Differential temperature adaptation in marine *Synechococcus* lineages: ecological distribution, molecular and physiological acclimation mechanisms

Deepa Ruth Varkey

This thesis is presented as a partial fulfilment to the requirements for a

Doctorate of Philosophy

Department of Chemistry and Biomolecular Sciences

Macquarie University, NSW, Australia

April 2016

Table of Contents

List of figures

List of tables

List of Supplementary files (in CD)

Chapter 2

SF 1 Label-free shotgun proteomics

SF S1 Distribution of log transformed NSAF values of strain WH8102 at 18°C and 22°C

SF S2 Distribution of log transformed NSAF values of strain BL107 at 18°C and 22°C

ST S1 Target genes and primers for qRT-PCR

ST S2 Total numbers of proteins and peptides identified in *Synechococcus* strains

ST S3 List of proteins differentially expressed at 18°C compared to 22°C in *Synechococcus* isolates

SF 3.1 Correlation of differential gene expression at the transcript and protein levels in *Synechococcus* strains at 26°C relative to 22°C

SI 3.1 Label-free shotgun proteomics

ST 3.1 Relative expression of proteins grouped into functional categories at 26°C and 22°C in *Synechococcus* strains

ST 3.2 List of hypothetical genes potentially involved in high temperature tolerance

ST 3.3 All differentially expressed genes in *Synechococcus* strain CC9311 at 26°C relative to 22°C

ST 3.4 All differentially expressed genes in *Synechococcus* strain BL107 at 26°C relative to 22°C

ST 3.5 All differentially expressed genes in *Synechococcus* strain WH8102 at 26°C relative to 22°C

ST 3.6 All differentially expressed genes in *Synechococcus* strain WH8109 at 26°C relative to 22°C

ST 4.1 Sequences of specific primers for *petB* phylogenetic marker and the attached Illumina sequencing adaptors and barcodes

SI 4.1 Commands used to process *petB* amplicon sequencing reads

ST 4.2 Matrix of community composition of Sydney Harbour estuary from *petB* marker

ST 4.3 Matrix of community composition of Sydney Harbour estuary from 16S rDNA marker

ST 4.4 Environmental parameters measured in the Sydney Harbour estuary

Declaration

I declare that the work in this thesis was conducted by me under the supervision of Prof. Ian T. Paulsen. The assistance and contribution of others have been appropriately acknowledged. The work presented here has not been previously submitted for any other degree.

Deepa Ruth Varkey

Acknowledgements

I thank my supervisor, Prof Ian Paulsen for his guidance and the opportunity to work as a part of his research team.

I thank my mentors, Sophie and Martin, for their incredible guidance, input, support and patience. I thank them for everything they have done for me.

Thank you to Sasha for all her help and support.

Thank you to Prof Paul Haynes for his help with the mass spectrometer and guidance with proteomics.

Thanks to all the other members of the Paulsen research group.

Lastly, I thank my friends and my wonderful family for being there when I needed them most.

Conference proceedings

Oral presentations

- Varkey D, Ostrowski M, Mazard S, Tetu S, Paulsen I. Exploring temperature adaptation and niche selection in marine cyanobacteria, *Synechococcus*. Sydney Micro, Sydney, Jan 2015.

- Varkey D, Ostrowski M, Mazard S, Tetu S, Paulsen I. Temperature niche selection in marine cyanobacteria, *Synechococcus*. Joint Academic Microbiology Seminars (JAMS), Sydney, July 2014.

Posters

- Varkey D, Ostrowski M, Mazard S, Tetu S, Haynes P, Paulsen I. Effect of low temperature on temperate and tropical isolates of marine *Synechococcus*. 9th European Workshop on Molecular Biology of Cyanobacteria, Netherlands, Sept 2014.

- Varkey D, Ostrowski M, Mazard S, Tetu S, Haynes P, Paulsen I. Temperature-related acclimation responses in isolates of marine *Synechococcus*. 4th Annual JAMS Symposium, Sydney, July 2014.

Achievements

Awarded the 'Best conference poster prize' at the 9th European Workshop on Molecular Biology of Cyanobacteria, Texel, The Netherlands, September 2014

Awarded the Postgraduate Research Fund ($5000 AUD) for conference travel from Macquarie University, Australia

Awarded the International Postgraduate Research Scholarship to undertake the Doctor of Philosophy degree

Contributions

Chapter 2:

Effects of low temperature on temperate and tropical isolates of marine *Synechococcus*

(Accepted for publication in the ISME Journal on 24 August 2015)

The work was conceptualised by Paulsen, Varkey, Ostrowski, Mazard and Tetu. All experimental work including design and troubleshooting was conducted by Varkey. The mass spectrometer was run by Haynes. All data analyses were performed by Varkey. The manuscript was written by Varkey with contributions from Ostrowski, Mazard and Paulsen.

Chapter 3:

Will a rise in temperature affect lineages of picocyanobacteria differently? – Growth physiology and cellular responses of marine *Synechococcus* isolates to high temperature

The work was conceptualised by Paulsen, Varkey, Ostrowski, Mazard and Tetu. All experimental work including design and troubleshooting was conducted by Varkey. The mass spectrometer was run by Varkey with support from the technical staff at the Australian Proteome Analysis Facility. All data analyses were performed by Varkey. The manuscript was written by Varkey with contributions from Ostrowski, Mazard and Paulsen.

Chapter 4:

Seasonal variation in marine cyanobacterial community in the Sydney Harbour estuary and the influence of temperature on community structure

The work was conceptualised by Varkey, Ostrowski, Jeffries and Paulsen. Samples from the Sydney Harbour estuary were collected by Jeffries. All experimental work with the *petB* amplicon library for

the Sydney Harbour estuary samples was conducted by Varkey. The 16S rDNA amplicon library preparation for the Sydney Harbour estuary was done by Jeffries and has been submitted for publication elsewhere. The *petB* amplicon library for the EAC/Port Hacking samples was prepared by Mazard. Bioinformatic and statistical data analyses were performed by Varkey with input from Ostrowski and Mazard. The manuscript was written by Varkey with significant contributions from Ostrowski, Mazard and Paulsen.

Chapter 5:

Microbial primary producers of Oceania: molecular characterisation of prokaryotic and eukaryotic communities across a range of habitats

Conceptualisation of the work on the Antarctic region was by Ostrowski and Armand. The Antarctic samples were collected by Armand on-board R.V. *Nathaniel B Palmer*. All experimental work with the 18S rDNA and *petB* amplicon libraries was conducted by Varkey, including DNA extraction. The 16S rDNA library preparation was done by Jason Woodhouse. Bioinformatic and statistical data analyses were performed by Varkey with input from Ostrowski and Mazard. The draft manuscript was written by Varkey and Mazard, with significant contributions from Ostrowski and essential input from Paulsen.

The work on the Sydney Harbour estuary was conceptualised by Ostrowski, Jeffries and Paulsen. The samples were collected by Jeffries. All experimental work with the 18S rDNA amplicon library was conducted by Varkey. Bioinformatic and statistical data analyses were performed by Varkey, Ostrowski and Mazard. The draft manuscript was written by Varkey and Mazard with significant contributions from Ostrowski.

Chapter 6:

Impact of DNA damaging agents on genome-wide transcriptional profiles in two marine *Synechococcus* species

The experimental work was carried out by Tetu, Johnson, Varkey, Phillippy, Stuart, Dupont and Hassan. Varkey performed growth curves, MIC analyses and PAM fluorometry. Data analyses were done by Tetu, Palenik, Paulsen and Varkey. The manuscript was written by Tetu, Palenik and Paulsen.

Abstract

Marine picocyanobacteria are abundant photosynthetic prokaryotes contributing significantly to global primary production and nutrient cycling. The genus *Synechococcus* is ubiquitous in the marine environment. Their habitat ranges from the polar regions to the equator and mesotrophic to oligotrophic environments. Such a widespread occurrence across a broad array of environmental conditions is facilitated by the diverse genetic complement of lineages of *Synechococcus*.

This work explores the influence of temperature on the lineages of *Synechococcus* which occupy different temperature niches. Responses and acclimation strategies employed by individual lineages were examined with gene expression analyses including global cellular proteomics and transcriptomics. Comparisons of growth physiology at different temperature conditions provide evidence for the specific temperature preferences of lineages. This is the first study to compare temperature acclimation responses across multiple lineages of marine *Synechococcus*. The acclimation responses to temperature involved the light harvesting complex, photosynthesis, membrane fluidity and protein synthesis with distinct differences between lineages.

This is the first study to determine the composition and structure of the *Synechococcus* community across different temperature regimes in regions around Australia were studied using various phylogenetic markers. Distinctive spatial and temporal partitioning of lineages is observed with temperature as a potential key factor shaping the population. Other co-varying factors such as nutrients and mixing were also determined to influence the partitioning of lineages.

This work provides insights into temperature acclimation and the distinct niche preferences of marine *Synechococcus* lineages. The differences in their distribution, growth and acclimation suggest that changes in temperature regimes can significantly alter *Synechococcus* community structure. Culture-based studies in addition to environmental distribution provide valuable information for predictive models. As significant contributors to primary production and biogeochemical cycling, it is important

to understand the influence of temperature and other factors on their diversity and distribution for better monitoring of ecosystem health.

Abbreviations

ABC-type	ATP-binding cassette type
ANOVA	analysis of variance
APC	allophycocyanin
ASW	artificial seawater
ATPase	adenylpyrophosphatase
ATP	adenosine triphosphate
bp	base pair
AQIS	Australian Quarantine and Inspection Service
Ca	carboxysome
CA	correspondence analysis
Cbp	chlorophyll-binding protein
CCA	canonical correspondence analyses
cDNA	complementary DNA
COG	clusters of orthologous genes
CPR	continuous plankton recorder
C_T	cycle threshold
CTD	conductivity, temperature and depth
Cyt	cytochrome
Dbp	DNA-binding protein
DNA	deoxyribonucleic acid
dsDNA	double-stranded DNA
dNTP	deoxynucleotide triphosphate
EAC	East Australian Current
ECH	East Central Harbour
EDGE-pro	estimated degree of gene expression in prokaryotes
EDTA	ethylenediaminetetraacetic acid
ETC	electron transport chain
FBPase/SBPase	fructose-1,6-biphosphatase/sedoheptulose-1,7-biphosphate phosphatase
FISH	fluorescence in-situ hybridisation
FTR	ferredoxin-thioredoxin reductase
g	Acceleration of gravity
g	grams

GOE	great oxidation event
GPM	Global Proteome Machine
Gyr	gigayears
HL	high light
Hli	high light inducible
HNLC	high nutrient, low chlorophyll
IPCC	Intergovernmental Panel on Climate Change
ITS	internal transcribed spacer
KCl	potassium chloride
km	kilometre
L	length
LC	Lane Cove
LL	low light
LPR	Lower Parramatta River
Lrc	rod-core linker polypeptide
m	metre
M	Mertz Glacier
Mbp	million base pairs
mg	milligram
MH	Middle Harbour
MHH	Marine/Harbour Heads
Min	minutes
ml	millilitre
MLD	mixed layer depth
mM	milliMolar
mRNA	messenger RNA
NADPH	nicotinamide adenine dinucleotide phosphate
ng	nanogram
nm	nanometre
NRS	national reference station
NSAF	normalized spectral abundance factor
NTR	NADPH-dependent thioredoxin reductase
OD	optical density
ODO	optical dissolved oxygen

OTU	operational taxonomic unit
PAGE	polyacrylamide gel electrophoresis
PAM	pulse amplitude modulation
PB	phycobilin
PBP	phycobiliprotein
PBS/PS	phycobilisomes
PC	phycocyanin
PCB	phycocyanobilin
PCR	polymerase chain reaction
PE	phycoerythrin
PEB	phycoerythrobilin
PH	Port Hacking
PON	particulate organic nitrogen
POP	particulate organic phosphorus
PPE	photosynthetic picoeukaryotes
ppt	parts per thousand
PR	Parramatta River
PS I	photoystem I
PS II	photosystem II
PUB	phycourobilin
qRT-PCR	quantitative reverse transcription PCR
Rbps	RNA-binding proteins
ROS	reactive oxygen species
rRNA	ribosomal RNA
RT	reverse transcription
RuBisCO	ribulose 1, 5 bisphosphate carboxylase/oxygenase
RV	research vessel
s	seconds
SDS	sodium dodecyl sulphate
SH	Sydney Harbour
sHSP	small heat shock protein
SO	Southern Ocean
SO-M	Southern Ocean to Mertz Glacier
SO-T	Southern Ocean to Totten Glacier

SOW	synthetic ocean water
SpC	spectral count
STX	Saxitoxin
T	Totten Glacier
TDN	total dissolved nitrogen
TDP	total dissolved phosphorus
TN	total nitrogen
TP	total phosphorus
UCYN A	uncultured cyanobacteria A
UCYN C	uncultured cyanobacteria C
UV	ultraviolet
V	volts
WCH	West Central Harbour
w/v	weight/volume
µm	micrometre
µg	microgram
µmol	micromoles

Chapter 1:

Introduction

1.1 Cyanobacteria

Cyanobacteria are ancient organisms with a long evolutionary history. Fossil evidence for the existence of autotrophic cyanobacteria dates back to more than 2.7 gigayears (Gyr) (Cavalier-Smith, 2006). Over their long evolutionary period, they have survived and adapted to a remarkable array of changing environmental conditions. Cyanobacteria are a large diverse group occurring in unicellular as well as filamentous forms, ubiquitous on the surface of the Earth. Their occurrence in diverse terrestrial, freshwater and marine habitats is attributed to their ability to tolerate a range of environmental conditions. Their success and abundance is evidenced by their global biomass estimated to be 10^{15} g wet biomass (Carr and Wyman, 1986; Schopf, 2012; Waterbury et al, 1986; Whitton and Potts, 2012).

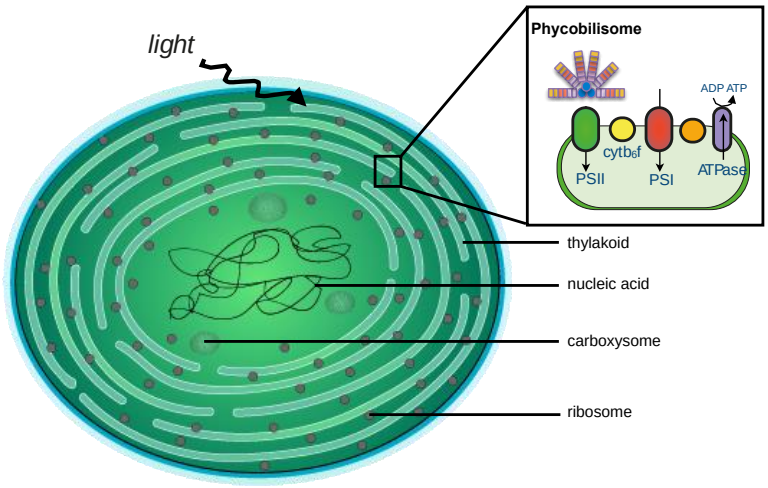


Figure 1.1: Schematic representation of a cyanobacterial cell showing the major constituents. The insert details components of the photosynthetic electron transport chain on the thylakoid membrane. This figure is modified from the Togo picture gallery maintained by Database Center for Life Science (DBCLS) (commons.wikimedia.org/wiki/File:Cyanobacteria.png).

Cyanobacteria are photosynthetic prokaryotes which utilise light as a source of energy for metabolism. Their photosynthetic apparatus, similar to chloroplasts in higher photosynthetic eukaryotes, comprises photosystem reaction centres bound to thylakoid membranes with the main photosynthetic pigment

being chlorophyll *a* (Figure 1.1). A distinguishing feature of most cyanobacteria is their accessory light-harvesting complex called phycobilisomes. These complexes contain the phycobiliproteins, phycocyanin and allophycocyanin, whilst some cyanobacteria produce an additional phycobiliprotein called phycoerythrin. This accessory light-harvesting complex is located on the outer surface of the thylakoid membrane. It absorbs and transfers light energy to the photosystem reaction centres (Carr and Wyman, 1986; Cohen-Bazire and Bryant, 1982). The array of pigment compositions within phycobilisomes provides cyanobacteria with the ability to take advantage of and successfully colonize a broad range of light conditions across the marine environment (Scanlan et al, 2009; Six et al, 2007). These pigment arrays distinguish cyanobacteria from the eukaryotic phototrophs.

The photosynthetic capacity, diversity, abundance and distribution make cyanobacteria key components of the ecosystem and thus important organisms to understand their biogeography and physiological adaptations.

1.2 Significance of cyanobacteria

Cyanobacteria occupy a pivotal position in Earth's evolutionary history. Geological and palaeontological records suggest that photosynthesis involving oxidation of water and production of oxygen first evolved in cyanobacteria (Giovannoni et al, 1988; Knoll, 2008). The Great Oxidation Event (GOE) approximately 2.4 Gyr ago is attributed to oxygenic photoautotrophy with cyanobacteria as the dominant oxygenic photoautotroph. This resulted in Earth's environmental evolution into an aerobic atmosphere and the partition of the biosphere into aerobic and anaerobic divisions (Cavalier-Smith, 2006; Giovannoni et al, 1988; Hayes, 1983). This in turn allowed the rapid expansion of multicellular and complex life-forms. The evolution of photosynthetic eukaryotes, such as algae and plants, is attributed to the engulfment and retention of a cyanobacterium by a eukaryotic cell (Giovannoni et al, 1988; Whitton and Potts, 2012; Waterbury et al, 1986; Carr and Wyman, 1986).

Cyanobacteria are key primary producers contributing to a significant proportion of global productivity. Their abundance and contribution to primary production make them one of the most

important organisms on Earth. At the base of the food web, cyanobacteria are the primary source of energy and organic matter sustaining other members of the ecosystem, thus contribute significantly to ecosystem structure and function. They play vital roles in biogeochemical cycling of nutrients including carbon, nitrogen and phosphorus. As photosynthetic organisms, cyanobacteria assimilate inorganic forms of carbon such as CO_2 and HCO_3^- thus making them a key component of the carbon cycle. Indeed, a major proportion of inorganic carbon assimilation is carried out by marine cyanobacteria (Raven, 2012; Whitton and Potts, 2012). Estimates of carbon fixation rates suggest that marine cyanobacteria contribute approximately 25% of the net primary productivity of oceans (Flombaum et al 2013). Several species of cyanobacteria such as *Trichodesmium*, *Nostoc* and *Anabaena* also have the ability to fix atmospheric nitrogen, contributing significantly to the so-called input of "new nitrogen" or biologically fixed nitrogen, and maintaining the balance of this key nutrient, particularly in the marine environment (Stahl, 2008; Zehr, 2011). Their roles in carbon and nitrogen fixation as well as other biogeochemical cycles make them indispensable for proper ecosystem functioning.

In addition to oxygen production and nutrient cycling, certain species of cyanobacteria have harmful effects on other organisms including humans. Some secondary metabolites produced by cyanobacteria are toxic (including cytotoxins, hepatotoxins, neurotoxins and dermatotoxins) and are particularly harmful when they occur in blooms, both in freshwater and marine environments (Dittmann and Wiegand 2006). Determining the factors that influence bloom formation and toxin production are important for monitoring ecosystem health.

Thus cyanobacteria are key organisms that influence various ecosystem functions. It is therefore important to understand their distribution and physiological adaptations.

1.3 Classification of cyanobacteria

Cyanobacteria are broadly classified into 5 subsections – Chroococcales, Pleurocapsales, Oscillatoriales, Nostocales and Stigonemetales. Their classification is based on several characteristics

including morphology, physiology, habitat as well as genetic content (Herdman et al, 2001). Unicellular cyanobacteria occurring in solitary or colonial forms are classified as Chroococcales (Subsection I). Cyanobacteria capable of binary and multiple fission with forms ranging from unicellular to pseudo-filamentous (thallus-forming) are grouped as Pleurocapsales (Subsection II). The group of Oscillatoriales (Subsection III) is composed of filamentous cyanobacteria with non-differentiated cells. Nostocales (Subsection IV) consists of filamentous forms with cellular differentiation whilst Stigonematales (Subsection V) has the most complex organisation among cyanobacteria with cell differentiation and multicellular forms (Knoll, 2008).

1.4 Unicellular cyanobacteria

Unicellular cyanobacteria belong to subsections I (Chroococales) and II (Pleurocapsales). They are oxygenic phototrophic prokaryotes with spherical or rod-shaped cells, and some genera occurring as colonial forms (Figure 1.2). The two sub-groups are differentiated based on the mode of reproduction with cyanobacteria of subsection I reproducing through binary fission or budding whilst subsection II through multiple fission or both binary and multiple fission (Callieri et al, 2012; Carr and Wyman, 1986; Herdman et al, 2001).

These unicellular cyanobacteria are key primary producers found in a broad range of habitats, from marine and freshwater ecosystems to specialised and extreme environments, contributing significantly to ecosystem services.

1.4.1 *Freshwater habitats*

Unicellular cyanobacteria are major components of aquatic systems across the Earth. Their distribution is widespread across freshwater systems such as lakes and other non-marine aquatic ecosystems in tropical, temperate and polar environments. They occur in a wide variety of trophic regimes. Some freshwater cyanobacteria are halotolerant with the ability to withstand high salinity.

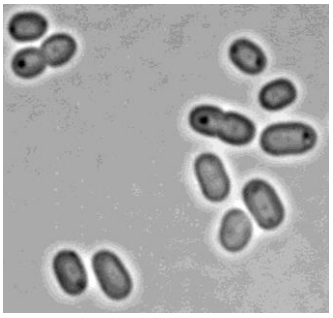
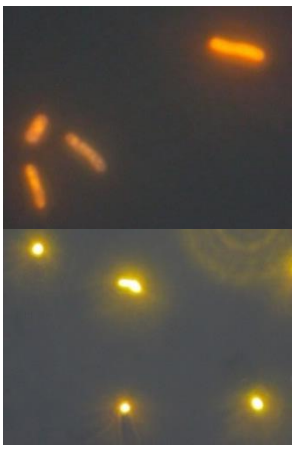
a)  b)  c) 

Figure 1.2: Microscopy images of unicellular cyanobacteria: a) euryhaline *Synechococcus* strain PCC 7002 (100X; commons.wikimedia.org/wiki/File:Synechococcus_PCC_7002_BF.jpg); b) marine *Synechococcus* sp. (40X; image by Dr Mazard; c) colonial *Chroococcus* sp. (800X; image reprinted with the permission of Dr Oyadomari; http://www.keweenawalgae.mtu.edu).

Environmental parameters of abiotic and biotic nature influence the diversity, distribution and abundance of the different genera. Freshwater cyanobacteria have been found to occur in high abundances particularly in eutrophic lakes and are known to produce a variety of toxins which can be harmful to humans. The diversity of these freshwater cyanobacteria is rich with numerous morphological types ranging from unicellular to colonial forms (comprising 5 to 50 cells). Unicellular forms belong to genera *Cyanobium*, *Synechococcus* and *Cyanothece* whilst colonial forms include *Aphanocapsa*, *Aphanothece* and *Chroococcus* amongst other genera (Callieri and Stockner, 2002; Callieri et al, 2012).

1.4.2 *Marine habitats*

Unicellular marine cyanobacteria are the most abundant photosynthetic organisms on Earth dominating vast areas of the marine photic zone. Some unicellular cyanobacteria such as *Cyanothece*, UCYN-A, UCYN-C and *Crocosphaera* have nitrogen-fixing capability and contribute significantly to nitrogen cycling (Stahl, 2008; Zehr, 2011). Picocyanobacteria of the genera *Synechococcus* and *Prochlorococcus* are highly abundant unicellular marine cyanobacteria. They are key primary

producers particularly in oligotrophic regions where nutrient availability is scarce, carrying out up to 25% of the global carbon fixation. Their biomass and significant contribution to global primary production, oxygen evolution and biogeochemical nutrient cycling make them important members of the marine community (Garczarek et al, 2008; Jardillier et al, 2010; Li, 1994; Palenik et al, 2009; Partensky and Vaulot, 1999; Scanlan et al, 2009).

Marine picocyanobacteria are small, unicellular and coccoid/rod-shaped. They are halophilic and have an obligate requirement for high salt concentrations (Carr and Wyman, 1986). The marine picocyanobacteria mainly comprise the two genera *Prochlorococcus* and *Synechococcus*. The former was first reported in 1979 by Waterbury et al and Johnson and Sieburth, while the latter was reported nearly a decade later by Chisholm et al (1988).

Prochlorococcus is the smallest (less than 0.8 µm) and the most abundant photosynthetic organism on Earth (Partensky et al, 1999). Ranging from 1.9 – 2.0 Mbp, *Prochlorococcus* has the smallest and highly streamlined genome among photosynthetic prokaryotes. They occur at high density up to 10^6 cells ml^{-1} and can be found from the surface to beyond the euphotic zone at depths down to 200 m (Partensky et al, 1999). The light harvesting complex in *Prochlorococcus* is a chlorophyll-based antenna complex utilising divinyl chlorophyll *a* and *b* as the main photosynthetic pigments. These pigments, unique to *Prochlorococcus*, enable optimum absorption of blue light (460-480 nm) prevalent in their open ocean habitat and at depth, since the shorter wavelength of blue light can reach deeper down the water column (Chisholm et al, 1988; Goericke and Repeta, 1992; Partensky et al, 1999).

Marine *Synechococcus* cells are mostly larger than *Prochlorococcus* with sizes ranging from 0.6 to 1.6 µm. Maximum cell numbers of *Synechococcus* range from 10^3 cells ml^{-1} in oligotrophic regions to 10^5-10^6 cells ml^{-1} in nutrient rich waters (Saito et al, 2005; Zwirglmaier et al, 2008). The *Synechococcus* genome consists of a single chromosome with size ranging from 2.2 to approximately 2.86 Mbp. The detection and identification of *Synechococcus* were based on their auto-fluorescence

detected by epifluorescence microscopy. The main distinguishable feature of *Synechococcus* is the orange fluorescence emitted by phycoerythrin which is their major light harvesting pigment (Partensky et al, 1999; Stockner, 1988; Waterbury et al, 1979; Waterbury et al, 1986). The light harvesting complex, phycobilisomes, comprises polypeptides bound to bilin chromophores such as phycocyanin, phycoerythrin and phycourobilin. The genus is a phylogenetically diverse group which is abundant across the euphotic zone ranging from coastal to open ocean waters. *Synechococcus* cells occur in higher abundance close to the surface but are also found down to depths of 200 m, albeit at very low numbers (Urbach et al, 1998; Waterbury et al, 1979).

1.4.3 *Specialised and extreme environments*

Representatives of cyanobacteria including *Synechococcus*, *Synechocystis* and *Cyanothece* were shown to have successfully colonised a variety of specialised or extreme habitats including hot springs, cryosphere, hypersaline lakes, subaerial surfaces, rock/cliff faces and desert soil crusts (Figure 1.3). This wide diverse environmental distribution implies the development of mechanisms to cope with extreme conditions of temperature, pH, sulphide, water stress and UV irradiance. Some unicellular cyanobacteria in these environments such as *Gloeothece* and *Synechococcus* are nitrogen-fixers (Pentecost and Whitton, 2012; Stahl, 2008; Ward et al, 2012). Other specialised environments in which unicellular cyanobacteria occur are as symbionts of a variety of organisms including algae, fungi, protists, sponges and others wherein they provide the hosts with fixed carbon and nitrogen in return for a stable and protected environment within the host (Adams et al, 2012; Foster et al, 2011; Hoffman, 1999).

The occurrence of unicellular cyanobacteria across such an array of environments including extreme and specialised habitats illustrates the remarkable diversity and adaptability with which these organisms have flourished on the Earth for several billion years.

a). 

b). 

Figure 1.3: Specialised environments in which cyanobacteria may occur, a) Grand Prismatic Spring (hot spring), Yellowstone National Park (image by James St. John, Ohio–Flickr; commons.wikimedia.org/wiki/File:Grand_Prismatic_Spring_2013.jpg; licensed under CC BY 2.0 via Commons), and b) stromatolite (image by Didier Descouens; licensed under CC BY-SA 3.0 via Commons; commons.wikimedia.org/wiki/File:Stromatolithe_Pal%C3%A9oarch%C3%A9en_-_MNHT.PAL.2009.10.1.jpg).

1.5 Phylogenetic diversity in marine picocyanobacteria

The marine picocyanobacterial genera *Prochlorococcus* and *Synechococcus* are a diverse group comprising several lineages (Figure 1.4). The genus *Prochlorococcus* was originally grouped with other chlorophyll *b*-producing cyanobacteria but phylogeny based on 16S rRNA suggests that marine

Synechococcus and *Prochlorococcus* are closely related. The marine picophytoplankton clade diverged more recently from a common ancestor with the *Cyanobium*-cluster containing freshwater *Synechococcus* (Urbach et al, 1998). There are two main ecotypes of *Prochlorococcus*, high-light adapted (HL) and low-light adapted (LL), with the former more recently evolved based on 16S rRNA phylogeny (Moore et al, 1998; Rocap et al, 2003; Urbach et al, 1998). These ecotypes have specific physiology and genetic makeup which enable them to occupy distinct niches down the water column (Bouman et al, 2006). The most widespread HL and LL ecotypes are HLI, HLII and LLI – LLIV. The LL ecotypes are genetically more diverse than HL ecotypes (Scanlan et al, 2009; Zinser et al, 2006) and additional lineages of *Prochlorococcus* are being detected, including NC1, HNLC, HLIII - HLVI and LLV – LLVI (Huang et al, 2012; Martiny et al, 2009; Lavin et al, 2010; Rocap et al, 2002; West et al, 2011).

The genus *Synechococcus* is a phylogenetically diverse group comprising both freshwater and marine lineages (Waterbury et al, 1986; Scanlan, 2012). The marine *Synechococcus* lineage was originally classified into marine clusters A, B and C based on isolation site characteristics, light-harvesting pigment composition, motility, salt requirement and G+C nucleotide content (Waterbury et al, 1989; Fuller et al, 2003; Herdman et al, 2001). Under the current classification, strains of marine *Synechococcus* are grouped together as cluster 5 which comprises three sub-clusters 5.1, 5.2 and 5.3. Sub-cluster 5.1 is the largest and most diverse consisting of strains with elevated salt requirement and mainly phycoerythrin as a major pigment in phycobilisomes. Coastal halotolerant strains with phycobilisomes containing phycocyanin are grouped together as sub-cluster 5.2. These strains do not possess phycoerythrin. Members of sub-cluster 5.3 were previously classified as clade X of sub-cluster 5.1, however, phylogenetic analyses suggest their divergence prior to the differentiation of other *Synechococcus* sub-clusters and *Prochlorococcus*. Whilst sub-clusters 5.2 and 5.3 are less diverse, 5.1 has a high diversity and together these sub-clusters are further classified into more than twenty phylogenetically distinct lineages (Ahlgren and Rocap, 2012; Dufresne et al, 2008; Fuller et al, 2003; Herdman et al, 2001; Scanlan, 2012; Urbach et al, 1998).

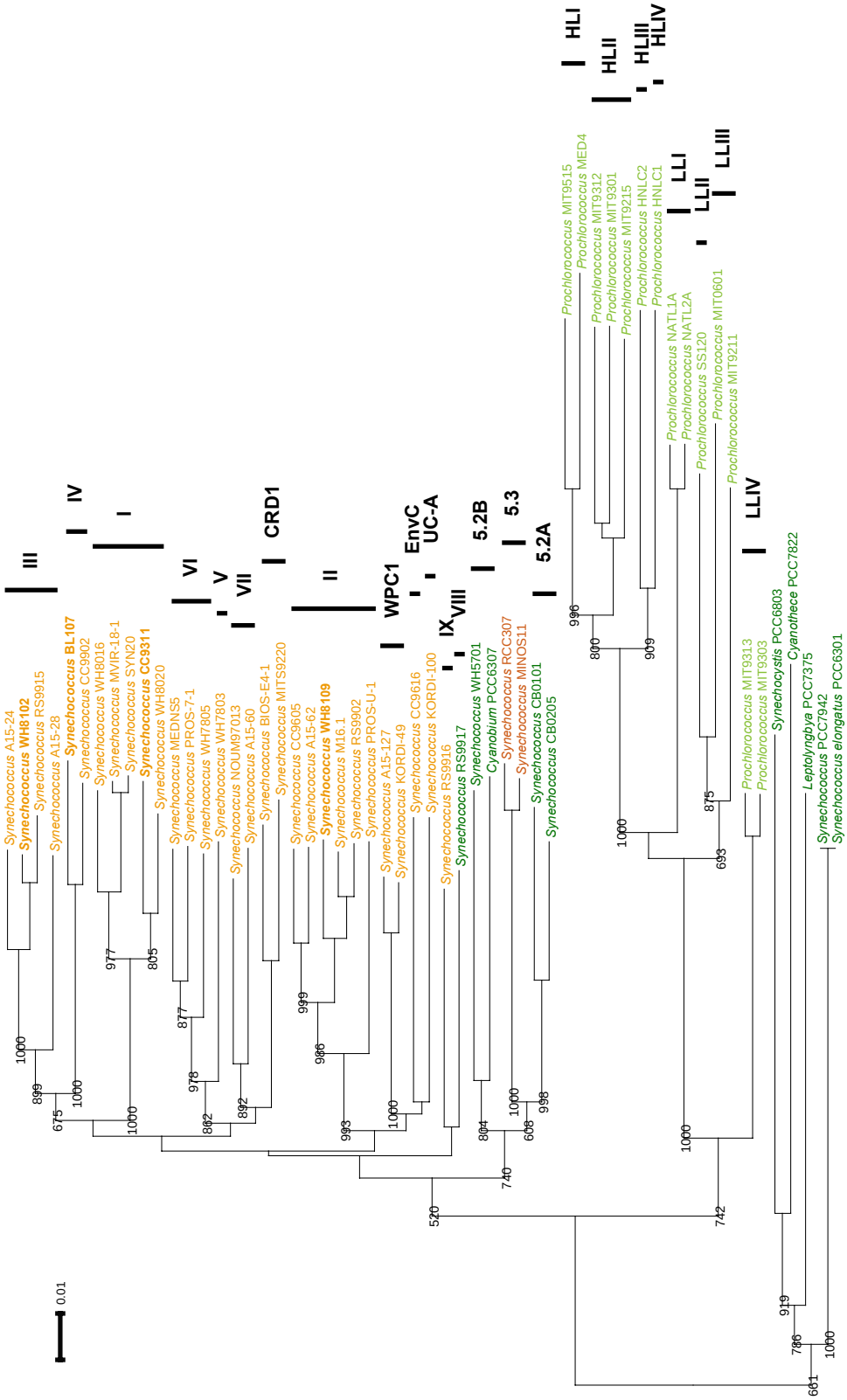


Figure 1.4: Phylogenetic tree of marine picocyanobacteria based on the nucleotide sequences of gene *petB*, encoding cytochrome b_6 subunit of cytochrome b_6f complex (Mazard et al, 2012). The tree was constructed using the Neighbor-Joining algorithm with 1000 repetitions for bootstrap values in ClustalW (Thompson et al, 1994).

Though closely related, *Synechococcus* subgroups have a wide distribution across the marine environment with each adapted to specific niches. The functional diversity and characteristics that enable such niche specialisation warrants further investigation.

1.6 Distribution of picocyanobacteria in the marine environment

The marine environment comprises habitats ranging from brackish estuarine to open ocean waters. Picocyanobacteria are distributed ubiquitously across the marine environment (Figure 1.5; Table 1.1). Their spatial partitioning is dependent on the prevalent environmental conditions, with lineages displaying niche preferences enabled by genetic, and therefore physiological differences. The genetic diversity in picocyanobacteria has likely been influenced by horizontal gene transfer mediated by phages and possibly plasmids (Dufresne et al, 2008; Scanlan et al, 2009). *Prochlorococcus* appears to be mainly influenced by light, temperature and water column stability whilst for *Synechococcus* nutrients and temperature have a significant role in shaping their distribution. In regions where the two genera co-occur, *Prochlorococcus* dominates stratified nutrient-poor (oligotrophic) regions whilst *Synechococcus* is more abundant in the mesotrophic and nutrient-rich well-mixed waters (Scanlan, 2012).

Prochlorococcus is dominant in the warm subtropical/tropical oligotrophic regions within the 40°N to 40°S latitudinal range (Campbell et al, 1994; Partensky et al, 1999). They occur at high abundances of more than 10^5 cells ml^{-1} with numbers reducing outside of that latitudinal range (Partensky and Vaulot, 1999). *Prochlorococcus* lineages spatially partition vertically down the water column along the light gradient with high concentrations of cells found down to depths of 200 m (Garczarek et al, 2007; Moore et al, 1998; Partensky and Vaulot, 1999; Scanlan et al, 2009). HL ecotypes are predominant in surface waters while LL ecotypes preferentially occur further down the water column. Latitude preference is also observed among HL ecotypes wherein HLI occurs above and below 30° and HLII within that latitude range (Bouman et al, 2006; Fuller et al, 2006; Johnson et al, 2006; Moore et al, 1998; Scanlan et al, 2009).

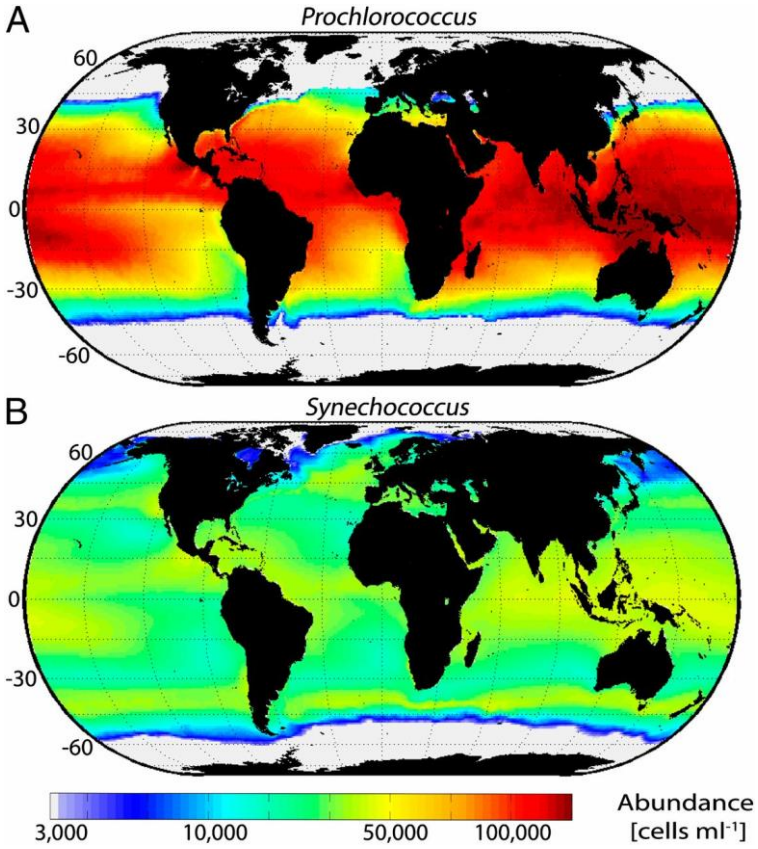


Figure 1.5: Global distribution of a) *Prochlorococcus* and b) *Synechococcus* depicting mean annual abundance. This figure is reprinted from Flombaum et al (2013), PNAS (copyright 2013, National Academy of Sciences, USA).

Synechococcus is abundant in the euphotic zone with higher abundance of cells close to the surface and very low numbers found down to depths of 200 m (Waterbury et al, 1979). *Synechococcus* lineages partition along the spatial scale with their numbers varying by several orders of magnitude between coastal and open ocean waters (Partensky et al, 1999; Waterbury et al, 1986; Scanlan et al, 2009; Zwirglmaier et al, 2008). The broad distribution from the polar to the equatorial regions is attributed to the genetic diversity in *Synechococcus* lineages (Zwirglmaier et al, 2008). This diversity potentially enables specific lineages to respond to environmental conditions and withstand the selective pressures to varying degrees (Mazard et al, 2012). Clades I and IV are predominant in coastal regions in latitudes above 30°N and below 30°S. Clade II is most prevalent in coastal and/or continental shelf regions in latitudes bordered by 30°N and 30°S. More recent studies (Sohm et al,

Table 1.1: Lineages of *Prochlorococcus* and *Synechococcus* and their distribution in the marine environment (adapted from Ahlgren and Rocap, 2012 and Huang et al, 2012).

Clade	Biogeography
Prochlorococcus	
HLI	Temperate, open ocean regions; upper to mid euphotic zone
HLII	Tropical, open ocean regions; upper to mid euphotic zone in stratified waters
HLIII/HNLC1, HLIV/HNLC2, HLV	Equatorial/tropical, iron-depleted waters; upper to mid euphotic zone
HLVI	Mid to lower euphotic zone
LLI	Temperate; mid to lower euphotic zone in stratified waters and entire euphotic zone in well-mixed waters
LLII	Lower euphotic zone
LLIII	Lower euphotic zone
LLIV	Tropical/sub-tropical; lower euphotic zone
LLV, LLVI	Tropical regions in intermediate depth in oxygen-minimum waters
NC1	Lower euphotic zone
Synechococcus	
Sub-cluster 5.1	
I	Temperate, coastal mesotrophic waters
II	Tropical/sub-tropical, coastal/open-ocean/oligotrophic waters
III	Tropical/sub-tropical oligotrophic waters
IV	Temperate, coastal mesotrophic waters
V/VI/VII	Wide distribution in low abundance; abundant in some upwelling regions
VIII	Euryhaline regions
IX	Detected at low abundance in the Red Sea and East China Sea
XV, XVI	Low abundance in transitional sites, well-mixed waters and sub-surface waters
CRD1	Tropical and sub-tropical waters, upwelling regions and HNLC provinces
EnvA, EnvB, EnvC	Transitional regions in the Atlantic Ocean
XVII, XVIII, XIX	Low abundance in South China Sea and equatorial Pacific
CRD2	Nearshore and oligotrophic regions, upwelling regions and HNLC provinces
CB1	Estuarine and coastal waters, polar and sub-polar regions
CB3	Chesapeake Bay and East China Sea
WPC1	Open ocean and nearshore waters
Sub-cluster 5.2	
CB4	Temperate, coastal waters
CB5	Temperate, estuarine/coastal waters and polar/sub-polar regions
Sub-cluster 5.3	
X/ 5.3-I	Tropical/ sub-tropical, off-shore and oligotrophic open-ocean waters
5.3-II – 5.3-VI	Sargasso Sea

2015) have reported clade II as the dominant lineage in warm, open ocean waters. Another *Synechococcus* clade primarily found in oligotrophic open ocean waters is clade III (Fuller et al, 2003; Huang et al, 2012; Scanlan, 2012; Sohm et al, 2015; Zwirglmaier et al, 2007, 2008). Other lineages of *Synechococcus* have been detected either in low abundance or as the predominant lineage in certain regions (Table 1.1), however, determining their global biogeography requires further oceanic surveys. *Synechococcus* sub-clusters 5.2 and 5.3 are mostly found in estuarine environments. Their global distribution is less known (Ahlgren and Rocap, 2006, 2012; Cai et al, 2010; Choi and Noh, 2009; Huang et al, 2012; Penno et al, 2006; Saito et al, 2005; Scanlan et al, 2009).

Though the occurrence of marine picocyanobacteria is well understood across several regions in the marine environment (namely Atlantic and Pacific oceans), attaining a global biogeographical perspective requires further surveys particularly in regions of the southern hemisphere such as the Southern Ocean.

1.7 Factors that influence the distribution of picocyanobacteria

The oceans are diverse and dynamic environments. Marine ecosystems vary in physical, chemical and biological characteristics across spatial and temporal scales. The occurrence and distribution of an organism within an ecosystem is dependent on its ability to cope with the environmental conditions prevalent in its ecological niche. A marine ecosystem is no exception. For a photosynthetic organism within the marine environment, limiting factors include light and nutrient availability. Other contributing factors include temperature, water column stability, competition and predators (Partensky and Vaulot, 1999; Scanlan et al, 2009; Sohm et al, 2015; Zwirglmaier et al, 2008).

1.7.1 *Irradiance*

As photosynthetic organisms, picocyanobacteria are dependent on quantity and quality of available light energy. The light environment is dependent on water turbidity as well as the depth of occurrence

of the organism. Picocyanobacteria distributed across different light regimes have specific adaptations to their light harvesting machinery.

Along the vertical gradient of their distribution, light is an important limiting factor (Partensky and Vaulot, 1999). The partitioning of the two main ecotypes of *Prochlorococcus* is primarily influenced by light conditions. These ecotypes have distinct physiology and genetic makeup which enable them to occupy distinct niches down the water column (Bouman et al, 2006; Fuller et al, 2006; Moore et al, 1998). *Prochlorococcus* has the ability to grow at irradiances that range over several orders of magnitude with high- and low-light adapted ecotypes varying in the growth optima for irradiance (Partensky et al, 1999; Moore et al, 1995). One of the adaptations distinguishing the two ecotypes of *Prochlorococcus* is the ratio of divinyl-chlorophyll *a* to divinyl-chlorophyll *b*. The ratio is higher in LL ecotypes enabling them to absorb efficiently at low irradiance (Moore et al, 1995; Scanlan, 2012). Differences in the number of genes that encode divinyl-chlorophyll *a/b*– binding proteins may also be involved in their light adaptation ability. Some were also found to have a type of phycobiliprotein similar to phycoerythrin which could contribute to their light adaptability (Garczarek et al, 2007).

Synechococcus is more abundant in surface waters with numbers reducing down the water column. Below the euphotic zone their numbers are extremely low (Partensky and Vaulot, 1999). However, the occurrence of *Synechococcus* cells further down the water column (up to 200 m) is observed in well mixed waters (clade II in Red Sea) (Zwirglmaier et al, 2008). *Synechococcus* lineages vary in their phycobilisome composition and the ratio of the chromophores, phycourobilin (PUB) and phycoerythrobilin (PEB). This difference in pigment type is reflective of their distribution. The trophic conditions, turbidity and level of mixing change the transmittance of light wavelengths affecting both the intensity and quality of available light. The presence/absence of PEB and the PUB to PEB ratio determines the wavelength of maximal absorption. The dominant strain in estuarine, coastal, oligotrophic and mesotrophic waters comprise pigment types enabling optimal absorption of the incident wavelength. Strains dominant in oligotrophic waters have high PUB to PEB ratio since PUB has maximum absorption at 475 nm which is the transmitted wavelength. Higher PEB is found in

those strains occupying mesotrophic waters where the transmitted wavelength is 525 – 550 nm (maximum absorption range of PEB) (Scanlan, 2012; Six et al, 2007).

1.7.2 *Temperature*

Ocean temperature is an important environmental factor with geographical, seasonal and temporal variation (Figure 1.6). Temperature can influence chemical and biochemical processes and thus, can affect the growth, physiology and metabolism of an organism. Adaptability to temperature conditions can define habitat ranges. With a wide distribution across the ocean from the poles to the equator, marine picocyanobacteria are exposed to a range of temperatures. Fluctuations in temperature are largest at higher latitudes and least at the equator.

Prochlorococcus appears to be sensitive to low temperatures, preferentially occurring in warmer tropical latitudes and cell numbers decreasing drastically above 45°N and below 40°S. Temperatures below 10°C are found to be inhibiting to *Prochlorococcus* (Partensky and Vaulot, 1999; Zubkov et al, 2000). The upper temperature limit for *Prochlorococcus* appears to vary from 25°C for individual cultures grown to higher than 29°C in the field (Partensky and Vaulot, 1999; Moore et al, 1995). Temperature optima vary among the *Prochlorococcus* ecotypes. Latitude preference of HL ecotypes may be correlated with temperature preferences (Johnson et al, 2006; Scanlan et al, 2009). HLI is more tolerant of temperatures below 15°C whereas HLII shows preference (in terms of maximum abundance and growth rate) for higher temperatures of approximately 25°C (Johnson et al, 2006; Zinser et al, 2007).

Synechococcus occur over a larger range of temperature from 0°C to 30°C with individual strains showing different growth temperature optima (Partensky and Vaulot, 1999). Cells have been detected at temperatures below 0°C albeit at low numbers (10 cells ml^{-1}) (Cottrell and Kirchman, 2009; Scanlan, 2012). *Synechococcus* lineages change from tropical to temperate latitudes and temperature could influence their distribution. The distribution of clades I and IV is predominantly temperate, mostly confined to latitudes above 30°N and below 30°S (Scanlan, 2012; Sohm et al, 2015;

Zwirglmaier et al, 2008). Higher abundance of clades I and IV are found at a lower temperature range in comparison to other clades such as clade II which occur at higher abundance at temperatures ranging from 22°C to 28°C. Clade II and III are the dominant lineages in warm, open-ocean regions within the 30°N and 30°S latitudes. At higher latitudes clade II is found at very low abundance (Sohm et al, 2015; Zwirglmaier et al, 2008).

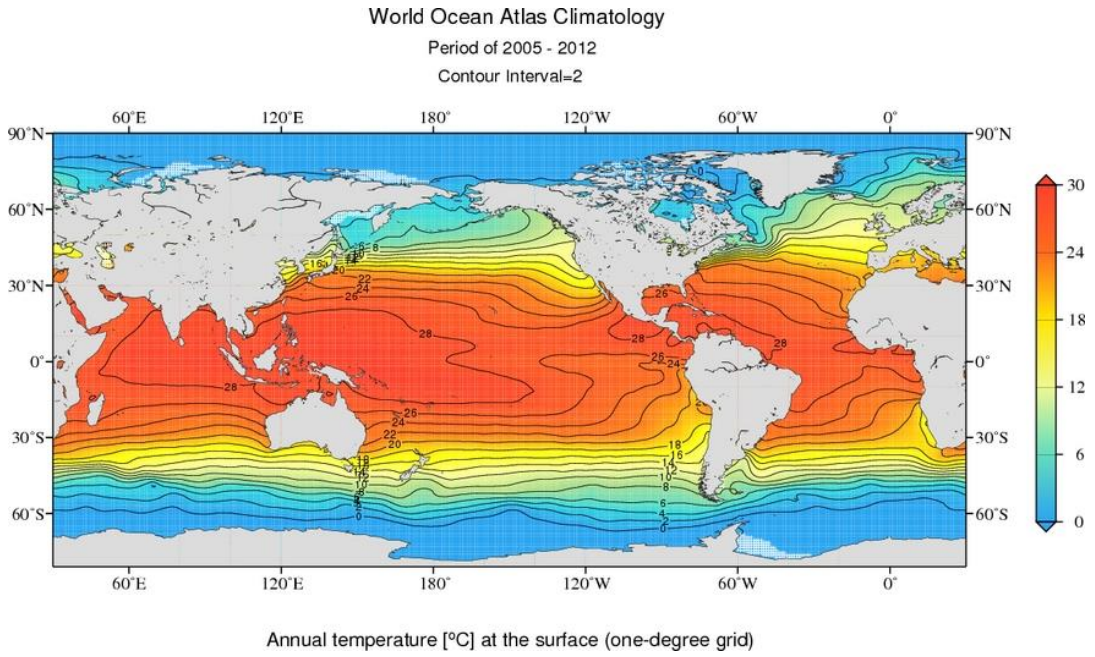


Figure 1.6: Average annual sea surface temperature based on satellite data (World Ocean Atlas 2013). This map was reproduced from National Centers for Environmental Information: National Oceanic and Atmospheric Administration (www.nodc.noaa.gov).

Recent studies by Mackey et al (2013) and Pittera et al (2014) have reported differences in temperature tolerance among cultured *Synechococcus* isolates which reflect their biogeographical distribution. These studies highlight differences in photophysiology and expression of photosynthetic genes that could contribute to temperature preferences among isolates. The work undertaken in this thesis further uses multiple model strains and explores the physiology as well as gene expression differences at the whole cellular level which could influence temperature preference and niche specialisation.

1.7.3 *Water column stability*

The level of stratification and mixing of water column influences the light and nutrient availability. Water column stability can vary geographically as well as seasonally. Stability of the water column appears to affect the distribution of marine picocyanobacteria. *Prochlorococcus* tends to prefer stratified water columns with their numbers reducing in well mixed waters and *Synechococcus* abundance become higher. Mixing due to physical forcing of the water column changes nutrient concentrations (particularly nitrate) which favour higher growth rates in *Synechococcus* (Scanlan, 2012).

Among the *Prochlorococcus* ecotypes, HLII appears to prefer stratified water columns whilst HLI occurs in less stable conditions (Bouman et al, 2006; Partensky and Vaulot, 1999; Zwirglmaier et al, 2008). Exceptions to the partitioning of HL and LL ecotypes have been observed particularly in mixed water columns where they have been reported to co-occur. *Prochlorococcus* population structure varies in response to the changed light and nutrient conditions found in mixed water columns (Bouman et al, 2006). *Synechococcus* clades V/VI/VII, as a group, have been detected at low abundances over large areas of the oceanic environment. In certain regions, however, these lineages appear to abound. These are regions of upwelling with nutrient rich waters. The oligonucleotide probe (SYN1280) used to detect these clades targets the 16S rRNA gene but does not provide sufficient resolution to differentiate the individual clades and therefore distinguishing their specific preferred niches is limited (Zwirglmaier et al, 2008).

1.7.4 *Nutrients*

Macronutrients such as nitrogen and phosphorus vary in their concentration and bioavailability across oceanic regions (Figure 1.7) and are important limiting factors to picocyanobacterial growth. In addition to macronutrients, trace metals such as iron, copper, nickel, manganese and zinc among others can be limiting to growth. Significant regions of the oceans are nutrient-scarce oligotrophic environments. In such oligotrophic open ocean regions where picocyanobacteria occur in high

abundance, surface concentrations of bioavailable nitrogen and phosphorus can be at nanomolar levels and is rapidly turned over. Marine picocyanobacteria utilise distinct strategies to cope with low and fluctuating concentrations of bioavailable nutrients. These adaptive strategies include altered elemental stoichiometry, reduced nutrient requirements, regulatory systems as well as high affinity binding proteins and transporters. There are variations in the strategies employed by different lineages of picocyanobacteria which reflects on differences in specific nutrient niches (Ammerman et al, 2003; Morel and Price 2003; Ostrowski et al, 2010; Palenik, 2015).

Prochlorococcus thrives in nutrient-scarce regions. Their high surface to volume ratio and low nutrient requirement make them well adapted to oligotrophic waters (Partensky et al, 1999). Though *Prochlorococcus* has been detected in nutrient-replete waters their cell concentrations do not vary with increase in nutrients. In eutrophic regions they do not occur in detectable numbers (Partensky and Vaulot, 1999). The availability of different nitrogen and phosphorus sources changes down the water column. *Prochlorococcus* ecotypes appear to be adapted to utilise the different available nutrient sources (Bouman et al, 2006).

Synechococcus is more abundant than *Prochlorococcus* in nutrient-replete oceanic regions. From mesotrophic to oligotrophic environments, the dominant *Synechococcus* lineages vary with the prevalent trophic conditions. The variations in the genomic content of lineages correlate with the sources and concentrations of nutrients prevailing in their specific niche. Clades I and IV which co-occur in several oceanic regions including Arctic Ocean, South Pacific Ocean and Mediterranean Sea are mostly restricted to coastal waters (Mella-Flores et al, 2011; Scanlan et al, 2009; Zwirglmaier et al, 2008). In their coastal habitat, clades I and IV are exposed to a range of nutrient concentrations. Clade I shows adaptations for a highly variable coastal habitat (Zwirglmaier et al, 2008; Palenik et al, 2006; Scanlan et al, 2009). Clade III appears to be strictly oligotrophic and well adapted to the stable narrow range of macronutrient concentrations. Adaptations include multiple nutrient transporters and low regulatory components (Zwirglmaier et al, 2008; Palenik et al, 2003; Scanlan et al, 2009; Sohm et al, 2015).

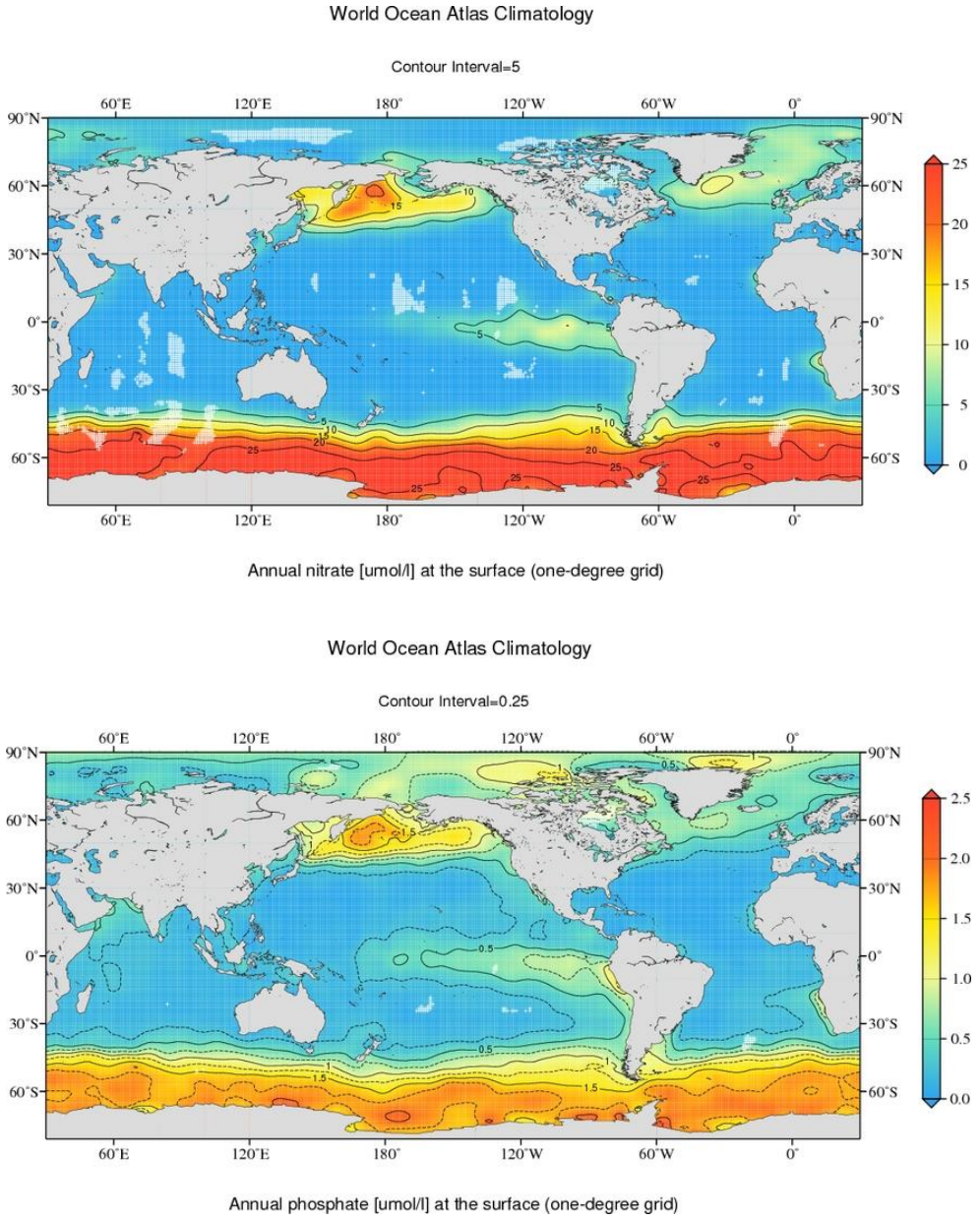


Figure 1.7: Mean annual surface distribution of nutrients (µmol l^{-1}) a) Nitrate and b) Phosphate, based on satellite data (World Ocean Atlas 2013). This map was reproduced from National Centers for Environmental Information: National Oceanic and Atmospheric Administration (www.nodc.noaa.gov).

1.7.5 *Biotic factors*

Viral communities are abundant and dynamic, and have the potential to influence their host population. Viruses that infect cyanobacteria, cyanophages, are ubiquitous and diverse in the marine

environment displaying both narrow and broad host ranges. Their distributions correlate with cyanobacterial communities with variations in their community composition across different marine habitats. Lysis as a result of cyanophage infection contributes significantly to mortality rates in picocyanobacterial populations. Phage specificity and host susceptibility can in turn influence community composition of picocyanobacteria. Viruses are considered to be large reservoirs of genetic diversity and cyanophages are no exception. Since they carry host genetic material, they are suggested to be mediators of lateral gene transfer amongst cyanobacteria. Therefore cyanophages potentially generate and maintain genetic diversity amongst marine picocyanobacteria (Hambly and Suttle, 2005; Huang et al, 2015; Mühling et al, 2005; Partensky and Vaulot, 1999; Suttle and Chan, 1994; Zhong et al, 2002).

Grazing is another biotic factor contributing to the control of picocyanobacterial populations. *Synechococcus* and *Prochlorococcus* are susceptible to grazing by protists such as heterotrophic nanoflagellates, ciliates and dinoflagellates. Picocyanobacteria exhibit differences in their susceptibility towards protist grazing and vary in predator-associated mortality rates. Preferential grazing of different *Synechococcus* isolates has been attributed to differences in cell surface properties and grazer evasion strategies (Scanlan, 2012; Strom et al, 2012; Zwirglmaier et al, 2009).

Another important biotic factor influencing the distribution of picocyanobacteria is competitive pressure, both within the picocyanobacterial groups as well as with other competing groups such as heterotrophic bacteria and larger photosynthetic organisms. The competitive pressure for bioavailable nutrients is likely high in oligotrophic open ocean environments. Both *Synechococcus* and *Prochlorococcus* have adopted strategies for nutrient acquisition, utilisation and storage which enable them to cope with the competitive pressure. In oligotrophic regions where the two picocyanobacterial genera co-occur, *Prochlorococcus* has a competitive advantage due to reduced cell size as well as specific adopted strategies (Ostrowski et al, 2010; Scanlan et al, 2009). In addition to specific acquisition and utilisation mechanisms, isolates of *Synechococcus* have been reported to produce

secondary metabolites which have inhibitory effects on the growth of other strains, thus potentially functioning as a defense mechanism against competition (Paz-Yepes et al, 2013).

Though each individual environmental variable could play a significant role in shaping the distribution of marine picocyanobacteria, multiple variables could interact and influence niche selection. Teasing out the variable with the most influence would require *in situ* complementation studies and modelling.

1.8 'Omics'based- tools to explore picocyanobacterial ecology and molecular biology

1.8.1 *Amplicon sequencing, genomics and metagenomics*

Since their initial identification and strain differentiation based on morphological and fluorescence characteristics, there is currently a much better understanding of the considerable genetic diversity within the marine picocyanobacterial group. Conserved phylogenetic markers are useful tools in determining genetic diversity and molecular phylogeny. Phylogenetic markers commonly used to determine molecular diversity in picocyanobacteria include 16S rRNA, 16S-23S internal transcribed spacer region (ITS), *rpoC1* (RNA polymerase subunit), *petB* (cytochrome b_6 subunit of cytochrome b_6f complex), *rbcL* (RuBisCO large subunit), *psbA* (photosystem II), *cpeB* (phycoerythrin), *narB* (nitrate reductase) and *ntcA* (nitrogen regulator) genes (Ahlgren and Rocap, 2006; Ahlgren and Rocap 2012; Herdman et al, 2001; Mazard et al, 2012; Paerl et al, 2011, 2012; Penno et al, 2006; Post et al, 2011; Steglich et al, 2003; Toledo and Palenik, 1997; Zeidner et al, 2003). These markers are useful to distinguish *Synechococcus* and *Prochlorococcus* lineages. Molecular markers are thus used to determine the distribution and abundance of picocyanobacteria as well as their community composition and structure in various oceanic environments. The phylogeny and distribution provide insights into the evolutionary pressures underlying their niche partitioning.

Comparisons of whole genomes enable the elucidation of the diversity of genetic repertoires and niche adaptation as well as taxonomy and evolution. Community metagenomics is another technique used extensively to explore the diversity, interactions and functional roles of members in microbial assemblages from complex environments, without potential isolation and culturing biases

(Handelsman, 2004). Genomic tools have been used to study the picocyanobacterial community diversity and the shifts in population structure in response to environmental changes (Villar et al, 2015). These techniques are also valuable in elucidating genetic exchange and rearrangements as well as the potential selective pressures that drive such changes (Palenik et al, 2009; Tai et al, 2011). The genetic repertoire of a population provides insights into the metabolic versatility and functional potential for adaptation to specific niches (Mazard et al, 2011).

1.8.2 *Gene expression analyses: transcriptomics and proteomics*

The ability of an organism to colonise an ecological niche is dependent on the genetic makeup of the organism and its capacity to respond and acclimate to the prevalent conditions. For a ubiquitous organism such as *Synechococcus*, comparisons of the genome content among the multiple lineages, which occupy specific ecological niches, can provide useful information regarding the genetic diversity and the presence or absence of genes. However, to gain more insight into the regulation and functional capacity, gene expression profiles such as expressed transcriptomes are required. Changes in RNA abundance profile at the global cellular level in response to changes in environmental conditions provide an insight into the acclimation responses and the mechanisms employed to cope with prevalent conditions. Apart from the transcriptional level, gene expression can be regulated at the post-transcriptional, translational as well as post-translational levels, therefore exploring the expression changes at the protein level is also important. For a better and more thorough understanding of the diversity in functional capacity of an individual lineage, global RNA and protein expression profiles in response to changes in environmental conditions need to be investigated. High-throughput techniques such as genome microarrays, RNA sequencing and quantitative label-free shotgun proteomics are powerful tools for profiling gene expression changes. Examining the transcriptional and proteomic responses to heat shock undertaken in the freshwater cyanobacteria, *Synechocystis* sp. PCC 6803 revealed several heat responsive genes (Suzuki et al, 2006). This combined approach highlighted the multiple levels of regulation at the RNA and protein expression stages. In marine *Synechococcus*, global expression studies have focussed on transcriptional responses

on exposure to stresses such as nutrient limitation, heavy metals and other chemicals. These studies have revealed an array of general and stress-specific acclimation responses as well as distinct differences amongst isolates inhabiting diverse oceanic niches. These studies have highlighted the role of two-component systems and other regulatory mechanisms as well as the adjustments of transport and other metabolic processes utilised to cope with environmental stresses (Ostrowski et al, 2010; Stuart et al, 2009; Tetu et al, 2009, 2013). More recent approaches such as metatranscriptomics and metaproteomics aim to understand gene expression and stress response *in situ* instead of laboratory-based cultures of model organisms (Palenik, 2015).

Thus a combined approach using phylogenetic marker profiling to determine distribution along with genome comparisons and gene expression profiles will provide a comprehensive view on the biogeography and physiological adaptations that enable niche specialisation.

1.9 Rationale and objectives

Representatives of the genus *Synechococcus* can be found across diverse marine environments spanning polar and equatorial latitudes. These picocyanobacteria are ideal model systems to understand evolution and niche partitioning. The genetic diversity in *Synechococcus* lineages enables their ability to tolerate the selective pressures prevalent in their environment of choice and therefore facilitate their widespread distribution. Their distribution across different thermal regimes suggests an ability to withstand a wide range of temperatures (0°C – 30°C). Individual lineages of *Synechococcus* vary in their distribution and abundance across temperature ranges suggesting intrinsic differences in temperature tolerance and growth optimum. Furthermore, lineages which occur in sub-tropical and high latitudes are exposed to wider fluctuations in temperature, hence, might require particular acclimation mechanisms to cope with these changes.

Global mean surface temperatures are predicted to rise by 1°C – 5°C by the year 2100 (IPCC, 2014). Changes in temperature and CO_2 concentrations can influence the physiology, biotic interactions and ecological function of picocyanobacteria. Fu et al (2007) showed that increased temperature and CO_2

had a significant influence on the growth and photosynthesis of *Synechococcus* strain WH7803 whilst in *Prochlorococcus* strain MED4 the effect was not evident. Environmental changes such as rising CO_2 and temperature can have specific effects on lineages within the picocyanobacterial genera (Fu et al, 2007). If changes in temperature affect individual lineages of picocyanobacteria differently, it might lead to changes in their distribution and community composition. As key members of the marine ecosystem, changes in their community structure could impact on primary production and biogeochemical cycles. Therefore, it is important to understand how temperature affects these photosynthetic microorganisms.

Different growth temperature optima have been observed in lineages confined to particular latitudinal range (Mackey et al, 2013; Pittera et al, 2014). These studies determined temperature associated changes in growth and photosynthesis in several *Synechococcus* isolates. Based on photophysiology the study by the two studies reported that adjustments of phycobilisomes, state transitions and regulation of photosynthetic electron flow in response to changes in metabolic demand could be important temperature acclimation strategies.

In addition to photophysiology, changes in gene expression in response to temperature will provide further insights to acclimation responses to temperature change. Mackey et al (2013) reported changes in the expression of phycobilisomes and photosynthetic proteins in response to changes in temperature. In addition to proteins related to photosynthesis, organisms regulate the expression of other cellular proteins in response to temperature which may be important for acclimation. Furthermore, comparative analyses of temperature-related molecular changes across multiple lineages is key to fully appreciate acclimation strategies specific to each lineage and to further understand temperature-niche preference and distribution across specific temperature ranges.

Temperature stress responses and acclimation strategies are well explored in freshwater *Synechococcus* and other genera of cyanobacteria. Exposure to temperatures above and below the optimum growth temperature has been reported to evoke physiological and metabolic changes in these

organisms. High and low temperatures can affect photosynthesis and electron transport chain, energy metabolism, protein structure and function as well as membrane fluidity. Temperature stress also evokes responses that are similar to other stresses such as light and oxidative stress suggesting cross-talk amongst stress regulatory mechanisms. Several acclimation strategies are employed by other cyanobacteria to tolerate temperature stresses. These include prevention of over-reduction of photosynthetic electron transport, and protection and repair of photosynthetic machinery, adjustments of membrane fluidity and protection from reactive species.

Mechanisms employed by marine picocyanobacteria to acclimate to temperature changes require further investigation. This study aims to explore the responses of marine *Synechococcus* lineages to different temperature conditions and determine the genetic heterogeneity among lineages that enable adaptation to particular growth temperature optima. Genetic factors influencing the adaptation of distinct lineages to different temperature niches are explored by determining gene expression at the translational and transcriptional levels.

Marine picocyanobacteria are significant contributors to primary production, biogeochemical cycling and ecosystem functioning, therefore understanding their distribution and diversity across the marine environment is central to predictive global models of biogeochemical cycles and environmental change. In order to better understand the environmental parameters and processes influencing their distribution, changes in the spatial and temporal structure of picocyanobacterial community is required. The distribution of marine picocyanobacteria in regions around the coast of Australia and the Southern Ocean are less explored. This study aims to determine the picocyanobacterial community composition in these regions and to investigate the influence of temperature in shaping the distribution of various picocyanobacterial lineages. Understanding the environmental factors that shape microbial community structure will enable better prediction of community responses to environmental variations.

1.10 Specific questions addressed:

a. How does low temperature influence the growth and physiology of marine *Synechococcus* isolates? What are the strategies utilised by tropical compared to temperate isolates to acclimate to low temperature? [Chapter 2]
b. What are the growth patterns of marine *Synechococcus* isolates at high temperature? How do the responses to high temperature vary between representatives of the major marine *Synechococcus* clades (I, II, III and IV)? [Chapter 3]
c. What is the community composition and structure of marine picocyanobacteria within environments presenting various thermal characteristics? How does temperature influence the picocyanobacterial community composition and what other factors influence the distribution of lineages of picocyanobacteria? [Chapter 4 and 5]
d. What are the effects of other factors such as anthropogenic chemical pollutants on marine *Synechococcus* [Chapter 6]

1.11 References

Adams D, Duggan P, Jackson O. (2012). Cyanobacterial symbioses. In: Whitton B (ed.). Ecology of Cyanobacteria II: Their Diversity in Space and Time. Springer Science+Business Media. pp 593-646.

Ahlgren N, Rocap G. (2006). Culture isolation and culture-independent clone libraries reveal new marine *Synechococcus* ecotypes with distinctive light and N physiologies. Appl Environ Microbiol 72: 7193-7204.

Ahlgren N, Rocap G. (2012). Diversity and distribution of marine *Synechococcus*: multiple gene phylogenies for consensus classification and development of qPCR assays for sensitive measurement of clades in the ocean. Front Microbiol 3: 1-24.

Ammerman J, Hood R, Case D, Cotner J. (2003). Phosphorus deficiency in the Atlantic: An emerging paradigm in oceanography. Eos Trans AGU 84: 165-170.

Bouman H, Ulloa O, Scanlan D, Zwirglmaier K, Li W, Platt T, *et al*. (2006). Oceanographic basis of the global surface distribution of *Prochlorococcus* ecotypes. Science 312: 918-921.

Cai H, Wang K, Huang S, Jiao N, Chen F. (2010). Distinct patterns of picocyanobacterial communities in winter and summer in the Chesapeake Bay. Appl Environ Microbiol 76: 2955–2960.

Callieri C, Cronberg G, Stockner J. (2012). Freshwater picocyanobacteria: Single cells, microcolonies and colonial forms. In: Whitton B (ed.). Ecology of Cyanobacteria II: Their Diversity in Space and Time. Springer Science+Business Media, pp 229-269.

Callieri C, Stockner J. (2002). Freshwater autotrophic picoplankton: a review. J Limnol 61: 1-14.

Campbell L, Nolla H, Vaulot D. (1994). The importance of *Prochlorococcus* to community structure in the central North Pacific. Limnol Oceanogr 39:954–961.

Carr N, Wyman M. (1986). Cyanobacteria: Their biology in relation to the oceanic picoplankton. Can Bull Fish Aquat Sci 214: 159-204.

Cavalier-Smith T. (2006). Cell evolution and Earth history: stasis and revolution. Phil Trans R Soc B 361: 969-1006.

Chisholm S, Olson R, Zettler E, Waterbury J, Goericke R, Welschmeyer N. (1988). A novel free-living prochlorophyte occurs at high cell concentrations in the oceanic euphotic zone. Nature 334:340–343.

Choi D, Noh J. (2009). Phylogenetic diversity of *Synechococcus* strains isolated from the East China Sea and the East Sea. FEMS Microbiol Ecol 69: 439–448.

Cohen-Bazire G, Bryant D. (1982). Phycobilisomes composition and structure. In: Carr N, Whitton B (eds.). The Biology of Cyanobacteria. Blackwell, Oxford, pp 143–190.

Cottrell M, Kirchman D. (2009). Photoheterotrophic microbes in the Arctic Ocean in summer and winter. Appl Environ Microbiol 75:4958–4966.

Dufresne A, Ostrowski M, Scanlan D, Garczarek L, Mazard S, Palenik B, *et al*. (2008). Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. Genome Biology 9:R90.

Flombaum P, Gallegos J, Gordillo R, Rincón J, Zabala L, Jiao N, *et al*. (2013). Present and future global distributions of the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. Proc Natl Acad Sci USA 110: 9824-9829.

Foster R, Kuypers M, Vagner T, Paerl R, Musat N, Zehr J. (2011). Nitrogen fixation and transfer in open ocean diatom-cyanobacterial symbioses. ISME J 5: 1484-1493.

Fu F, Warner M, Zhang Y, Feng Y, Hutchins D. (2007). Effects of increased temperature and CO2 on photosynthesis, growth, and elemental ratios in marine *Synechococcus* and *Prochlorococcus* (cyanobacteria). J Phycol 43: 485-496.

Fuller N, Marie D, Partensky F, Vaulot D, Post A, Scanlan D. (2003). Clade-specific 16S ribosomal DNA oligonucleotides reveal the predominance of a single marine *Synechococcus* clade throughout a stratified water column in the Red Sea. Appl Environ Microbiol 69: 2430-2443.

Fuller N, Tarran G, Yallop M, Orcutt K, Scanlan D. (2006). Molecular analysis of picocyanobacterial community structure along an Arabian Sea transect reveals distinct spatial separation of lineages. Limnol Oceanogr 51: 2515-2526.

Garczarek L, Dufresne A, Blot N, Cockshutt A, Peyrat A, Campbell D, *et al*. (2008). Function and evolution of the *psbA* gene family in marine *Synechococcus*: *Synechococcus* sp. WH7803 as a case study. ISME J 2: 937-953.

Garczarek L, Dufresne A, Rousvoal S, West N, Mazard S, Marie D, *et al*. (2007). High vertical and low horizontal diversity of *Prochlorococcus* ecotypes in the Mediterranean Sea in summer. FEMS Microbiol Ecol 60:189–206.

Giovannoni S, Turner S, Olsen G, Barns S, Lane D, Pace N. (1988). Evolutionary relationships among cyanobacteria and green chloroplasts. J Bacteriol 170**:** 3584-3592.

Goericke R, Repeta D. (1992). The pigments of *Prochlorococcus marinus*: the presence of divinyl chlorophyll *a* and *b* in a marine prochlorophyte. Limnol Oceanogr 37:425–433.

Hambly E, Suttle C. (2005). The virioshpere, diversity, and genetic exchange within phage communities. Curr Opin Microbiol 8:444-450.

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

Hayes J. (1983). Geochemical evidence bearing on the origin of aerobiosis, a speculative hypothesis. In: Schopf J (ed.). The Earth's earliest biosphere, its origins and evolution. Princeton University Press, Princeton, New Jersey, pp 291-300.

Herdman M, Castenholz R, Iteman I, Waterbury J, Rippka R. (2001). Subsection I (Formerly *Chroococcales* Wettstein 1924, emend. Rippka, Deruelles, Waterbury, Herdman and Stanier 1979). In: Boone D, Castenholz R, Garrity G (eds.). Bergey's manual of systematic bacteriology, vol 1, 2nd edn, The archaea and the deeply branching and phototrophic bacteria. Springer, New York, pp 493–514.

Hoffman L. (1999). Marine cyanobacteria in tropical regions: diversity and ecology. Eur J Phycol 34: 371-379.

Huang S, Wilhelm S, Harvey H, Taylor K, Jiao N, Chen F. (2012). Novel lineages of *Prochlorococcus* and *Synechococcus* in the global oceans. ISME J 6: 285-297.

Huang S, Zhang S, Jiao N, Chen F. (2015). Marine cyanophages demonstrate biogeographic patterns throughout the global ocean. Appl Environ Microbiol 81:441– 452.

IPCC, 2014: Climate Change 2014: Synthesis report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core writing team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Jardillier L, Zubkov M, Pearman J, Scanlan D. (2010). Significant CO_2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. ISME J 4:1180–1192.

Johnson P, Sieburth J. (1979). Chroococcoid cyanobacteria in the sea: a ubiquitous and diverse phototrophic biomass. Limnol Oceanogr 24:928–935.

Johnson Z, Zinser E, Coe A, McNulty N, Woodward E, Chisholm S. (2006). Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. Science 311: 1737-1740.

Knoll A. (2008). Cyanobacteria and Earth history. In: Herrero A (ed.). The Cyanobacteria: Molecular biology, Genomics, and Evolution. Horizon Scientific Press, pp 1-19.

Lavin P, Gonzáles B, Santibáñez J, Scanlan D, Ulloa O. (2010). Novel lineages of *Prochlorococcus* thrive within the oxygen minimum zone of the eastern tropical South Pacific. Environ Microbiol Rep 2:728–738.

Li W. (1994). Primary productivity of prochlorophytes, cyanobacteria, and eucaryotic ultraphytoplankton: measurements from flow cytometric sorting. Limnol Oceanogr 39:169–175.

Mackey K, Paytan A, Caldeira K, Grossman A, Moran D, McIlvin M, *et al*. (2013). Effect of temperature on photosynthesis and growth in marine *Synechococcus* spp. Plant Physiol 163: 815-829.

Martiny A, Tai A, Veneziano D, Primeau F, Chisholm S. (2009). Taxonomic resolution, ecotypes and the biogeography of *Prochlorococcus*. Environ Microbiol 11:823–832.

Mazard S, Ostrowski M, Garczarek L, Scanlan D. (2011). A targeted metagenomic approach to determine the "population genome" of marine *Synechococcus*. In: de Bruijn F (ed.). Handbook *of Molecular Microbial Ecology, Volume II: Metagenomics in Different Habitats*, First Edition. Wiley-Blackwell, pp 301-307.

Mazard S, Ostrowski M, Partensky F, Scanlan D. (2012). Multi-locus sequence analysis, taxonomic resolution and biogeography of marine *Synechococcus*. Environ Microbiol 14: 372-386.

Mella-Flores D, Mazard S, Humily F, Partensky F, MahéF, Bariat L, *et al*. (2011). Is the distribution of *Prochlorococcus* and *Synechococcus* ecotypes in the Mediterranean Sea affected by global warming? Biogeosciences 8: 2785-2804.

Moore L, Goericke R, Chisholm S. (1995). Comparative physiology of *Synechococcus* and *Prochlorococcus*: influence of light and temperature on growth, pigments, fluorescence and absorptive properties. Mar Ecol Prog Ser 116: 259-275.

Moore L, Rocap G, Chisholm S. (1998). Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. Nature 393:464–467.

Morel F, Price N. (2003). The biogeochemical cycles of trace metals in the oceans. Science 300: 944-947.

Mühling M, Fuller N, Millard A, Somerfield P, Marie D, Wilson W, *et al*. (2005). Genetic diversity of marine *Synechococcus* and co-occurring cyanophage communities: evidence for viral control of phytoplankton. Environ Microbiol 7:499–508.

Ostrowski M, Mazard S, Tetu S, Phillippy K, Johnson A, Palenik B, *et al*. (2010) PtrA is required for coordinate regulation of gene expression during phosphate stress in a marine *Synechococcus* . ISME J 4:908–921.

Paerl R, Johnson K, Welsh R, Worden A, Chavez F, Zehr J. (2011). Differential distributions of *Synechococcus* sub-groups across the California current system. Front Microbiol 2:59.

Paerl R, Turk K., Beinart R, Chavez F, Zehr J. (2012). Seasonal change in the abundance of *Synechococcus* and multiple distinct phylotypes in Monterey Bay determined by *rbcL* and *narB* quantitative PCR. Environ Microbiol 14: 580–593.

Palenik B. (2015). Molecular mechanisms by which marine phytoplankton respond to their dynamic chemical environment. Annu Rev Mar Sci 7: 325-340.

Palenik B, Brahamsha B, Larimer F, Land M, Hauser L, Chain P, *et al*. (2003). The genome of a motile marine *Synechococcus*. Nature 424: 1037-1042.

Palenik B, Ren Q, Dupont C, Myers G, Heidelberg J, Badger J, *et al*. (2006). Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment. Proc Natl Acad Sci USA 103: 13555-13559.

Palenik B, Ren Q, Tai V, Paulsen I. (2009). Coastal *Synechococcus* metagenome reveals major roles for horizontal gene transfer and plasmids in population diversity. Environ Microbiol 11: 349-359.

Partensky F, Blanchot J, Vaulot D. (1999). Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. Bull Inst Oceanogr 19: 457–475.

Partensky F, Vaulot D. (1999). Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. Bulletin de l'Institut océanographique de Monaco 19: 457-475.

Paz-Yepes J, Brahamsha B, Palenik B. (2013). Role of a microcin-C-like biosynthetic gene cluster in allelopathic interactions in marine *Synechococcus*. Proc Natl Acad Sci USA 110: 12030-12035.

Penno S, Lindell D, Post A. (2006). Diversity of *Synechococcus* and *Prochlorococcus* populations determined from DNA sequences of the N-regulatory gene *ntcA*. Environ Microbiol 8: 1200–1211.

Pentecost A, Whitton B. (2012). Subaerial cyanobacteria. In: Whitton B (ed.). Ecology of Cyanobacteria II: Their diversity in space and time. Springer Science+Business Media, pp 291-316.

Pittera J, Humily F, Thorel M, Grulois D, Garczarek L, Six C. (2014). Connecting thermal physiology and latitudinal niche partitioning in marine *Synechococcus*. ISME J 1-16.

Post A, Penno S, Zandbank K., Paytan A, Huse S, Welch D. (2011). Long term seasonal dynamics of *Synechococcus* population structure in the Gulf of Aqaba, Northern Red Sea. Front Microbiol 2:131.

Raven J. (2012). Carbon. In: Whitton B (ed.). Ecology of Cyanobacteria II: Their diversity in space and time. Springer Science+Business Media, pp 443-460.

Rocap G, Distel D, Waterbury J, Chisholm. (2002). Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. Appl Environ Microbiol 68: 1180-1191.

Rocap G, Larimer F, Lamerdin J, Malfatti S, Chain P, Ahlgren N, *et al*. (2003). Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. Nature 424: 1042–1047.

Saito M, Rocap G, Moffett J. (2005). Production of cobalt binding ligands in a *Synechococcus* feature at the Costa Rica upwelling dome. Limnol Oceanogr 50: 279–290.

Scanlan D. (2012). Marine picocyanobacteria. In: Whitton B (ed.). Ecology of Cyanobacteria II: Their diversity in space and time. Springer Science + Business Media, pp 503-533.

Scanlan D, Ostrowski M, Mazard S, Dufresne A, Garczarek L, Hess W, *et al*. (2009). Ecological genomics of marine picocyanobacteria. Microbiol Mol Biol Rev 73: 249-299.

Schopf J. (2012). The fossil record of cyanobacteria. In: Whitton B (ed.). Ecology of Cyanobacteria II: their diversity in space and time. Springer Science+Business Media, pp 15-36.

Six C, Thomas J, Garczarek L, Ostrowski M, Dufresne A, Blot N, *et al*. (2007). Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. Genome Biology 8: R259.

Sohm J, Ahlgren N, Thomson Z, Williams C, Moffett J, Saito M, *et al*. (2015). Co-occurring *Synechococcus* ecotypes occupy four major oceanic regimes defined by temperature, macronutrients and iron. ISME J 1-13.

Stal L. (2008). Nitrogen Fixation in Cyanobacteria. In: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd, Chichester.

Steglich C, Post A, Hess W. (2003). Analysis of natural populations of *Prochlorococcus* spp. in the northern Red Sea using phycoerythrin gene sequences. Environ Microbiol 5: 681–690.

Stockner J. (1988). Phototrophic picoplankton: an overview from marine, freshwater ecosystems. Limnol Oceanogr 33: 765–775.

Strom S, Brahamsha B, Fredrikson K, Apple J, Rodríguez A. (2012). A giant cell surface protein in *Synechococcus* WH8102 inhibits feeding by a dinoflagellate predator. Environ Microbiol 14: 807–816.

Stuart R, Dupont C, Johnson A, Paulsen I, Palenik B. (2009). Coastal strains of marine *Synechococcus* species exhibit increased tolerance to copper shock and a distinctive transcriptional response relative to those of open-ocean strains. Appl Environ Microbiol 75: 5047-5057.

Suttle C, Chan A. (1994). Dynamics and distribution of cyanophages and their effect on marine *Synechococcus*. Appl Environ Microbiol 60: 3167-3174.

Tai V, Burton R, Palenik B. (2011). Temporal and spatial distributions of marine *Synechococcus* in the Southern California Bight assessed by hybridization to bead-arrays. Mar Ecol Prog Ser 426: 133–147.

Tetu S, Brahamsha B, Johnson D, Tai V, Phillippy K, Palenik B, *et al*. (2009). Microarray analysis of phosphate regulation in the marine cyanobacterium *Synechococcus* sp. WH8102. ISME J 3: 835-849.

Tetu S, Johnson D, Varkey D, Phillippy K, Stuart R, Dupont C, *et al*. (2013). Impact of DNA damaging agents on genome-wide transcriptional profiles in two marine *Synechococcus* species. Front Microbiol 4.

Thompson J, Higgins D, Gibson T. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 22: 4673-4680.

Toledo G, Palenik B. (1997). *Synechococcus* diversity in the California current as seen by RNA polymerase (*rpoC1*) gene sequences of isolated strains. Appl Environ Microbiol 63: 4298–4303.

Urbach E, Scanlan D, Distel D, Waterbury J, Chisholm S. (1998). Rapid Diversification of Marine Picophytoplankton with Dissimilar Light-Harvesting Structures Inferred from Sequences of *Prochlorococcus* and *Synechococcus* (Cyanobacteria). J Mol Evol 46: 188-201.

Villar E, Farrant G, Follows M, Garczarek L, Speich S, Audic S, *et al*. (2015) Environmental characteristics of Agulhas rings affect interocean plankton transport. Science 348.

Ward D, Castenholz R, Miller S. (2012). Cyanobacteria in geothermal habitats. In: Whitton B (ed.). Ecology of Cyanobacteria II: Their diversity in space and time. Springer Science+Business Media, pp 39-63.

Waterbury J, Watson S, Guillard R, Brand L. (1979). Widespread occurrence of a unicellular, marine, planktonic, cyanobacterium. Nature 277: 293-294.

Waterbury J, Watson S, Valois F, Franks D. (1986). Biological and ecological characterization of the marine unicellular cyanobacteria *Synechococcus*. Can Bull Fish Aquat Sci 214: 71–120.

West N, Lebaron P, Strutton P, Suzuki M. (2011). A novel clade of *Prochlorococcus* found in high nutrient low chlorophyll waters in the South and Equatorial Pacific Ocean. ISME J 5: 933–944.

Whitton B, Potts M. (2012). Introduction to the Cyanobacteria. In: Whitton B (ed.). Ecology of Cyanobacteria II: Their Diversity in Space and Time. Springer Science+Business Media, pp 1-14.

Zehr J. (2011). Nitrogen fixation by marine cyanobacteria. Trends Microbiol 19: 162-173.

Zeidner G, Preston C, Delong E, Massana R, Post A, Scanlan D, *et al*. (2003). Molecular diversity among marine picophytoplankton as revealed by *psbA* analyses. Environ Microbiol 5: 212–216.

Zhong Y, Chen F, Wilhelm S, Poorvin L, Hodson R. (2002). Phylogenetic diversity of marine cyanophage isolates and natural virus communities as revealed by sequences of viral capsid assembly protein gene g20. Appl Environ Microbiol 68: 1576–1584.

Zinser E, Coe A, Johnson Z, Martiny A, Fuller N, Scanlan D, *et* al. (2006). *Prochlorococcus* ecotype abundances in the North Atlantic Ocean as revealed by an improved quantitative PCR method. Appl Environ Microbiol 72: 723–732.

Zinser E, Johnson Z, Coe A, Karaca E, Veneziano D, Chisholm S. (2007). Influence of light and temperature on *Prochlorococcus* ecotype distributions in the Atlantic Ocean. Limnol Oceanogr 52: 2205-2220.

Zubkov M, Sleigh M, Burkill P, Leakey R. (2000). Picoplankton community structure on the Atlantic Meridional Transect: a comparison between seasons. Prog Oceanogr 45: 369–386.

Zwirglmaier K, Heywood J, Chamberlain K, Woodward E, Zubkov M, Scanlan D. (2007). Basin-scale distribution patterns of picocyanobacterial lineages in the Atlantic Ocean. Environ Microbiol 9: 1278-1290.

Zwirglmaier K, Jardillier L, Ostrowski M, Mazard S, Garczarek L, Vaulot D, *et al*. (2008). Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. Environ Microbiol 10: 147-161.

Zwirglmaier K, Spence E, Zubkov M, Scanlan D, Mann N. (2009). Differential grazing of two heterotrophic nanoflagellates on marine *Synechococcus* strains. Environ Microbiol 11: 1767–1776.

Chapter 2:

Effects of low temperature on tropical and temperate isolates of marine *Synechococcus*

Pages 41-52 of this thesis have been removed as they contain published material. Please refer to the following citation for details of the article contained in these pages:

Varkey, D., Mazard, S., Ostrowski, M., Tetu, S. G., Haynes, P., & Paulsen, I. T. (2016). Effects of low temperature on tropical and temperate isolates of marine *Synechococcus*. The ISME journal, 10(5), 1252-1263. DOI: 10.1038/ismej.2015.179

Chapter 3:

Growth physiology and cellular responses of marine *Synechococcus* isolates to high temperature

Title:

Growth physiology and cellular responses of marine *Synechococcus* isolates to high temperature

Authors:

Deepa Varkey, Martin Ostrowski, Sophie Mazard, Sasha G. Tetu and Ian T. Paulsen*

Authors' affiliation:

Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, NSW, Australia 2109

*Corresponding author: Prof Ian T Paulsen, ian.paulsen@mq.edu.au, Telephone: +61 2 9850 8152, Fax: +61 2 9850 8313.

Running Title: High temperature responses in marine *Synechococcus*

Synechococcus is an abundant photosynthetic prokaryote with a global distribution. The predicted rise in global sea surface temperature would likely significantly impact these key primary producers. The distinct partitioning of *Synechococcus* clades across latitudes suggests preferential temperature niches. This study aimed to determine how lineages with distinct temperature preferences cope with increased temperatures. Representative isolates of clades I, II, III and IV were grown at 26°C and 22°C to explore growth physiology and global cellular responses at the transcriptional and translational levels. The rates of growth at 26°C relative to 22°C, lower in temperate clades I/IV and higher in (sub)tropical clades II/III, provide evidence for temperature preferences. There were distinct differences amongst the isolates in their gene expression. The (sub)tropical clades, II and III mostly induced transcripts of phycobilisomes, photosystems and electron transport whilst temperate clades I and IV repressed the majority of genes associated with either phycobilisomes or photosystems. The induction of genes involved in protein turnover and photoprotection in all four strains suggests an important role in high temperature acclimation. Overall, the induction of photosynthesis-associated genes and higher rates of growth in (sub)tropical clades II and III at higher temperatures would provide them with a competitive advantage over their temperate counterparts.

Keywords: cyanobacteria/*Synechococcus*/proteomics/temperature/adaptation

INTRODUCTION

Cyanobacteria are photosynthetic prokaryotes that carry out a major part of the Earth's carbon fixation and primary production. They are a diverse and ubiquitous group of organisms found from terrestrial to freshwater environments as well as marine habitats. The marine picocyanobacteria, *Prochlorococcus* and *Synechococcus* are abundant photosynthetic microorganisms in the marine environment with cell numbers up to 10^6 cells ml^{-1} for *Prochlorococcus* and 10^5 cells ml^{-1} for *Synechococcus* (Scanlan et al, 2009; Waterbury et al, 1979; Zwirglmaier et al, 2008). Their abundance and ubiquitous occurrence make them

numerically dominant primary producers forming the base of marine food webs. They contribute significantly to oxygen production and biogeochemical cycling (Garczarek et al, 2008; Jardillier et al, 2010; Li, 1994; Palenik et al, 2009; Partensky and Vaulot, 1999; Scanlan et al, 2009).

Picocyanobacteria are distributed across the oceanic environment from the poles to the equator in nutrient-rich and nutrient-scarce regions. Their widespread occurrence requires the various lineages to withstand a range of abiotic and biotic environmental factors including light, nutrients, temperature, water column stability, predators and competition.

The genus *Synechococcus* is classified into three sub-clusters and at least twenty phylogenetically distinct lineages. The partitioning of these lineages across the oceanic environment suggests preferential niche selection due to exposure and adaptation to prevalent conditions. Niche differentiation is evident in the four most abundant lineages of *Synechococcus*, clades I, II, III and IV. Clades I and IV co-occur at higher abundances within the temperate latitudinal range above 30°N and below 30°S in coastal and/or mesotrophic waters (Paerl et al, 2011; Scanlan, 2012; Tai and Palenik, 2009; Zwirglmaier et al, 2008). In contrast, clade II is more abundant within the tropical/subtropical regions. This group is widespread across mesotrophic coastal and continental shelf zones as well as offshore oligotrophic waters (Ahlgren and Rocap, 2006; Ahlgren and Rocap, 2012; Scanlan et al, 2009; Zwirglmaier et al, 2008). Clade III is most abundant in oligotrophic open ocean waters in sub-tropical and tropical waters (Scanlan, 2012; Zwirglmaier et al, 2008).

Lineages of *Synechococcus* occur at temperatures as low as 2°C. The upper temperature limit for both *Synechococcus* and *Prochlorococcus* appears to be 30°C (Johnson et al, 2006; Moore et al, 1995; Partensky and Vaulot, 1999; Scanlan, 2012). Temperature is potentially a prominent factor influencing the distribution of lineages in distinct latitudinal ranges. Temperature preference among these clades is indicated by their relative abundance at different temperature ranges. Highest abundance of clades I and

IV have been detected between temperatures 7°C and 20°C while clade II and III occur at higher abundance at temperatures above 20°C up to 28°C. For other clades preferential temperature ranges are unclear, and other environmental factors may be important for their niche speciation (Scanlan, 2012). Changes in sea surface temperatures could influence *Synechococcus* and *Prochlorococcus* distribution and community composition. Global distribution models of marine cyanobacteria predict significant changes in their abundance and distribution under future climate conditions of increased sea surface temperature (Flombaum et al, 2013). Fu et al (2007) showed that higher CO_2 and temperature conditions affected the growth and photosynthetic performance of one isolate of *Synechococcus* (strain WH7803) and *Prochlorococcus* (strain MED4).

Growth and survival of photosynthetic organisms is dependent on the balance in energy absorption, conversion and metabolic utilisation. Photosynthetic reaction centres which utilise light energy for the production of chemical energy in the form of ATP and NADPH are highly sensitive to changes in environmental parameters including light, UV radiation and temperature (Allakhverdiev et al, 2008; Fu et al, 2007; Huner et al, 1998). Disturbance of the balance of incident versus utilised light energy results in the inactivation of photosystems, disruption of electron transport and oxidative stress. Temperatures above growth optimum affect photosynthetic reactions centres, pigment composition, energy metabolism and general metabolic pathways including amino acid and protein biosynthesis. High temperature affects lipid membrane composition and organisation, protein structure and functioning (Allakhverdiev et al, 2008; Fu et al, 2007; Huner et al, 1998; Wen et al, 2005). Heat-shock proteins such as chaperones, and proteases as well as sigma factors and signal transduction genes play an important role in cellular adaptation to heat stress in freshwater cyanobacteria (Castielli et al, 2009; Glatz et al, 1999; Suzuki et al, 2006).

With the predicted rises in global sea surface temperatures associated with climate change, it is important to understand how temperature influences the different picocyanobacterial strains and how they cope with changes in temperature. Acclimation and adaptation to temperature change may involve changes in

cellular physiology as well as affect the distribution of strains based on their specific local niches and the ability to cope.

Representative strains from clades I to IV were subjected to a growth temperature higher than the reference temperature (26°C compared to 22°C) and their response analysed through transcriptomics and proteomics. This study aims to gain an understanding of the physiological and molecular changes associated with growth at high temperatures in strains which are adapted to different temperature niches. Flombaum et al 2013 predicted significant changes in *Synechococcus* populations in the 30°S to 60°S latitudinal regions in 2050. Since populations within temperate regions appear to be most affected, this study aimed to determine how high temperature affects temperate isolates. The study also aimed to compare the responses between isolates to determine if the tropical isolates grow better at 26°C due to the expression of certain genes that are not expressed by temperate isolates.

MATERIALS AND METHODS

Cell culture and growth

Synechococcus isolates CC9311 (clade I), WH8109 (clade II), WH8102 (clade III) and BL107 (clade IV) were grown in synthetic ocean water-based (SOW) medium with 13.4 µM $Na_2EDTA.2H_2O$, 87.6 µM K_2HPO_4, 96.8 µM Na_2CO_3 and 9 mM $NaNO_3$ (Morel et al, 1979; Su et al, 2006). Cultures were grown at 26°C (high) and 22°C (reference) under constant illumination of 20 µmol photons m^{-2} s^{-1} and agitation at 100 rpm. The temperatures at which clades I and IV occur at highest relative abundance range from 7°C-20°C, therefore 26°C was chosen as sufficiently high to evoke thermal stress responses in these temperate clades. Though 26°C is within the temperature range (20°C to 28°C) at which clades II and III occur at high relative abundance, the molecular responses produced at this temperature would be useful to compare tropical and temperate isolates. Cultures were acclimated to experimental temperature conditions through three successful serial transfers (8 generations) prior to harvesting cells in logarithmic phase. Experiments

were performed in biological triplicates for each condition. Growth was measured based on optical density at 750 nm (Beckmann DU 640 spectrophotometer). For strains CC9311 and WH8102, photosynthetic efficiency was measured during the entire growth period as the quantum yield (Fv/Fm) of photosystem II (PSII) using a Phytoplankton analyser, Phyto-PAM with PhytoWin V1.45 software (Heinz Walz GmbH, Germany). An aliquot of the culture (1 ml) was dark-adapted (5 min) prior to measuring yield using a saturated pulse of light. Averages for growth rates and photosynthetic yield during exponential phase were calculated and statistical analyses were performed using independent two-sample *t*-tests.

Absence of heterotrophic bacterial contamination was verified by plating an aliquot of the cultures on ASW medium solidified with 10.0 g l^{-1} agar and supplemented with 100 mg l^{-1} peptone.

RNA-Seq Transcriptomics

Gene expression at 26°C compared to 22°C was determined using RNA-Seq transcriptomics. Cells were harvested from 100 ml of culture in logarithmic phase by centrifugation at 7000 x *g*, 5 min at 4°C (to ensure consistency between control and treatment and minimize metabolic changes). Cells were washed with Tris-EDTA buffer (pH 8.0), centrifuged at 3100 x *g*, 5 min and processed immediately. Cells were lysed using TRIzol reagent (Ambion, Australia) at 55°C for 40 min. RNA purification was performed using the miRNeasy Mini kit (Qiagen, Australia). DNase digestion was carried out using TURBO DNA-free kit (Ambion, Australia) with the addition of ribonuclease inhibitor (RNaseOUT, Invitrogen, Australia). Absence of genomic DNA contamination was confirmed using PCR. Following purification, RNA was stored at -80°C until further processing. For RNA samples of all strains except that of strain WH8102, ribosomal RNA (rRNA) was depleted using RiboZero rRNA depletion kit (Epicentre, Australia) and purified using RNeasy MinElute cleanup kit (Qiagen, Australia). For strain WH8102, non-depleted RNA samples were used since depletion resulted in the recovery of very low RNA concentrations. The cDNA library was constructed using TruSeq stranded mRNA library preparation kit (Illumina, Australia).

Library construction and RNA sequencing was performed at the Ramaciotti Centre for Genomics. Sequencing was performed on a HiSeq 2500 system 100bp paired-end run (BL107, CC9311 and WH8109) and on a NextSeq 500 system 75bp paired-end run (WH8102).

Sequencing reads were processed and quantitated using gene expression analysis software EDGE-pro v1.3.1 (Estimated Degree of Gene Expression in PROkaryotes) (Magoc et al, 2013). Default parameters of EDGE-pro, except for number of threads (changed from default 1 to 12), were used for all analysis (Magoc et al, 2013). Low quality reads were filtered out and the remaining were mapped to the reference genome of each strain (BL107, Dufresne et al, 2008; CC9311, Palenik et al, 2006; WH8102, Palenik et al, 2003; WH8109, Sabehi et al, 2012)using Bowtie2 (Langmead and Salzberg, 2012). The output from EDGE-pro software containing gene expression levels is further analysed using the DESeq2 software. Quantitation and statistical analyses of differential expression was performed across triplicate samples from the two growth temperatures 22°C and 26°C. Statistical significance of differential expression is inferred based on the Wald test for significance of generalised linear model coefficients. To account for multiple hypotheses testing, the Benjamini-Hochberg multiple testing adjustment is applied to obtain adjusted *p*-values. Changes in expression were considered significant if adjusted *p*-value was less than 0.05 and $\log_2$ fold change was greater than 0.5

Label-free shotgun proteomics

Cells were harvested from 500 ml of culture in logarithmic phase by centrifugation at 7000 x *g*, 10 min. Cells were washed with Tris-EDTA buffer (pH 8.0), centrifuged at 3100 x *g*, 10 min and stored at -80°C. Proteins were extracted from cell pellets using sodium dodecyl sulphate (SDS) lysis buffer (4.6% w/v SDS, 0.12 M Tris (pH 6.8), 1 mM EDTA, 4% w/v glycerol and 0.1% v/v ß-mercaptoethanol) with bead beating followed by centrifugation (10,000 x *g*, 10 min). Extracted proteins were precipitated using methanol/chloroform/water protocol (Wessel and Flügge, 1984) and resuspended in 2% w/v SDS and 50

mM Tris (pH 6.8). Protein samples were quantitated using the bicinchoninic acid protein assay (ThermoFisher Scientific, USA). Samples were diluted with sample loading buffer, denatured by boiling (95°C, 5 min) and fractionated using SDS-PAGE (10% precast gel; Bio-Rad, Australia). Sample loading and running buffers were prepared using suggested protocols (Bio-Rad, Australia). For each sample, 100µg of total protein was resolved into 16 fractions at 60V for 15 min followed by 110V for 45 min. After gel de-staining, peptides were extracted and digested with trypsin as described in Mirzaei et al, (2012). Peptide extracts resuspended in 0.1% formic acid were analysed using nanoflow liquid chromatography tandem mass spectrometry (Supplementary file 3.1) performed on a Q Exactive hybrid quadrupole-orbitrap mass spectrometer (Thermo Scientific, USA) at the Australian Proteome Analysis Facility (NSW, Australia).

Raw spectral files converted to MGF format were processed using the global proteome machine (GPM) software with the X!Tandem algorithm (Craig and Beavis, 2003; Craig and Beavis, 2004). Protein sequences extracted from the sequenced genome of each individual strain were used as the search database (Palenik et al, 2003, 2006; Dufresne et al, 2008). To ensure data quality, identified proteins were filtered based on two criteria, namely, reproducible identification across three replicates and a total spectral count of more than 5 (Gammulla et al, 2010; Mirzaei et al, 2011). A database of reversed sequences was searched to determine the false discovery rate.

Quantitative analyses of expressed proteins were performed based on spectral counts and normalised spectral abundance factor (NSAF) (Zybailov et al, 2006). NSAF for each protein (represented by k) in an experiment is calculated as the number of spectral counts (SpC) identifying k divided by its length (L), divided by sum of SpC/L of all identified proteins (Zybailov et al, 2006). Such normalisation accounts for differences in spectral counts dependent on protein length (Neilson et al, 2014). For statistical analyses, NSAF values were log-transformed to ensure normal distribution and checked using Q-Q plot and density estimation. Independent two-sample *t*-tests were performed on log transformed NSAF values comparing

protein expression at 26°C relative to 22°C using the program Scrappy, based on R modules (Neilson et al, 2014).

RESULTS AND DISCUSSION

Growth physiology

The growth of the four representative strains of *Synechococcus* clades I, II, III and IV was compared between high temperature (26°C) and reference temperature (22°C) conditions under constant irradiance (Figure 3.1; Table 3.1). Growth rates were lower for the temperate strains BL107 (clade IV) and CC9311 (clade I) at high temperature (18-20% lower than 22°C, $p < 0.05$). The significantly higher rates of growth ($p < 0.05$) in the (sub)tropical strains WH8109 (clade II) and WH8102 (clade III) at 26°C compared to 22°C suggests a higher temperature preference. The observed trends in growth correlate well with the ecological niche partitioning of these *Synechococcus* clades wherein preferred temperature range for clades I and IV is below 20°C and above 20°C for clades II and III (Zwirglmaier et al, 2008).

Photosynthetic yield of PSII (Fv/Fm) was measured using PAM fluorometry for strains CC9311 and WH8102 at 22°C and 26°C. The yield was higher at 26°C compared to 22°C in tropical strain WH8102 (p 0.05), which suggests that 26°C is closer to optimum than the control temperature (22°C). For the temperate strain CC9311, there was no significant difference in photosynthetic yield at the two temperature conditions (Figure 3.2).

Protein investment in functional groups

Protein investment across different functional categories could vary across strains based on their lifestyle strategies. Protein expression using whole cell lysate was determined through quantitative label-free shotgun proteomics and investment patterns were compared across three strains CC9311 (clade I), WH8102 (clade III) and BL107 (clade IV) grown at 22°C and 26°C. Proteins detected at the two

temperature conditions were grouped into functional categories based on Clusters of Orthologous Genes (COG, Tatusov et al, 2003) with custom groupings for proteins associated with photosynthesis (phycobilisomes (PS), photosystems, carboxysomes) and membrane transport.

PS proteins comprise a major proportion (18-36%) of the cellular investment in all strains, BL107, CC9311 and WH8102 (Supplementary table 3.1). In other cyanobacteria, PS has been reported to account for a significant fraction of total cellular protein (Tandeau de Marsac and Cohen-Bazire, 1977). At 26°C, there was a significant reduction in the overall abundance of PS in strain WH8102 (p <0.001) (Figures 3.3).

At 22°C, the percentage of proteins associated with PSI was lower in strain CC9311 in comparison to the other two strains (p <0.05; Supplementary table 3.1). The change in relative abundance of PSI and PSII proteins showed different trends at the two temperatures in strains BL107 and CC9311. A higher relative abundance of PSII proteins was observed in strain BL107 at 26°C in comparison to 22°C, whilst PSI was more abundant at 22°C. In strain CC9311, the differences were not significant. The relative abundance of carboxysome proteins did not vary significantly between the two temperatures in any of the strains.

The proportion of proteins associated with membrane transport (referred to as 'A' in figure 3.3) was higher (p <0.05) in strain WH8102 (6.1 – 7.3%) than strains BL107 (2.8 – 3.3%) and CC9311 (1.5 – 2.2%) at both temperatures (Figure 3.3; Supplementary table 3.1). The higher abundance of transporter proteins in WH8102 probably reflects a greater need for nutrient scavenging in the (sub)tropical oligotrophic open ocean, where clade III genotypes are characteristically found (Palenik et al, 2003; Zwirglmaier et al, 2008).

Differential gene expression in response to high temperature

Differential gene expression at 26°C in comparison to reference temperature (22°C) was determined for the four strains through transcriptomic analyses (RNA-Seq). Genes were considered to be significantly differentially expressed at high temperature in comparison to 22°C if adjusted *p*-value was less than 0.05 and $\log_2$ fold change was greater than 0.5. Gene expression change at 26°C relative 22°C was also determined at the proteomic level for strains BL107, CC9311 and WH8102. All changes with *p*-values less than 0.05 were considered significant.

Correlation in gene expression between transcript and protein levels was determined for strains BL107, CC9311 and WH8102 by plotting all genes that were differentially expressed at both levels (Supplementary figure 3.1). The number of genes that were differentially expressed at both levels was low (39-135). Positive but low correlation was obtained for strains BL107 and CC9311 whilst in strain WH8102 the correlation was limited. The reasons for limited correlation may be differences in regulation at the transcriptional and translational levels or differences in the processing methods for RNA and protein.

Selected genes showing differential expression at 26°C relative to 22°C at the transcript or protein levels in the four strains are discussed in the following sections.

Hypothetical genes

About 40 – 60% percent of genes that were differentially expressed at the transcript level are hypothetical or of unknown function. These genes were significantly differentially expressed and could play an important but as yet unknown role in temperature acclimation. There are six hypothetical genes (Supplementary table 3.2) that were differentially expressed in both temperate strains, suggesting a potential role in high temperature acclimation in cold-adapted strains. The two tropical strains also showed

common responses with six upregulated and two downregulated hypothetical/uncharacterised genes (Supplementary table 3.2). One of the upregulated hypothetical genes (SYNW1579, Syncc8109_1366) is conserved in the genomes of clades II and III and not present in temperate clades I and IV. This gene could potentially be involved in the acclimation of *Synechococcus* strains to high temperature conditions prevalent in lower- to mid-latitude niches.

Chaperones and proteases

High temperature can result in protein misfolding and aggregation. Chaperones such as *groEL*, *groES*, *dnaJ*, *dnaK*, *htpG* and small heat shock (sHSP) proteins are induced under conditions of stress and are important for acclimation to high temperature in cyanobacteria. Chaperones confer resistance to high temperature by ensuring protein quality. They bind to structurally unstable proteins that are unfolded or misfolded, prevent their aggregation and facilitate refolding and reactivation (Glatz et al, 1999; Rhee et al, 2012; Tanaka and Nakamoto, 1999). At 26°C, several heat shock proteins such as chaperones and proteases were differentially expressed in the four strains at the transcript and protein levels (Figure 3.4). The induction of chaperones and proteases at high temperature suggests an important role in high temperature tolerance through degradation of damaged proteins and/or regulation of protein expression. In the tropical strains, increased expression of chaperones and proteases could be to support faster protein turnover during faster growth. The *groEL* gene was induced in all four strains at the transcript/protein level suggesting a conserved role in high temperature acclimation across *Synechococcus* lineages. In addition to chaperones, protein quality and homeostasis is ensured by proteases. Proteases function by degrading unwanted/damaged proteins and are induced at high temperature in freshwater cyanobacteria (Yu and Houry, 2007). Among the four strains, the tropical strain WH8102 had the highest number of proteases that were induced at 26°C at both the transcript and protein levels. The induction of proteases suggests greater protein turnover when grown at 26°C compared to 22°C in strain WH8102, potentially supporting faster growth under more optimal conditions. None of the other chaperones or proteases

showed similar changes in expression across all four strains suggesting inherent differences in the response to temperature among the strains.

Phycobilisomes

Phycobilisomes are the light harvesting antenna complex in *Synechococcus* composed of phycobiliproteins which bind chromophore called phycobilins. Phycobiliproteins include allophycocyanin forming the core with phycocyanin (PC) and/or phycoerythrin (PE) as the rods. PE can bind the green light absorbing phycoerythrobilin (PEB) or a combination of PEB and blue light absorbing phycourobilin (PUB) (Six et al, 2007). Fu et al (2007) showed that in response to increase in temperature there was an associated increase in pigment content and PE to PC ratio in *Synechococcus* strain WH7803. The adjustment of the pigment content and composition would increase the absorption cross-section and efficiency of light harvesting, thereby increasing photosynthesis (Fu et al, 2007). Mackey et al (2013) also reported increase in PS content in response to growth at higher temperatures (23°C relative to 18°C and 15°C).

At 26°C, the expression changes of PS-related genes displayed varied patterns across the four strains (Figure 3.4). All PS-related genes showing differential expression were upregulated at the transcript level in tropical strain WH8102 whilst at the protein level an equal number was up- and downregulated. However, the overall abundance of PS proteins was considerably lower at 26°C than at 22°C (p<0.001; Figures 3.3 and Supplementary table 3.1). This difference in expression may be due to distinct regulation at the transcriptional and translational level or a faster rate of protein turnover. In strain WH8109, PS genes that were differentially expressed more than 0.5 $\log_2$ fold had similar numbers up- and downregulated at 26°C. However, there were several more upregulated genes which were just below the cutoff of 0.5 $\log_2$ fold. Induction of PS genes suggests an increase in light harvesting and utilisation

supporting growth at high temperature which in the case of tropical strains may be a more optimal condition.

In contrast to the tropical strains, the temperate strain BL107 downregulated genes associated with PS at the transcript level. A lower growth rate in BL107 could lower the requirement for light absorption and thus less PS expression. The downregulation could also imply a stress coping mechanism wherein light absorption is lowered to prevent oxidative damage if the utilisation of light energy is inefficient. Reduction in PS biosynthesis would also reduce the biosynthetic commitment potentially enabling the cells to better cope stressful conditions (Huang et al, 2002). At the protein level, though a few were repressed at 26°C, overall abundance of PS proteins did not change at the two temperatures which could suggest altered protein stability. In temperate strain CC9311, equal numbers of PS-associated transcripts were up- and downregulated whilst the difference in overall protein abundance was not statistically significant (Supplementary table 3.1). Changes in expression of PS-related genes suggests potential remodeling of PS in response to high temperature which has previously been observed in freshwater cyanobacterium, *Synechocystis* strain PCC6803 (Rowland et al, 2010). The distinct responses at the transcription and translation levels suggest differences in regulation. Additional fluorescence assays and pigment analyses will aid better interpretation of the high temperature-associated adjustment in PS structure.

Photosystems

In cyanobacteria and higher plants, photosynthesis and carbon fixation are highly sensitive to high temperature. Damage due to high temperature includes reduced carbon fixation and PSII activity through inactivation of reaction centre and oxygen-evolving complex. High temperature disrupts linear electron transport and causes an increase in the generation of reactive oxygen species (ROS) and inhibition of

repair cycle. Inhibition of photosynthesis is followed by impairment of other metabolic processes (Allakhverdiev et al, 2008; Murata et al, 2007; Wen et al, 2005).

At the transcript level, most of differentially expressed subunits of photosystems are induced in both tropical isolates WH8102 and WH8109 at 26°C (Figure 3.4). This could imply maximal utilisation of the available light conditions and faster growth. The induction of photosystem subunits could also be due to faster protein turnover to overcome thermal damage and thus better ability to cope with higher temperature. In strain WH8102, though a few genes at the protein level were upregulated at 26°C, overall relative abundance is not significantly different (Figure 3.3 and 3.4). In strain WH8102 photosynthetic yield, measured using PAM fluorometry was higher at 26°C which may be in response to increased expression of photosystem genes.

In the temperate strain CC9311 most transcripts of PSII subunits were downregulated at 26°C (Figure 3.4). The downregulation does not affect photosynthetic performance of PSII as indicated by the maintenance of photosynthetic yield measured using PAM fluorometry (Figure 3.2). The lowered expression of PSII may be due to increased efficiency of energy conversion. Alternatively, since the growth rate of strain CC9311 is lower at 26°C, the expression change may be due to lower requirement for energy conversion. Despite the downregulation of PSII, gene *ctpA* encoding a peptidase responsible for processing of D1 protein (core protein of PSII) is induced, possibly suggesting increased turnover of D1 (Rowland et al, 2010). Whilst most PSII genes were downregulated, the expression of genes encoding subunits of photosystem I (PSI) was mostly unchanged in strain CC9311 (Figure 3.4). This may suggest that strain CC9311 is more dependent on PSI for energy conversion at high temperature which is evident in the trend towards a higher overall protein abundance of PSI constituents (Figure 3.3). PSI subunit IV (*psaE*) was induced at 26°C. The PsaE subunit is involved in binding electron carriers to PSI. In the freshwater cyanobacterium *Synechocystis* PCC 6803, the PsaE subunit is involved in preventing the production of ROS (Jeanjean et al, 2008). The gene is also induced in tropical strains WH8102 and WH8109. Changes

in photosynthetic activity and/or energy utilisation may result in increased ROS at higher temperature, hence the increased expression of *psaE* gene may help the strains CC9311, WH8102 and WH8109 counter this increased stress.

The majority of ferredoxin genes were downregulated at 26°C in strain CC9311 (Figure 3.4). Ferredoxins are a key component of the photosynthetic electron transport and are involved in redox reactions in a variety of metabolic pathways. Ferredoxins are regulated by environmental conditions and are important in stress tolerance (Cassier-Chauvat and Chauvat 2014). The downregulation of ferredoxin genes correlates with the downregulation of components of PSII and other metabolic pathways and slower growth at 26°C. However, two ferredoxin genes and two flavodoxin genes were significantly induced. These may help maintain linear and increase cyclic electron transport around PSI under high temperature. In plants, PSI activity increases at high temperature and a lack of cyclic electron flow increases heat sensitivity (Sharkey and Zhang, 2010). Genes associated with the electron transport chain such as cytochrome b_6f complex and cytochrome c oxidase showed a differential expression in strain CC9311. Genes encoding subunits of cytochrome b_6f were downregulated whilst several genes encoding cytochrome c oxidase subunits were upregulated in strain CC9311. The induction of cytochrome c oxidase genes is consistent with a role of cytochrome c oxidase as terminal electron sink under low photosystem activity and in avoiding the over-reduction of the ETC. The increased expression of cytochrome c oxidase may also help maintain the proton gradient across the membrane, essential for both ATP synthesis as well as maintenance of membrane structure, particularly if high temperature increases membrane leakage (Fu et al, 2007; Lea-Smith et al, 2013; Sharkey and Zhang, 2010).

In the temperate strain BL107, transcripts of four genes encoding various forms of the core protein D1 and one gene encoding D2 were induced at 26°C relative to 22°C (Figure 3.4). Induced D1 genes (*psbA2*) encode an alternate form of the protein, D1.2. Two alternative forms D1.1 and D1.2 are exchanged during acclimation, with the D1.2 form predominating under less favorable conditions (Campbell et al, 1998).

Induction of D1.2 under high light intensities is observed in *Synechococcus* strain WH7803 as well in the freshwater cyanobacterium *Synechocystis* (Campbell et al, 1998; Garczarek et al, 2008). Induction of all four genes encoding D1.2 protein in strain BL107 at 26°C may be in response to over-excitation of PSII and the need for replacement of damaged D1.1 with D1.2 at high temperature. The gene encoding D1.2 remains unchanged in strains CC9311 and WH8102 whilst in WH8109, D1.2 genes are differentially expressed. Efficient functioning of PSII requires a continuous repair cycle involving turnover of D1 through proteolytic digestion and replacement with newly synthesised protein. Factors such as reactive oxygen species which inhibit polypeptide synthesis interfere with PSII repair cycle in cyanobacteria and higher plants (Murata et al, 2007; Nishiyama et al, 2004). Induction of all four copies of the *psbA2* gene at high temperature may help overcome the interruption of *de novo* D1 protein synthesis.

Mechanisms for photoprotection

In photosynthetic organisms, mechanisms to avoid photoinhibition and over-reduction of the electron transport chain are important under conditions that alter the balance between energy absorption and utilisation and increase the production of ROS. These mechanisms include energy dissipation (*isiA*), alternate electron flow through PSI and Mehler reaction (flavoproteins), ROS protection (superoxide dismutase, thioredoxin, peroxiredoxin, glutathione-related enzymes, high-light inducible (Hli) proteins) and photorespiration (He et al, 2001; Scanlan et al, 2009). Cross-signaling between high temperature and oxidative stress has been shown in plants and freshwater cyanobacteria (Allakhverdiev et al, 2008; Kotak et al, 2007; Rowland et al, 2010; Takahashi and Murata, 2008). The gene encoding Psb32, an integral membrane subunit of PSII, is induced at 26°C at the transcript level in three strains (BL107, CC9311 and WH8102; Figure 3.4). The homologous protein encoded gene *sll1390* in *Synechocystis* PCC6803 is involved in PSII assembly and repair and provides protection from oxidative stress (Wegener et al, 2011). Its role in the repair cycle and dimerisation of PSII monomers was also shown in the higher eukaryote *Arabidopsis* (gene TLP18.3) (Shi et al, 2012; Sirpio et al, 2007). The upregulation of the gene *psb32* at

26°C in three of the four strains (except strain WH8109), could imply an increased need for PSII repair at high temperature possibly due to oxidative stress. Several other genes associated with photoprotection are induced at 26°C (Figure 3.4). Among the four strains, the tropical strain WH8102 induced the most number of genes, including both the transcript and protein levels, involved in photoprotection. The mechanisms employed by the strain WH8102 include ROS-scavenging systems as well as flavoproteins which were induced at 26°C particularly at the protein level. Strain WH8109 induced genes involved in cyclic electron flow, ROS-scavenging and dissipation of excess energy (*isiA*). The protein encoded by *isiA* is known to be essential for high temperature tolerance in *Synechocystis* PCC6803. IsiA was shown to dissipate energy as heat and may quench reactive chlorophyll molecules (Havaux et al, 2005; He et al, 2001; Kojima et al, 2006; Scanlan et al, 2009). In temperate strain BL107 several genes involved in ROS-protection were differentially expressed at the transcript and proteomic level. In temperate strain CC9311, though most thioredoxin-related and glutathione peroxidase genes were repressed, genes encoding Hli proteins and flavoproteins involved in alternate electron flow were induced. Though the specific role of Hli proteins is unclear, it has been suggested that they may be involved in energy dissipation, chlorophyll binding and ROS-protection (Havaux et al, 2005; He et al, 2001). The induction of these genes suggest that high temperature results in a certain level of PS II over-reduction and oxidative stress, despite optimality in tropical strains, and the four *Synechococcus* strains employ several different mechanisms for protection.

Energy metabolism

Genes associated with carbon fixation (including RuBisCO and carbon concentrating mechanisms) and the Calvin cycle were repressed in the strain CC9311 at high temperature (Figure 3.4), correlating with its slower rate of growth. The activity of RuBisCO has been reported to be enhanced at high temperature (Fu et al, 2007), thus the downregulation of RuBisCO genes may be due to an increase in its efficiency. The downregulation of genes associated with carbon fixation on exposure to high temperature is observed in

Spirulina platensis strain C1 (Panyakampol et al, 2015). Studies in higher plants have reported a correlation between high temperature adaptation and deactivation of RuBisCO (Sharkey and Zhang, 2010). It is not known whether this strategy is employed by marine cyanobacteria. In the other strains, though a few genes associated with energy metabolism were differentially expressed, the trend is not as clear as in strain CC9311.

In strains WH8102 and CC9311, there was a significant level of downregulation of ATP synthase genes (Figure 3.4). This could indicate a change in the need for ATP and/or the efficiency of production. Lower carbon fixation could lower ATP utilisation (Sharkey and Zhang, 2010) and thereby reduce the need for expression of associated genes which may be the case for strain CC9311. In strain WH8102, lower expression of ATP synthase genes despite faster growth may be due to a change in efficiency of ATP production under optimal thermal conditions.

Fatty acid desaturases

High temperature increases the fluidity of lipid membranes in freshwater cyanobacteria (Los and Murata 2004). The level of lipid saturation and its associated role in high temperature acclimation is unclear in cyanobacteria and algae (Allakhverdiev et al, 2008). However, studies have suggested that there is an increase in saturated fatty acid content at higher temperatures (Balogi et al 2005; Los and Murata 2004). Balogi et al (2005) reported that thermotolerance in *Synechocystis* involved thylakoid membrane stabilisation with the accumulation of highly saturated fatty acyl chains. The cyanobacterium *Spirulina platensis* responds to high temperature fluidity changes by reducing the expression of desaturases (Hongsthong et al, 2009). Rowland et al (2010) reported a downregulation of desaturase genes in response to heat shock in *Synechocystis* strain PCC6803. Desaturases increase the degree of unsaturation of glycerol moiety fatty acids in glycerolipids. The gene *desC* encoding stearoyl-CoA desaturase (Δ9 desaturase) carries out the first step in desaturation through the introduction of a double bond at Δ9

position in stearic acid to produce oleic acid. Gene *desA* encoding Δ12 desaturase introduces the second double bond in oleic acid at (Los and Murata 1999). In the tropical strains, one desaturase gene was downregulated [$\log_2$ fold -0.85 in WH8102 (*desA*; SYNW0696) and -1.6 in WH8109 (Syncc8109_1775; possibly *desA* based on phylogenetic analyses (data not shown)]. In the temperate strain BL107, none were significantly differentially expressed. In strain CC9311, two genes encoding desaturases were downregulated [$\log_2$ fold -0.96 (sync_0396) and -0.67 (*desC*; sync_2791)] and one gene (*desC*; sync_2793) was upregulated (0.9 $\log_2$ fold). The downregulation of desaturase genes could reduce the level of unsaturation, thus possibly counteracting the high temperature-associated increase in membrane fluidity.

CONCLUSIONS

This study reports the influence of high temperature on the growth of *Synechococcus* strains CC9311 (clade I), WH8109 (clade II), WH8102 (clade III) and BL107 (clade IV). The environmental partitioning of these dominant clades is such that clades I and IV are restricted to temperate latitudes above 30°N and below 30°S with highest abundances at temperatures ranging from 10°C to 20°C. With higher relative abundances within tropical/ subtropical regions, clade II and III appears to have a higher temperature preference ranging from 20°C to 28°C (Sohm et al, 2015; Zwirglmaier et al, 2008). When grown at high temperature, temperate strains BL107 (clade IV) and CC9311 (clade I) reduce their rate of growth. The tropical strains WH8109 and WH8102 had a higher rate of growth at 26°C relative to 22°C which provides evidence for the higher temperature preference of the strains.

In plants and freshwater cyanobacteria, physiological and molecular responses to high temperature involve photosynthesis and carbon metabolism. The four marine *Synechococcus* strains used in this study varied in their expression of genes involved in photosynthesis and carbon metabolism.

At the molecular level the two temperate strains differ in their responses to high temperature. Temperate strain BL107 induced several copies of *psbA2* gene of PSII and repressed PS-associated genes. These may be mechanisms to cope with oxidative stress at high temperature through the increase in PSII repair and reduction in energy absorption. At high temperature, the temperate strain CC9311 downregulates most genes involved in energy conversion and production including PS, ATP synthase and carbon fixation which correlates with a greater reduction in growth rate.

At the molecular level, (sub)tropical strains WH8102 and WH8109 respond to high temperature with an increase in the expression of genes associated with PS I and II. This potentially enables faster growth in strains WH8109 and WH8102. Despite faster growth in strain WH8109, genes associated with PS, electron transport, carbon fixation and ATP synthesis either remain unchanged or have equal number of genes up- and down-regulated. In strain WH8102 at high temperature, induction of genes encoding photosynthetic proteins may be in response to faster protein turnover (as evidenced by the induction of chaperones and proteases) and thus enabling higher rate of growth as well as photosynthetic efficiency.

Since the four *Synechococcus* strains used in this study varied in their gene expression in response to temperature change, it is difficult to predict the molecular response profile for tropical and temperate strains in general. This would require examining more strains within each clade. The variability in gene expression amongst the strains may be because temperature affects multiple biological processes within the cell. The variability further highlights the functional diversity between strains and their specific adaptation to local niches. If temperature changes influence other variables such as the absorption/utilization of light and nutrient availability/utilization, strains adapted to specific conditions could respond very differently.

The ability to maintain higher rates of growth with increases in temperatures would enable tropical/ subtropical strains WH8102 and WH8109 to outcompete their temperate counterparts in regions where

their niches overlap. Thus a climate change associated rise in temperature would likely favor the growth of these tropical/sub-tropical clades over temperate strains and possibly expand their niche range but this would also be dependent on the changes in other environmental factors.

ACKNOWLEDGEMENTS

This work was supported by the Australian Research Council Discovery Grant DP110102718 and Laureate Fellowship FL40100021 to ITP, and an International Postgraduate Research Scholarship to DV from the Australian government. We thank D.J Scanlan and F.D. Pitt, University of Warwick, UK for the *Synechococcus* strains.

Supplementary information is provided in CD.

REFERENCES

Ahlgren N, Rocap G. (2006). Culture isolation and culture-independent clone libraries reveal new marine *Synechococcus* ecotypes with distinctive light and N physiologies. Appl Environ Microbiol 72: 7193-7204.

Ahlgren N, Rocap G. (2012). Diversity and distribution of marine *Synechococcus*: multiple gene phylogenies for consensus classification and development of qPCR assays for sensitive measurement of clades in the ocean. Front Microbiol 3: 1-24.

Allakhverdiev S, Kreslavski V, Klimov V, Los DA, Carpentier R, Mohanty P. (2008). Heat stress: an overview of molecular responses in photosynthesis. Photosynth Res 98: 541-550.

Balogi Z, Török Z, Balogh G, Jósvay K, Shigapova N, Vierling E, *et al*. (2005). "Heat shock lipid" in cyanobacteria during heat/light acclimation. Arch Biochem Biophys 436: 346-354.

Campbell D, Eriksson M, Oquist G, Gustafsson P, Clarke A. (1998). The cyanobacterium *Synechococcus* resists UV-B by exchanging photosystem II reaction-center D1 proteins. Proc Natl Acad Sci USA 95: 364-369.

Cassier-Chauvat C, Chauvat F. (2014). Function and regulation of ferredoxins in the cyanobacterium, *Synechocystis* PCC6803: recent advances. Life (Basel) 4: 666-680.

Castielli O, De la Cerda B, Navarro J, Hervás M, De la Rosa M. (2009). Proteomic analyses of the response of cyanobacteria to different stress conditions. FEBS Lett 583: 1753-1758.

Craig R, Beavis R. (2003). A method for reducing the time required to match protein sequences with tandem mass spectra. Rapid Commun Mass Spectrom 17: 2310-2316.

Craig R, Beavis R. (2004). TANDEM: matching proteins with tandem mass spectra. Bioinformatics 20: 1466-1467.

Dufresne A, Ostrowski M, Scanlan D, Garczarek L, Mazard S, Palenik B, *et al*. (2008). Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. Genome Biol 9: R90.

Fu F, Warner M, Zhang Y, Feng Y, Hutchins D. (2007). Effects of increased temperature and CO2 on photosynthesis, growth, and elemental ratios in marine *Synechococcus* and *Prochlorococcus* (cyanobacteria). J Phycol 43: 485-496.

Gammulla C, Pascovici D, Atwell B, Haynes P. (2010). Differential metabolic response of cultured rice (*Oryza sativa*) cells exposed to high and low-temperature stress. Proteomics 10: 3001-3019.

Garczarek L, Dufresne A, Blot N, Cockshutt A, Peyrat A, Campbell D, *et al*. (2008). Function and evolution of the *psbA* gene family in marine *Synechococcus*: *Synechococcus* sp. WH7803 as a case study. ISME J 2: 937-953.

Glatz A, Vass I, Los D, Vígh L. (1999). The *Synechocystis* model of stress: From molecular chaperones to membranes. Plant Physiol Biochem 37: 1-12.

Havaux M, Guedeney G, Hagemann M, Yeremenko N, Matthijs H, Jeanjean R. (2005). The chlorophyll-binding protein IsiA is inducible by high light and protects the cyanobacterium *Synechocystis* PCC6803 from photooxidative stress. FEBS Lett 579: 2289-2293.

He Q, Dolganov N, Björkman O. (2001). The high light-inducible polypeptides in *Synechocystis* PCC6803: expression and function in high light. J Biol Chem 276: 306-314.

Hongsthong A, Sirijuntarut M, Yutthanasirikul R, Senachak J, Kurdrid P, Cheevadhanarak S, *et al*. (2009). Subcellular proteomic characterization of the high-temperature stress response of the cyanobacterium *Spirulina platensis*. Proteome Sci 7: 33

Huang L, McCluskey M, Ni H, LaRossa R. (2002). Global gene expression profiles of the cyanobacterium *Synechocystis* sp. Strain PCC 6803 in response to irradiation with UV-B and white light. J Bacteriol 184: 6845-6858.

Huner N, Öquist G, Sarhan F. (1998). Energy balance and acclimation to light and cold. Trends Plant Sci 3: 224-230.

Jardillier L, Zubkov M, Pearman J, Scanlan D. (2010). Significant CO_2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. ISME J 4: 1180–1192.

Jeanjean R, Latifi A, Matthijs H, Havaux M. (2008). The PsaE subunit of photosystem I prevents light-induced formation of reduced oxygen species in the cyanobacterium *Synechocystis* sp. PCC 6803. Biochim Biophys Acta 1777: 308-316.

Kojima K, Suzuki-Maenaka T, Kikuchi T, Nakamoto H. (2006). Roles of the cyanobacterial *isiABC* operon in protection from oxidative and heat stresses. Physiol Plant 128: 507-519.

Kotak S, Larkindale J, Lee U, von Koskull-Döring P, Vierling E, Scharf K. (2007). Complexity of the heat stress response in plants. Curr Opin Plant Biol 10: 310-316.

Langmead B, Salzberg S. (2012). Fast gapped-read alignment with Bowtie2. Nat Methods 9: 357-359.

Lea-Smith D, Ross N, Zori M, Bendall D, Dennis J, Scott S, *et al*. (2013). Thylakoid terminal oxidases are essential for the cyanobacterium *Synechocystis* sp. PCC 6803 to survive rapidly changing light intensities. *Plant Physiol* 162: 484-495.

Li W. (1994). Primary productivity of prochlorophytes, cyanobacteria, and eucaryotic ultraphytoplankton: measurements from flow cytometric sorting. Limnol Oceanogr 39: 169–175.

Los D, Murata N. (2004). Membrane fluidity and its roles in the perception of environmental signals. Biochim Biophys Acta 1666: 142-157.

Mackey K, Paytan A, Caldeira K, Grossman A, Moran D, McIlvin M, *et al*. (2013). Effect of temperature on photosynthesis and growth in marine *Synechococcus* spp. Plant Physiol 163: 815-829.

Magoc T, Wood D, Salzberg S. (2013). EDGE-pro: Estimated Degree of Gene Expression in Prokaryotic Genomes. Evol Bioinform 9: 127-136.

Mirzaei M, Pascovici D, Keighley T, George I, Voelckel C, Heenan P, *et al*. (2011). Shotgun proteomic profiling of five species of New Zealand *Pachycladon*. Proteomics 11: 166-171.

Mirzaei M, Soltani N, Sarhadi E, Pascovici D, Keighley T, Salekdeh G, *et al*. (2012). Shotgun proteomic analysis of long-distance drought signaling in rice roots. J Proteome Res 11: 348-358.

Moore L, Goericke R, Chisholm S. (1995). Comparative physiology of *Synechococcus* and *Prochlorococcus*: influence of light and temperature on growth, pigments, fluorescence and absorptive properties. Mar Ecol Prog Ser 116: 259-275.

Morel F, Rueter J, Anderson D, Guillard R. (1979). Aquil: A chemically defined phytoplankton culture medium for trace metal studies. J Phycol 15: 135-141.

Murata N, Takahashi S, Nishiyama Y, Allakhverdiev S. (2007). Photoinhibition of photosystem II under environmental stress. Biochim Biophys Acta 1767: 414-421.

Neilson K, George I, Emery S, Muralidharan S, Mirzaei M, Haynes P. (2014). Analysis of rice proteins using SDS-PAGE shotgun proteomics. In: Jorrin-Novo JV KS, Weckwerth W, Wienkoop S (ed.). Plant Proteomics: Methods and Protocols, Methods in Molecular Biology. Humana Press, pp 289-302.

Nishiyama Y, Allakhverdiev S, Yamamoto H, Hayashi H, Murata N. (2004). Singlet oxygen inhibits the repair of photosystem II by suppressing the translation elongation of the D1 protein in *Synechocystis* sp. PCC 6803. Biochemistry (Mosc) 43: 11321-11330.

Paerl R, Johnson K, Welsh R, Worden A, Chavez F, Zehr J. (2011). Differential distributions of *Synechococcus* sub-groups across the California current system. Front Microbiol 2:59.

Palenik B, Brahamsha B, Larimer F, Land M, Hauser L, Chain P, *et al*. (2003). The genome of a motile marine *Synechococcus*. Nature 424: 1037-1042.

Palenik B, Ren Q, Dupont C, Myers G, Heidelberg J, Badger J, *et al*. (2006). Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment. Proc Natl Acad Sci USA 103: 13555-13559.

Palenik B, Ren Q, Tai V, Paulsen I. (2009). Coastal *Synechococcus* metagenome reveals major roles for horizontal gene transfer and plasmids in population diversity. Environ Microbiol 11: 349-359.

Panyakampol J, Cheevadhanarak S, Sutheeworapong S, Chaijaruwanich J, Senachak J, Siangdung W, *et al*. (2015). Physiological and transcriptional responses to high temperature in *Arthrospira* (*Spirulina*) *platensis* C1. Plant Cell Physiol 56: 481-496.

Partensky F, Vaulot D. (1999). Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. Bulletin de l'Institut océanographique de Monaco 19: 457-475.

Rhee J, Ki J, Kim B, Hwang S, Choi I, Lee J. (2012). HspA and HtpG enhance thermotolerance in the cyanobacterium, *Microcystis aeruginosa* NIES-298. J Microbiol Biotechn 22: 118-125.

Rowland J, Pang X, Suzuki I, Murata N, Simon W, Slabas A. (2010). Identification of components associated with thermal acclimation of photosystem II in *Synechocystis* sp. PCC6803. PLOS ONE **5:** e10511.

Sabehi G, Feingersch R, Ferriera S, Johnson J, Kravitz S, Beeson K, *et al*. (2012). *Synechococcus* sp. WH8109, complete genome. NCBI Reference Sequence: NZ_CP006882.1.

Scanlan D. (2012). Marine picocyanobacteria. In: Whitton B (ed.). Ecology of Cyanobacteria II: Their diversity in space and time. Springer Science + Business Media, pp 503-533.

Scanlan D, Ostrowski M, Mazard S, Dufresne A, Garczarek L, Hess W, *et al*. (2009). Ecological genomics of marine picocyanobacteria. Microbiol Mol Biol Rev 73: 249-299.

Sharkey T, Zhang R. (2010). High temperature effects on electron and proton circuits of photosynthesis. J Integr Plant Biol 52: 712-722.

Shi L, Hall M, Funk C, Schröder W. (2012). Photosystem II, a growing complex: Updates on newly discovered components and low molecular mass proteins. Biochim Biophys Acta 1817: 13-25.

Sirpio S, Allahverdiyeva Y, Suorsa M, Paakkarinen V, Vainonen J, Battchikova N, *et al*. (2007). TLP18.3, a novel thylakoid lumen protein regulating photosystem II repair cycle. Biochem J 406: 415-425.

Six C, Thomas J, Garczarek L, Ostrowski M, Dufresne A, Blot N, *et al*. (2007). Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. Genome Biol 8.

Sohm J, Ahlgren N, Thomson Z, Williams C, Moffett J, Saito M, *et al*. (2015). Co-occurring *Synechococcus* ecotypes occupy four major oceanic regimes defined by temperature, macronutrients and iron. ISME J 1-13.

Su Z, Mao F, Dam P, Wu H, Olman V, Paulsen I, *et al*. (2006). Computational inference and experimental validation of the nitrogen assimilation regulatory network in cyanobacterium *Synechococcus* sp. WH 8102. Nucleic Acids Res 34: 1050-1065.

Suzuki I, Simon W, Slabas A. (2006). The heat shock response of *Synechocystis* sp. PCC 6803 analysed by transcriptomics and proteomics. J Exp Bot 57: 1573-1578.

Tai V, Palenik B. (2009). Temporal variation of *Synechococcus* clades at a coastal Pacific Ocean monitoring site. ISME J 3: 903-915.

Takahashi S, Murata N. (2008). How do environmental stresses accelerate photoinhibition? Trends Plant Sci 13: 178-182.

Tanaka N, Nakamoto H. (1999). HtpG is essential for the thermal stress management in cyanobacteria. FEBS Lett 458: 117-123.

Tandeau de Marsac N, Cohen-Bazire G. (1977). Molecular composition of cyanobacterial phycobilisomes. Proc Natl Acad Sci USA 74: 1635-1639.

Tatusov R, Fedorova N, Jackson J, Jacobs A, Kiryutin B, Koonin E, *et al.* (2003). The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4:41.

Waterbury J, Watson S, Guillard R, Brand L. (1979). Widespread occurrence of a unicellular, marine, planktonic, cyanobacterium. Nature 277: 293-294.

Wegener K, Bennewitz S, Oelmüller R, Prakasi H. (2011). The Psb32 protein aids in repairing photodamaged photosystem II in the cyanobacterium *Synechocystis* 6803. Molecular Plant 4: 1052-1061.

Wen X, Gong H, Lu C. (2005). Heat stress induces an inhibition of excitation energy transfer from phycobilisomes to photosystem II but not to photosystem I in a cyanobacterium *Spirulina platensis*. Plant Physiol Biochem 43: 389-395.

Wessel D, Flügge U. (1984). A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. Anal Biochem 138: 141-143.

Yu A, Houry W. (2007). ClpP: a distinctive family of cylindrical energy-dependent serine proteases. FEBS Lett 581: 3749-3757.

Zwirglmaier K, Jardillier L, Ostrowski M, Mazard S, Garczarek L, Vaulot D, *et al*. (2008). Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. Environ Microbiol 10: 147-161.

Zybailov B, Mosley A, Sardiu M, Coleman M, Florens L, Washburn M. (2006). Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. J Proteome Res 5: 2339-2347.

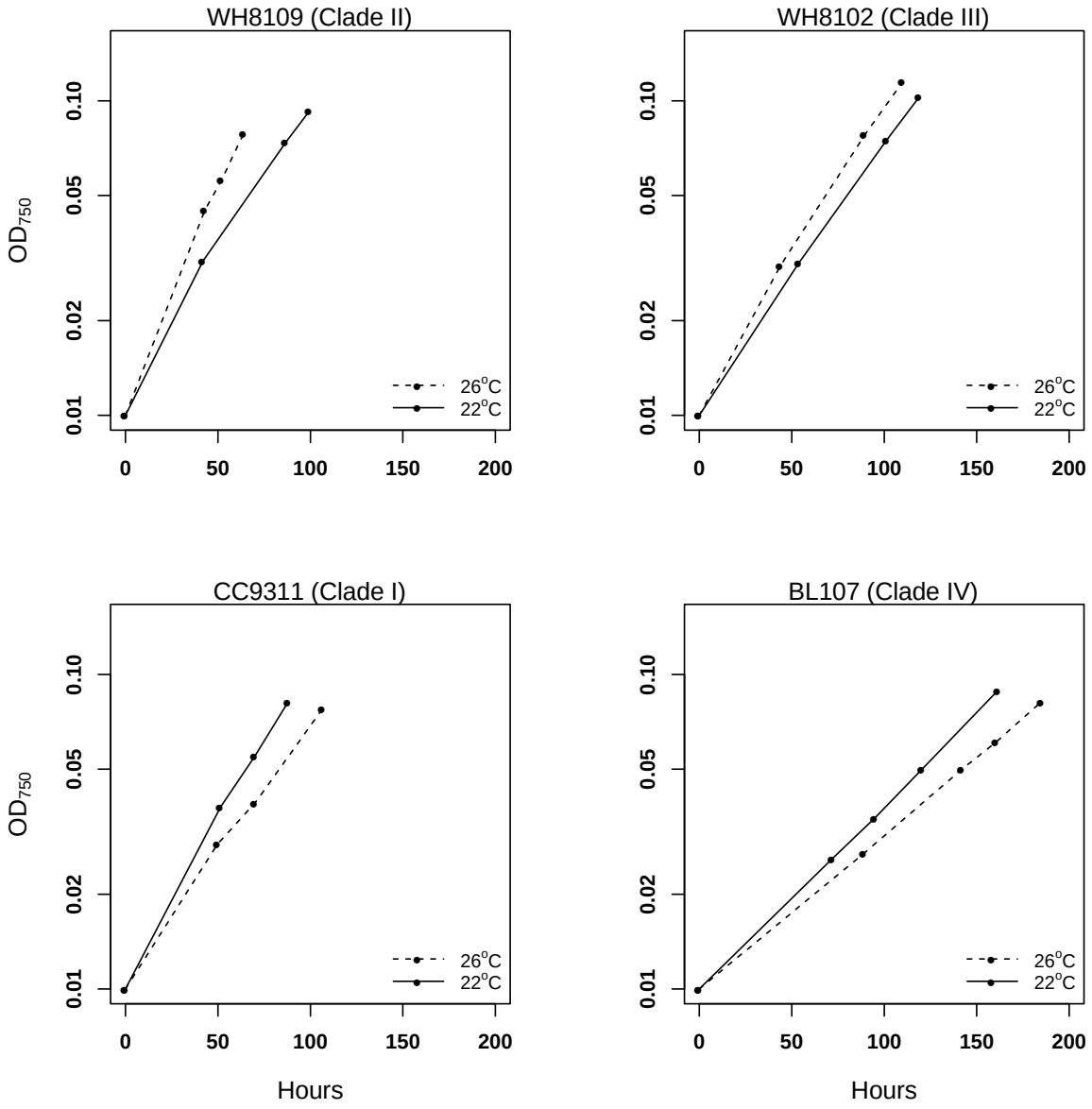


Figure 3.1: Growth physiology of marine *Synechococcus* isolates at 26°C and 22°C. Isolates CC9311 (clade I), WH8109 (clade II), WH8102 (clade III) and BL107 (clade IV). The *x*-axis represents growth period in hours and the *y*-axis represents optical density at 750 nm. Dashed lines represent growth at 26°C and solid lines represent growth at 22°C. Error bars, though included, are smaller than symbols and thus not visible.

Table 3.1: Growth physiology of *Synechococcus* strains grown at 22°C and 26°C

Clade	Strains	Average growth rate ±SD
I	CC9311 22°C	0.58 ± 0.015
	CC9311 26°C	0.46 ± 0.039 **
II	WH8109 22°C	0.47 ± 0.011
	WH8109 26°C	0.65 ± 0.039 **
III	WH8102 22°C	0.47 ± 0.015
	WH8102 26°C	0.53 ± 0.015 **
IV	BL107 22°C	0.34 ± 0.014
	BL107 26°C	0.27 ± 0.011 **

Statistical significance of difference represented as ** (p <0.01)

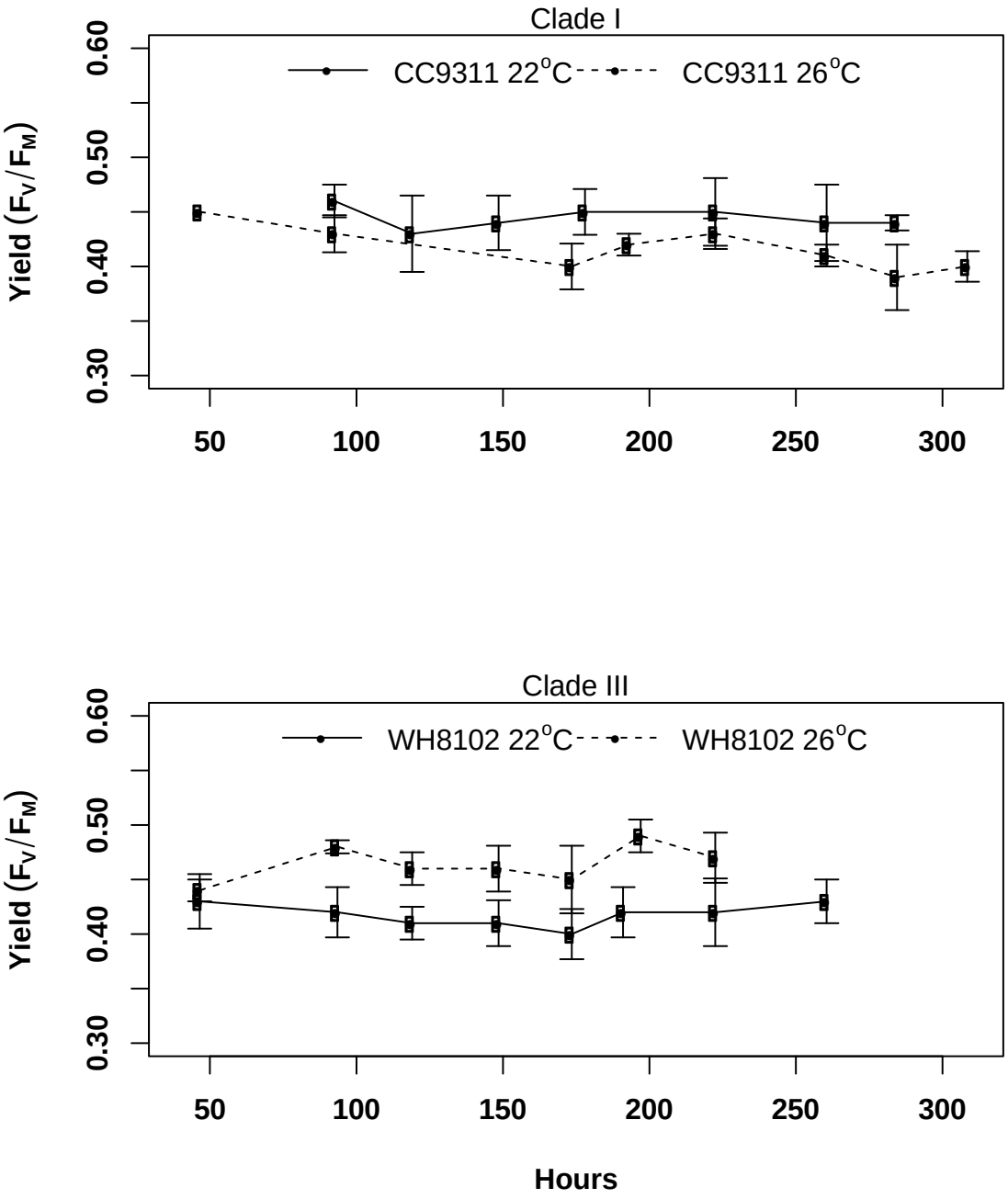
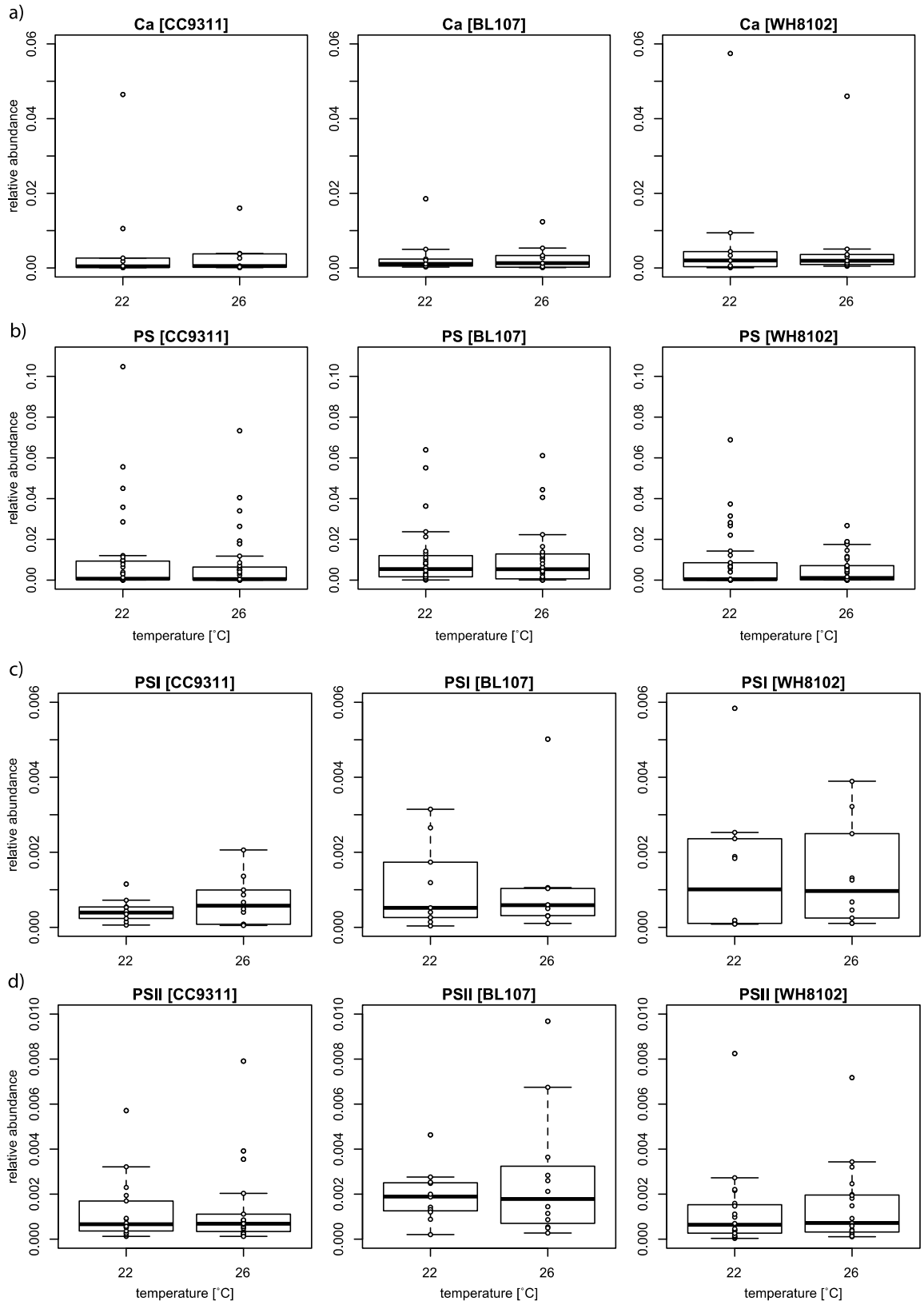


Figure 3.2: Photosystem II photosynthetic yield in marine *Synechococcus* isolates CC9311 (clade I) and WH8102 (clade III) at 26°C and 22°C. The *x*-axis represents growth period in hours and the *y*-axis represents mean quantum yield (F_v/F_m). Dashed lines represent growth at 26°C and solid lines represent growth at 22°C.


a)
Ca [CC9311]
Ca [BL107]
Ca [WH8102]
relative abundance
0.00
0.02
0.04
0.06
22
26
b)
PS [CC9311]
PS [BL107]
PS [WH8102]
relative abundance
0.00
0.02
0.04
0.06
0.08
0.10
22
26
temperature [˚C]
c)
PSI [CC9311]
PSI [BL107]
PSI [WH8102]
relative abundance
0.000
0.002
0.004
0.006
22
26
temperature [˚C]
d)
PSII [CC9311]
PSII [BL107]
PSII [WH8102]
relative abundance
0.000
0.002
0.004
0.006
0.008
0.010
22
26
temperature [˚C]

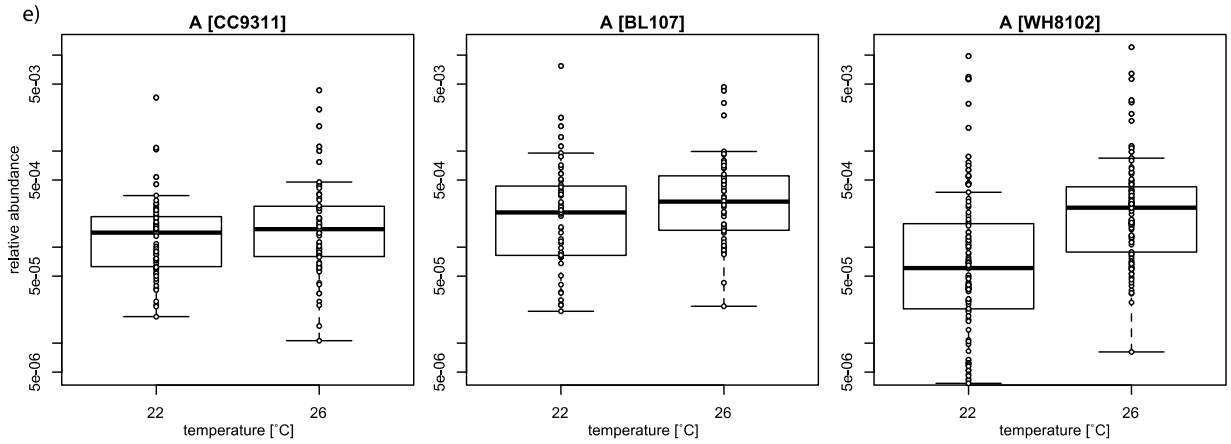


Figure 3.3: Relative abundance (NSAF) of proteins associated with a). carboxysomes (Ca), b). phycobilisomes (PS), c). photosystem I (PSI), d). photosystem II (PSII) and e). membrane transport (A) in marine *Synechococcus* isolates CC9311 (clade I), WH8102 (clade III) and BL107 (clade IV) when grown at 26°C and 22°C.

Gene/ Protein name	BL107 log2 fold change	CC9311 log2 fold change	WH8102 log2 fold change	WH8109 log2 fold change
Phycobilisomes				
APC, alpha-B chain	-	-	0.66	0.77
APC, alpha chain	-	-	-	0.59
APC-like protein	-	0.84	-	-
Core linker polypeptide (APC)	-#	-	-*	-
R-PC IIC alpha/beta subunit	-	-	-#(2)	-
Putative PB:PC lyase	-	-	-	-0.77
PCB:Cys-84 beta-PC lyase	-	-	-*	-
PEB Cys-84 alpha-PC lyase, *rpcE/rpcF*	-	-0.54	-	-0.52 (2)
PEB:Cys-84 alpha-R-PC-V lyase-isomerase	-0.50	-	-	-
C-PE I, alpha/beta subunits	-	> -0.89 (2)	0.72 #(2)	-
C-PE II, alpha/beta chains	-	> -0.90 (2)	-#(2)	-
Putative PEB:PE lyase, *cpeT*	-0.58	-	-*	-
PEB:Cys-82 alpha-PE lyase, *cpeY/cpeZ*	> -0.55 (2)	> 0.58 (2)	> 0.62 (2)	-
Putative PB:C-PE II lyase	-0.70	-	> 0.88 (2)	0.54
PEB:Cys-83 alpha-PE II lyase-isomerase	1.16	0.58	-	-
Rod linker polypeptide (C-PE I)	-	0.57	-	-
Rod linker polypeptide (C-PE II)	-0.92 *	0.61 #	-	-
Rod linker polypeptide (C-PE II)	-	-0.98	0.51	0.68
LrC (C-PE I), *cpeE*	-	-	0.54	-
LrC (C-PE II)	-	-	0.51	-
LrC, CpcG1/CpcG2/CpcG3	-0.55	1.68	-*	0.67
PEB:Cys-82 beta-PE lyase	-	-	-*	-
putative PE:PEB lyase	-#	-	-	-
PEB:ferredoxin oxidoreductase, *pebB*	-0.87	-0.63	-	-0.57
PCB:ferredoxin oxidoreductase	-	-	-*	-
Putative PB lyase	-	-	-#	-
PBP asparagine methyltransferase	-	-#	-*	-
PBP lyase or activator	-#	-	-	-
Photosystems				
PSI subunit X	-	-	-	0.71
PSI subunit IV (*psaE*)	-	0.69	0.67	0.50
PSI *psaF*	-	-0.59 #	-	0.50
PSI subunit II (*psaD*)	-	-	-*	0.86
PSI subunit VIII	-	-	-	0.52
PSI assembly protein *ycf4*	-	-	-*	-
PSII *psbI*	-	-	-	0.56
PSII 12 kDa extrinsic protein *psbU*	-	-	-	0.67
PSII protein *psbZ* (*ycf9*)	-	-	-	0.67
PSII *psbX* protein	-	-	-	0.64
PSII *psbH*	-	-	0.50	0.55
PSII *psbJ*	-*	-	0.70 *	-
PSII *psbM*	-	-0.56	0.93	-
PSII CyanoQ	-	-0.90	-	0.83
PSII biogenesis protein *psb29*	-	-1.16	-	-0.58
PSII chlorophyll-binding protein CP47	-	-0.55	-	-
PSII *psbY* protein	-	-0.68	-	0.53
PSII D2 protein	0.51	> -0.55 (3)	-*	-
PSII Mn-cluster assembly protein *psb27*	-	-1.21	-	-
PSII CyanoP	-	-0.52	-	-
PSII T protein	-	-0.59	-	-
PSII *psb28* protein	-0.60	-0.99 #	-	-
PSII protein D1.2 (*psbA2*)	> 0.59 (4)	-	-	> -0.64 (2)
PSII protein D1.2 (*psbA2*)	-	-	-	0.61
PSII PsbY-like protein	-	0.87	-	-
PSII stability/assembly factor	-#	-	-	-
Electron transport chain:				
2Fe-2S ferredoxin	-	> 1.06 (2)	0.76	> -0.53 (3)
2Fe-2S ferredoxin	-	> -0.61 (3) #(2)	-#	> 0.57 (2)
3Fe-4S ferredoxin	-	-0.65	-	-
Ferredoxin--NADP$^+$ reductase (FNR)	-	-0.84	0.64	-
Cyt b6-f complex subunits	-0.68	> -0.54 (2)	-*	0.58
Cyt c553 (*petJ*)	-	0.83	-#	-
Cyt cM	0.58	-	-	0.51
Plastocyanin precursor, *petE*	-	-	0.81	-
Cyt c oxidase subunits	0.53 *	> 0.62 (4) *	1.00	-
Cyt-c oxidase subunits		-0.56	-0.60	-
Putative cyt c oxidase assembly protein	0.81	-	-1.33	-
Fatty acid desaturase				
Δ12 desaturase (*desA*)	-	-	-0.85	-1.60
Δ9 desaturase (*desC*)	-	0.90	-	-
Δ9 desaturase (*desC*)	-	>-0.67 (2)	-	-

Gene/ Protein name	BL107 log2 fold change	CC9311 log2 fold change	WH8102 log2 fold change	WH8109 log2 fold change
ATP synthesis				
ATP synthase A chain	-	-0.54	-	-
ATP synthase B'	-	-1.50	-0.71	-
ATP synthase C chain	-	-1.90	-	-
ATP synthase delta chain	-	-1.79	-1.62	-
ATP synthase protein I	-	-1.09	-0.57	-
ATP synthase subunit B	-	-1.60	-1.45	-
ATP synthase alpha chain	-	-1.17	-#	-
ATP synthase beta chain	-	-1.80	-	-
ATP synthase epsilon chain	-	-1.26 #	-#	-
ATP synthase gamma chain	-	-1.18 #	-#	-
Carbon fixation				
Carboxysomal carbonic anhydrase	-	-0.85	-*	-
Carboxysome shell peptide	-0.52 #	> -0.83 (5) #	-#	-
RuBisCO, large subunit	-	-1.78	-	-
RuBisCO, small subunit	-#	-1.49	-	-
Phosphoribulokinase	-	-1.83	-#	-
FBPase/SBPase	-0.66	-1.09	-	-
Fructose-1,6-bisphosphate aldolase, class II	-	-1.02 #	-#	-
Transketolase	-	-0.81	-	-
Glyceraldehyde-3-phosphate dehydrogenase	-	-1.75	0.78 #	-
Ribulose-phosphate 3-epimerase	-	-*	-#	-
Phosphoglycerate kinase	-#	-	-	-
Chaperones and Proteases:				
Chaperone *dnaJ*	0.73 *	-*	-	-0.78
Molecular chaperone sHSP	0.64	2.18	-*	-
Chaperonin *groEL*	0.53 *	-*	0.69	0.58
Chaperonin *groES*	0.50	-0.55	-	0.53
Molecular chaperone *dnaK*	-	2.17 *	-~	-
Chaperone *clpB*	-	0.60 *	-	-
Chaperone protein *htpG*	-	-*	0.84 *	-
Putative protease *clpB1* /*clpB2*	0.71		-*(2)	
Clp protease adaptor protein *clpS*	-	-	-	0.56
Clp protease (*clpP* subunits)	-0.62 #	-	> 0.59 (3) *(2)	-
Cell division protein *ftsH1*	-	-0.60	-	-
Cell division protein *ftsH3*	-	0.53 *	-*	-
Cell division protein *ftsH*	-	-	-	-0.61
PS II D1 C-terminal peptidase *ctpA*	-#	0.99	-	-
PS II D1 repair protein *ftsH2*	-	-0.60	-	-
Membrane protease subunit	-*	-	0.63 *	-
C-terminal protease	-*	-	*	-
Serine protease	0.74 *#	-	-*	0.74
Possible metal-dependent protease	-	0.88	-	-
Possible metal-dependent peptidase	-0.55	> -0.61 (2)	-*	-0.59
Aminopeptidase	-	-	1.65 *	-
Peptidase family	-	-	0.95 *(2)	0.75
Signal peptidase I	-	-	-	-0.64
O-sialoglycoprotein endopeptidase	-	-	-0.77 *	-
Putative oligopeptidase A	-	-	-*	
Photoprotection:				
psb32	0.51	0.62	0.70	-
isiA (PSII Fe-stress-induced Cbp)	-	-	-	1.15
Flavoprotein (Mehler reaction)	-*	0.58	-*(2)	-
Flavodoxin	-0.74	> 0.95 (2)	n/a	0.63
Superoxide dismutase, *sodN*	-	-#	1.00	-
NTR	0.50	-	-*	0.77
FTR, variable chain	-	-0.69	-*	-
Thioredoxin	-#	-0.66	-	-
Thioredoxin	-	0.53	-	-
Peroxiredoxin	-#	-1.39	0.93 *(2)	-0.80
Rubredoxin	-	-0.59	-*	-
Starvation-inducible Dbp (*dspA*)	-	-	-	0.88
Glutathione peroxidase	-#	-1.28	-*	-
Glutathione synthetase	-#	-0.53	-	-
Glutathione reductase	-*	0.75	-*	-
Glutathione S-transferase			-*(4)	
High light inducible protein	0.50	> 0.50 (4)	-*	0.53
High light inducible protein	-0.50	-0.71	-	> -0.55 (2)

Protein names that are repeated diferentiate copies of genes, some of which are upregulated and some downregulated

Complete lists of differentially expressed genes are provided in Supplementary tables 3.3 - 3.6

Lrc - rod-core linker polypeptide; APC - allophycocyanin; PEB - phycoerythrobilin; PE - phycoerythrin; PC - phycocyanin; PB - phycobilin; PBP - phycobiliprotein; PSI - photosystem I; FBPase/SBPase - Fructose-1,6-biphosphatase/sedoheptulose-1,7-biphosphate phosphatase; NTR - NADPH-dependent thioredoxin reductase; Cbp - Chlorophyll-binding protein; Dbp - DNA-binding protein; FTR - Ferredoxin-thioredoxin reductase; Cyt - cytochrome; PSII - photosystem II

Figure 3.4: Selected list of genes differentially expressed at the trancriptional level in marine *Synechococcus* strains CC9311 (clade I), WH8109 (clade II), WH8102 (clade III) and BL107 (clade IV) at 26°C relative to 22°C ($p<0.05$). Expression changes are given as $\log_2$ fold change (upregulation: red; downregulation: blue). Differential expression at the translational (protein) level is indicated by * (induction) and # (repression). Numbers in brackets represent the number of differentially expressed genes.

Chapter 4:

Seasonal variation in the marine cyanobacterial community in the Sydney Harbour estuary and the influence of temperature on community structure

Title

Seasonal variation in the marine cyanobacterial community in the Sydney Harbour estuary and the influence of temperature on community structure

Authors

Deepa Varkey, Sophie Mazard, Thomas Jeffries, Justin Seymour, Ian Paulsen and Martin Ostrowski

Affiliation

Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, NSW, Australia 2109

Corresponding author:

Prof Ian T Paulsen, ian.paulsen@mq.edu.au, Telephone: +61 2 9850 8152, Fax: +61 2 9850 8313.

Running title:

Photosynthetic microbial populations from Sydney Harbour estuary

Conflict of interest

The authors declare no conflict of interest.

The Sydney Harbour estuary is a well-mixed system with significant marine influence and occasional large inflows of freshwater. The occurrence and distribution of cyanobacteria in the estuary and other coastal systems in Australia are poorly understood. Cyanobacteria could be a major contributor to primary production in the estuary, thus it is important to understand their role, distribution and factors influencing them. This study examined the marine *Synechococcus* community composition within the estuary using the phylogenetic marker, *petB* gene. The relative abundance of *Synechococcus* as a proportion of the total planktonic microbial community was determined with universal 16S rDNA marker. Cyanobacteria (average relative abundance of 12.7% across sampling sites compared to 10.2% for chloroplast) potentially contribute significantly to primary production in the Sydney Harbour estuary. The comparison of cyanobacterial communities using the 16S rDNA marker indicates that members of marine *Synechococcus* sub-cluster 5.1 were the predominant cyanobacteria throughout the estuary. The comparison between February (late summer) and September (early Spring) in 2013 showed a notable variation in the abundance and composition of the *Synechococcus* population. In February, the abundance of *Synechococcus* was significantly higher than in September. Clade II was the dominant lineage in February whilst in September clade I accounted for a higher proportion of the *Synechococcus* community. *Synechococcus* communities within the estuary were significantly different than those found immediately offshore, indicating that the estuarine community is geographically separated or subjected to different selective pressure. Temperature has a large influence on the seasonal partitioning of the *Synechococcus* community within the estuary. However, the observed variation in nutrients, salinity and mixing as well as other non-measured factors may also influence the changes in the *Synechococcus* population.

Keywords**:** Cyanobacteria/*Synechococcus*/estuary/EAC/coastal

INTRODUCTION

Estuarine environments are ecologically, culturally and economically important. These ecosystems are biologically diverse and important conservation areas. Industrialisation and urbanisation, however,

threaten these ecosystems through the alteration of water flow, increased input of nutrients, sediments and other pollutants and habitat degradation (Caroppo, 2015).

Sydney Harbour estuary (historically referred to as Port Jackson estuary), is located on the south-eastern coast of New South Wales, Australia. The estuary was formed about 17,000 years ago by the flooding and drowning of a river valley due to rising sea levels. It is 30 km in length draining approximately 500 km^2 (Hatje et al, 2003). Port Jackson was colonised by British settlers on 26th of January 1788 and since then the population of the Port Jackson catchment has risen from a few thousand and become the urban centre of a city with over 4.6M residents with accompanying increases in urbanisation and industrial activity. The irregular bathymetry comprises areas of deep (45 m) and shallow (<3.5 m) regions, with an average depth of 13 m, and large, shallow bays that connect to the main channel of the estuary (Das et al, 2000; Hatje et al, 2003; Hedge et al, 2013). With no large river systems that connect to the estuary, input of freshwater is low, occurring only during periods of rainfall which results in occasional high inflow of storm water runoff (Hose et al, 2005). Tidal turbulence caused by semi-diurnal tides results in well-mixed waters in the estuary. The waters are saline throughout the estuary except during large influx of storm water runoff wherein upstream regions of the estuary could be less saline (Das et al, 2000; Hose et al, 2005; Lee et al, 2011). Tidal flushing times of the estuary vary along its length, i.e. rapid flushing occurs at the mouth of the harbour whilst in regions furthest upstream flushing time has been estimated to be up to 225 days (Das et al, 2000).

Anthropogenic influence on the harbour catchment and estuary is significant due to its location surrounded by the Sydney metropolitan area. Historically, the estuary was used for disposal of industrial and urban effluent (Hose et al, 2005; McCready et al, 2000). Direct disposal of waste into the estuary continued until the 1970s. High concentrations of heavy metals and chemicals remain in the sediments particularly in the upper regions of the estuary (Hatje et al, 2003; McCready et al, 2000; McCready et al, 2003). Currently, the harbour is a busy shipping port with high maritime traffic, an important commercial fishery (east of the Sydney Harbour Bridge) as well as a popular recreational

area for swimming, boating and fishing (Hose et al, 2005). Proximity to urbanised and industrialised regions continues to affect the water quality of the estuary particularly through sewerage overflows and storm water contamination (Hose et al, 2005). Such overflows cause sediment influx, nutrient loading, increased suspended solids as well as heavy metal and chemical contamination (Beck and Birch, 2012). Current estimates show relatively high annual loadings of heavy metals Zn (69 t), Pb (39 t) and Cu (26 t) and total suspended solids (49,239 t) as well as elevated levels of pharmaceuticals and pesticides (Birch et al, 2015a; Birch et al, 2015b). Nutrient loading results in eutrophication and diminished oxygen concentrations (Das et al, 2000). There is also more suspended particulate matter upstream rather than at the mouth of the harbour. Suspended particulate matter increases the turbidity, thus affecting light quality and intensity (Hatje et al, 2003). These altered environmental conditions can have a significant impact on the organisms that inhabit the estuary, particularly the primary producers which in turn affects higher trophic levels.

With significant oceanic influence and periodic freshwater input from heavy rainfall events, the heterogeneous ecosystem of the Sydney Harbour estuary can support diverse prokaryotic and eukaryotic communities. Studies on the Sydney Harbour estuary have focused on water quality and pathogenic microbes in the estuary whilst those on phytoplankton and zooplankton are limited (Hedge et al, 2013; Hose et al, 2005). The photoautotrophic plankton species composition and the changes in community structure in response to environmental conditions have not been explored for the Sydney Harbour estuary.

Phytoplankton of the genus *Synechococcus* are important components of estuarine, coastal and open ocean ecosystems across the Earth. These photosynthetic prokaryotes are highly abundant and ubiquitous, contributing significantly to global primary productivity and nutrient cycling (Jardillier et al, 2010; Li, 1994, Scanlan et al, 2009). As primary producers at the base of food webs, it is important to understand the cyanobacterial diversity, community structure and ecology. Cyanobacteria in sub-tropical and temperate estuarine areas vary in abundance and contribution to the phytoplankton community (Murrell and Lores, 2004; Ning et al, 2000; Phlips et al, 1999). Environmental parameters

influencing the spatial and temporal variations in the abundance and composition of cyanobacteria have been reported in other estuarine environments with phycocyanin-rich *Synechococcus* lineages more abundant in less saline regions and phycoerythrin-rich lineages abundant at the marine end of the estuary (Cai et al, 2010; Murrell and Lores, 2004; Phlips et al, 1999).

This study explores the diversity of cyanobacteria within the Sydney Harbour estuary. The study also aims to determine the spatial and temporal changes in the genetic composition of cyanobacterial populations in the context of the environmental parameters that influence the structure and function of pelagic communities in this estuarine ecosystem.

MATERIALS AND METHODS

Sample collection

Samples were collected from 30 sites along the length of the Sydney Harbour estuary (30 km; Figure 4.1). Water conditions ranged from brackish in the uppermost reaches of the estuary to marine at mouth of the harbour. Samples were obtained from two time periods in 2013, February (late summer, moderate-high rainfall >50 mm) and September (early spring, low rainfall). Water samples (2 l) from a depth of ~0.5 m at each site were filtered onto 0.2 µm polycarbonate membrane filters (Millipore, Australia) and stored at -20°C until further processing. Physical and chemical parameters such as temperature, pH, salinity, turbidity, dissolved oxygen and nutrient concentrations were determined as in Jeffries et al, 2016.

Samples from the East Australian Current (EAC) were collected during the RV *Southern Surveyor* voyage SS2010 v09 that took place from 15 to 31 October 2010. Samples from the Port Hacking reference station (PH4) were collected in February/September 2010. The mixed layer depth (MLD) was calculated as the depth within the thermocline where temperature decreased from 0.5˚C compared to the surface temperature. Data profiles were obtained from the Australian ocean data network portal (portal.aodn.org.au). Water samples (2 l) from surface waters (EAC) or depth profiles from the PH4

long-term reference station, located above the 100m isobath, 3.2 miles offshore, and 20 miles south of Sydney Harbour mouth, were filtered onto 0.22 µm diameter pore size Sterivex filters (Millipore, Australia) flash frozen in liquid nitrogen and stored at -80˚C until processing.

Nucleic acid extraction

DNA was extracted from the polycarbonate membrane and Sterivex filters with a PowerWater DNA isolation kit (MO BIO, USA) using bead beating for cell lysis. DNA in nuclease-free water was stored at -80°C until used for amplification of phylogenetic markers. The concentration of DNA was measured using a NanoDrop 2000 UV spectrophotometer (Thermo Scientific, USA).

Library preparation

The conserved genetic marker *petB* was used to determine species composition of the picocyanobacterial community within the Sydney Harbour estuary. This core genome marker provides enhanced taxonomic resolution to distinguish phylogenetic lineages of marine *Synechococcus* (Mazard et al, 2012). The gene *petB*, encoding the cytochrome b_6 subunit of the cytochrome b_6f complex, also allows accurate automated alignment (Mazard et al, 2012). Specific primers for *petB* were attached with sequencing adaptors and indices that enabled multiplexed sequencing. These indices were designed based on the Nextera XT index kit (Illumina, USA) suggested in the 16S rDNA metagenomic library preparation guide (Supplementary Table 4.1). The *petB* sequences from the EAC voyage and from the PH4 reference station were generated with the same primers using 454 sequencing (500 bases) (Ostrowski et al, *submitted*). For the Sydney Harbour samples, generation of amplicon libraries for bacterial 16S rDNA genes, using universal eubacterial primers 926F (5'-aaactYaaaKgaattgacgg-3') and 1392R (5'-acgggcggtgtgtRc-3'), from the same samples is described in Jeffries et al, 2016.

Amplicon libraries of the *petB* marker for each site were prepared in 50 µl PCR reactions, Each reaction contained 20 pmol each of forward and reverse Illumina primers for the Sydney Harbour

estuary samples or 10 pmol each of forward and reverse 454 primers for the EAC and PH samples, 0.8 mM dNTPs, 2.5 mM $MgCl_2$, 1X reaction buffer, 1 unit Taq polymerase (Qiagen, Australia) and 5 – 10 ng of template DNA. The PCR program comprised a denaturation step at 95°C for 5 min, followed by 30 cycles of denaturation at 95°C for 30s, annealing at 55°C for 30s, extension at 72°C for 45s. The final step was an extension for 5 min at 72°C. Amplification products were quantitated using Quant-iT PicoGreen dsDNA assay kit (Life Technologies, Australia). For the Sydney Harbour samples, 10 ng of each sample were pooled and purified using Agencourt AMPure XP bead purification (Beckman Coulter, Inc., Australia). Purified amplicons were eluted using nuclease-free water (Ambion, Australia). Purified multiplexed samples were sequenced using the Illumina MiSeq platform at Ramaciotti Centre for Genomics (NSW, Australia). For the EAC and PH samples, 10 ng for *petB* of each sample were pooled and cleaned on preparative 2.5% agarose gel and the extracted product purified using the Wizard gel cleanup kit from following the manufacturer's instructions (Promega, Australia). Purified amplicons were eluted using nuclease-free water (Ambion, Australia). The multiplexed samples were sequenced using the Roche 454 platform at Ramaciotti Centre for Genomics (NSW, Australia).

Bioinformatic analyses

The *petB* amplicon sequencing reads for both 454 and Illumina data were processed using the USEARCH64 (Edgar, 2010) OTU clustering pipeline with some modifications. The commands used are included in Supplementary file 4.1. Sequencing adaptors and primers were removed using QIIME (Caporaso et al, 2010) split_libraies.py, including sequence trimming for length (280 bp), quality (19) and demultiplexing the sequencing libraries. The trimmed sequences were collected into a single fasta file, dereplicated (using usearch –derep) and filtered to remove sequences with less than 4 representatives before OTU clustering.

For *de novo* OTU picking, dereplicated sequences were clustered at 97% identity, chimeras were removed at this stage, and the resultant OTUs were further clustered at 94%. A mapping file was

created containing *de novo petB* OTUs and *petB* closed reference sequences (using an up to date *petB* sequence database containing sequences from complete genomes, cloned bidirectional sequences from the Warwick and Roscoff culture collections, and bidirectional sequences from clone libraries of environmental amplicons (Mazard et al, 2012; Humily et al, 2014)).

The quality trimmed sequence reads were then searched against the reference database to create a mapping file. The mapped sequences were used to generate an OTU table using the python script uc2otutab.py accompanying usearch. Representative sequences of *de novo* OTUs and the closed reference sequences were used to generate a multiple sequence alignment and a consensus phylogenetic tree in ARB (Ludwig et al, 2004) using Neighbor Joining and PhyML (Guindon et al, 2010), which was then used to assign taxonomy and examine the phylogenetic placement of *de novo* OTUs. The OTU matrix with taxonomic assignment is provided in Supplementary table 4.2.

Microbial community structure data determined using universal 454 sequencing of 16S rDNA amplicons (Jeffries et al, 2016) was re-analysed using the same USEARCH pipeline as above with minor modifications. Barcodes and primers were removed and all sequences were trimmed to 360 bp after quality filtering (notes). Sequences with Ns and any less than 360 bp in length were discarded. *De novo* OTUs were produced at 99% identity from the pool of dereplicated sequences after removal of singletons. Taxonomy was assigned against the Silva 119 release (Yilmaz et al, 2014) 99% non-redundant reference database. A mapping file and an OTU table (Supplementary table 4.3) were created and the relative abundance of sequences classified as Cyanobacteria; chloroplasts, and Cyanobacteria were extracted to estimate the relative proportion of phototrophic sequences at each site. Representative sequences classified as Cyanobacteria were assigned to *Prochlorococcus* and *Synechococcus* sub clusters 5.1, 5.2 and 5.3 using a 16S rRNA gene reference phylogenetic tree from Mazard et al (2012).

Statistical analyses

Ordination analyses (correspondence analysis, CA, and canonical correspondence analyses, CCA) and estimation of diversity indices were performed in R (R Core Team, 2015) using the Vegan package version 2.3-0 (Oksanen et al, 2015). The community composition in terms of *petB* sub-clade proportions were analysed with biotic and abiotic environmental parameter profiles at each site. The statistical significance of influence on the *Synechococcus* community structure, for each of the environmental parameters, was determined by ANOVA of the single factor CCA. The Simpson's reciprocal index was calculated for each sample and rarefaction curves produced rarefying samples for each of the datasets to the smallest sample in this dataset. All results were visualised using R software.

RESULTS AND DISCUSSION

Seasonal and spatial variation in environmental parameters

The sampling sites, ranging from the Parramatta River to the mouth of the harbour, are grouped into six regions namely Parramatta River, Lane Cove, Western Central Harbour, Eastern Central Harbour, Middle Harbour and Marine/Harbour Heads. Between the two seasons, temperature, salinity, dissolved oxygen and particulate organic nitrate are markedly different (Jeffries et al, 2016; Supplementary table 4.4). A comparison of the spatial heterogeneity of environmental parameters between the two months showed that there is more variability across sites in February whilst in September the conditions are more homogeneous. In February, the temperature ranges from 22.15°C – 27.59°C with highest temperatures at sites in the inland branches of the estuary. Temperature was at a lower range, from 17.16°C – 20.13°C, in September. Thus, there is a distinct difference in average temperatures between seasons with more variations between sites in February. As in the case of temperature, in February there is a salinity gradient along the estuary with highest salinity (~34 ppt) at the sites closest to the harbour and lowest furthest inland (<13 ppt). The salinity is more uniform along the estuary in September (~34 ppt). Nutrient (various forms of nitrogen and phosphorous such as oxidised, total and total dissolved) concentrations also showed higher heterogeneity between sites in February than

September. The variability in salinity and nutrient concentrations is attributed to high rainfall events in February. Such periodic inputs of freshwater from point sources result in temporary and localised increase in nutrient concentrations (Beck and Birch, 2012). Overall, there is less variation in temperature, salinity and nutrient concentrations along the estuary in September which suggests greater mixing and/or may be due to less influx from point sources.

Composition of picocyanobacteria in the estuary

Based on the 16S rRNA gene composition of bacterial communities in Sydney Harbour (Jeffries et al, 2016), cyanobacteria contributed up to 41.3% of the bacterioplankton community with higher abundance in sites at the mouth of the estuary in February whilst in September the proportion of cyanobacteria dropped dramatically to below 1%. Thus, in February cyanobacteria appear to be a prominent component of the bacterioplankton community. In comparison to the relative abundance of sequences assigned to chloroplast (up to 46.7%), cyanobacteria may also contribute significantly to primary production in the estuary. In February, the cyanobacterial population throughout Sydney Harbour mainly comprised marine *Synechococcus* sub-cluster 5.1 (97.03%) with only a very minor portion of sub-cluster 5.2 (0.52%), *Prochlorococcus* (1.89%) and other cyanobacteria (0.56%). In the comparatively well-studied temperate estuarine environment, the Chesapeake Bay (36-39˚N), the picocyanobacterial community is dominated by halotolerant *Synechococcus* lineages belonging to sub-cluster 5.2 in the upper bay area (further inland) whilst lineages of sub-cluster 5.1 were predominant in the regions closest to the mouth of the bay (Chen et al, 2006). The Chesapeake Bay has three main river systems draining into it with significant inflows of freshwater and salinities range from 5 ppt in the upper bay to 25 ppt at the marine end (Wang et al, 2011). Other brackish environments such as the Baltic Sea (53-66˚N) which display salinity gradients from freshwater to marine show a similar spatial partitioning of sub-clusters 5.1 and 5.2 (Larsson et al, 2014). Similar differences in the spatial partitioning of *Synechococcus* lineages were expected in the Sydney Harbour estuary with the predominance of coastal/oceanic lineages at the mouth of the estuary due to tidal flushing and more freshwater/euryhaline lineages further inland. However, *Synechococcus* sub-cluster 5.1 was the

predominant group throughout the estuary suggesting a greater oceanic influence in comparison to other estuarine environments wherein sub-cluster 5.2 is quite prevalent.

Seasonal partitioning of *Synechococcus* community in the estuary

The prokaryotic community composition, as determined using 16S rDNA by Jeffries et al (2016) showed significant shifts between February and September (Figure 4.2). The changes in abundance (averaged across all sites of the estuary) between the two months were distinct for different microbial groups. There is a higher abundance of chloroplast, cyanobacteria and SAR11 in February. In the early spring sample (September) there is a significant decrease in the relative abundance of sequences assigned to phototrophs. Average relative abundance of cyanobacteria reduces from 12.69% to 0.25% and chloroplasts from 10.2% to 0.41% across the estuary accompanied by a decrease in SAR11 (8% to 0.5%). The abundance of several heterotrophic bacterial groups, including SAR86, OM1 (marine Actinobacteria) and NS5 marine group (Flavobacteria), increases. To explore the cyanobacterial component further, the marine *Synechococcus* sub-cluster 5.1 community composition in the Sydney Harbour estuary was analysed through the high resolution *petB* phylogenetic marker (Mazard et al, 2012). In comparison to the *petB* marker, the level of taxonomic resolution provided by 16S rRNA marker is not sufficient to delineate lineages of *Synechococcus* (Mazard et al, 2012). The community structure was compared between February and September. Rarefaction curves of species diversity show a higher diversity of *Synechococcus* lineages in February than in September (Figure 4.3). Ordination analyses showed a clear distinction between the *Synechococcus* community composition from the two seasons with a statistically significant correlation to environmental parameters including temperature, salinity, dissolved oxygen, total nitrogen (TN) and particulate organic nitrogen (PON) (Figure 4.4; Table 4.1). Though the concentrations of other nutrient sources such as oxidised and total dissolved forms of nitrogen and phosphorous varied between the seasons, the influence of these parameters was not statistically significant. Parameters that are not mutually exclusive and other unmeasured factors such as light, metals and pollutants may influence the community.

The *Synechococcus* community profile based on *petB* amplicons showed a strong seasonal variation in community composition (Figure 4.5). In February, the predominant *Synechococcus* lineage was clade II (52.82 – 70.37%; with multiple distinct sub-lineages IIa, IIc and IIe), followed by clade Ib (13.91 to 41.23%) in all sites [except in P3 where clade II was 40.42% and clade Ib was 50.17%]. The *petB* OTUs that were phylogenetically related to clade VI were present in the estuary in February at up to 11.48% of *petB* sequences, however, the phylogenetic affiliation of these OTUs was not clearly resolved. Previous reports of clade VI in the environment have been presented as a combined detection of a group comprising clades V/VI/VII using 16S rDNA sequence-based methods of detection (Zwirglmaier et al, 2008; Ahlgren and Rocap, 2012). The use of the *petB* gene as the phylogenetic marker enables greater taxonomic resolution of these lineages and better detection of their individual distributions (Mazard et al, 2012). The detection of clade VI in the Sydney Harbour estuary and not in the coastal regions may suggest that this clade is more prevalent in estuarine environments (Figures 4.5 and 4.6).

In September, despite the low proportion of *Synechococcus* in the microbial community, there is a change in the *Synechococcus* composition (based on the *petB* marker; Figure 4.5) from the predominance of clade II to being mainly comprised of clade I (59.95 – 90.43%) and IV (6.79 – 17.62%). The presence of clade II sub-groups was marginal (1.65 – 8.83%) throughout the estuary, except at the sites near the mouth of the harbour, i.e. Middle Harbour and Marine Heads, where their proportion was higher (16.57 – 32.53%) than other sites. The other lineages detected in February were not present in September. Thus, there is an evident temporal variation in the *Synechococcus* community composition. Seasonal partitioning of *Synechococcus* lineages have been reported in other estuarine environments. In the Chesapeake Bay, winter populations mainly comprised isolates closely related to clade I and other sub-lineages closely related to the brackish/euryhaline genus *Cyanobium*. The summer populations were composed of sub-cluster 5.2 and clade CB1 of sub-cluster 5.1. Temperature was determined to be a significant parameter influencing cyanobacterial communities in these estuarine ecosystems (Cai et al, 2010; Murrell and Lores, 2004).

In the Sydney Harbour estuary, the seasonal variation in the *Synechococcus* community with higher proportions of clades I and IV in September (18.3°C) and II in February (24.9°C) correlates with the specific temperature niches previously defined by global oceanic surveys. On a global scale, relative abundances of clades I and IV are highest in temperate, mesotrophic regions. The preferred temperature profile of clades I and IV is at a lower range of 10°C to 20°C (Sohm et al, 2015; Tai and Palenik, 2009; Zwirglmaier et al, 2008). The global distribution of clade II with highest prevalence in warmer open ocean waters suggests a higher temperature profile (20°C to 28°C) than clades I and IV (Sohm et al, 2015; Zwirglmaier et al, 2008).

Partitioning of *Synechococcus* lineages on a global scale is influenced by nutrient availability. Clades I and IV are more abundant in coastal environments with higher nutrient content whilst clade II is more prevalent in open ocean waters of lower nutrient content (Sohm et al, 2015; Tai and Palenik, 2009; Zwirglmaier et al, 2008). Based on ordination analyses, PON and TN concentrations may partially explain the variation in *Synechococcus* community composition between the two seasons observed here. However, under higher conditions of PON and TN (February), clade II is predominant whilst the clades I and IV are more prevalent in September, which is different to the global trend. The concentrations of nutrients, namely, other sources of nitrogen (oxidised and reduced) and phosphate are different between the two seasons but based on ordination analyses, these parameters do not appear to significantly influence the variation in community.

It is likely that other parameters, which were not measured during this study, have an important influence. The well-mixed waters in September may have increased chemicals and metals, released from the sediments of the estuary. Concentrations of metals can affect the distribution of *Synechococcus* lineages with clades I and IV being more tolerant of higher metal availability (Mackey et al, 2012; Stuart et al, 2009). The prevalent conditions clearly have a severe impact on the abundance of the *Synechococcus* community (based on the 16S rDNA analysis). The higher relative proportions of clade I and IV in comparison to other *Synechococcus* lineages (based on *petB*) as well as the overall lower abundance of cyanobacteria may be influenced by changes in concentrations of chemicals

and/or nutrients. The conditions in September thus significantly influence both the structure of the *Synechococcus* community as well as the general microbial community (both phototrophic and heterotrophic).

Spatial partitioning of *Synechococcus* lineages along the estuary

Jeffries et al (2016) reported spatial shifts in the community composition of bacterioplankton (based on 16S rDNA marker) in the Sydney Harbour estuary. Spatial variability in the microbial community was more evident during periods of heavy rainfall (February) with greater shifts in inland sites. This variability is attributed to changes in environmental parameters such as nutrient concentrations and dissolved oxygen (Jeffries et al, 2016). The abundance of cyanobacteria varied from 0.4% to 28.3% in February and was less than 0.74% in September across the different sites in the estuary. In February, the sites (P1, P2, LC1 and MH1) furthest away from the mouth of the harbour had the lowest abundance of cyanobacteria with less than 1.5% whilst the remaining sites had more than 8%. The changes in nutrient concentrations and the influx of heavy metals (Zn, Pb, Cu) and pesticides associated with higher rainfall may influence these differences (Birch et al, 2015) and warrant further investigation. However since metal and pesticide concentrations were not measured, it is difficult to determine their influence in the two months.

Though the abundance of the cyanobacterial population varied, the composition of marine cyanobacterial community remained relatively stable throughout the estuary within both months of sampling. Therefore the high rainfall inputs resulting in variability of environmental parameters including salinity and nutrients do not appear to significantly alter the *Synechococcus* community composition across the spatial scale. Some level of spatial partitioning was observed. The highest diversity in *Synechococcus* lineages was observed in February in the main branch of the estuary including Eastern/Western Central Harbour and the Marine Heads regions (Figure 4.3). Patterns in spatial partitioning are evident in the sub-lineages of clade II (Figure 4.5). Sub-lineage IIf occurred at higher abundance at sites further away from the harbour whilst sub-lineage IIe showed an opposite

trend with abundance increasing towards the mouth of the harbour. The measured environmental parameters did not show a strong correlation with observed trends for clade II sub-lineages suggesting that other factors are likely involved. Ordination plots showed a distinct clustering of inland sites (including Paramatta River, Lane Cove River and eastern/western central harbour) and those of Middle Harbour and Marine Heads in September (Figure 4.4). This can be attributed to the marginal percentage of clade II in most sites and a higher proportion in the Middle Harbour and Marine Heads regions. This suggests greater coastal influence in those sites. This community structure was similar to that found at the coastal Sydney site (Harbord) outside of the harbour (Figures 4.5 and 4.6).

***Synechococcus* community along the east Australian coast**

To gain a better understanding of the factors affecting microbial communities in the Sydney Harbour estuary we compared cyanobacterial genotypes in marine waters off the east coast. The East Australian Current and its subsequent eddy field, which originates in the region south of Smoky Cape (30˚55'S), have a strong influence on coastal environments. The EAC transports warm, nutrient-depleted waters southward from the tropical Coral Sea in the South Pacific (Schaeffer et al, 2013; Seymour et al, 2012). The EAC and its surrounds were sampled during a voyage of the RV *Southern Surveyor* in October 2010. The cyanobacterial population within the main EAC current mostly comprised clade II (sub-lineages IIa and IIc; Figure 4.6) which is a known inhabitant of warmer tropical/subtropical regions. Several clades appearing at lower relative abundance, including WPC1 and CRD1a, have been observed in high relative proportions of *Synechococcus* communities in the South China Sea and the Costa Rica dome (Huang et al, 2012; Saito et al, 2005).

The *Synechococcus* community composition changes at sites closer to the coast (EAC11, EAC1) with an increase in clades IIh and Ia. At a coastal site further south, close to Sydney (Harbord, water depth 7m), clade I becomes predominant along with a small proportion of clade IV, which is indicative of temperate coastal ocean waters (Tai and Palenik, 2009; Mazard et al, 2012; Zwirglmaier et al, 2008). There is also a shift in the dominance of clades IIa and IIc in the EAC to increased prevalence of

clades IIh and Ia nearer the coast. This shift may be influenced by changes in temperature and/or nutrients. The upwelling of cold nutrient rich waters caused by the EAC interaction with the shelf particularly in summer is known to influence coastal ecosystems either as a result of Ekman pumping (Roughan and Middleton, 2004; Hedge et al, 2013; Schaeffer et al, 2013) but also due to seasonal wind-driven upwelling. Such upwelling could result in an influx of nutrients into the Sydney Harbour estuary and also a change in the tidal inflow of cyanobacterial seeding populations.

A national reference station (PH4), 20 miles further south of Sydney Harbour and 3.2 miles offshore (water column depth 100m), is located on a continental shelf region south of the EAC separation zone. The *Synechococcus* community at PH4 is predominantly composed of clade II lineages in both February and September. Clade IIa is the most abundant lineage, similar to the EAC. However, the presence of *petB* sub-clades IIh, IV and XVI indicate that communities at the coastal and shelf sites are quite distinct from those of the EAC. These observations may indicate some influence of the tropical current on local populations, but also suggest a succession of genotypes (i.e IIh, IIe and IIf replacing IIc and WPC1) as the EAC waters are transported south. It is interesting to note that the water column at PH4 is stratified for most of the year (Australian ocean data network portal; portal.aodn.org.au) and that *Synechococcus* genotypes partition down the water column. In particular, the temperate clades Ia, Ib and IV increase in abundance below the thermocline (i.e. below 25 m). The mixed layer is deeper down the water column in September (35m) than in February (12m). In September, clade I increased in abundance, particularly deeper (100m) in the water column. Ekman pumping and tidal mixing in close proximity to the coast may be mixing/selecting for temperate/mesotrophic genotypes that are being washed into the harbour.

CONCLUSIONS

This is the first study to describe the genetic composition of cyanobacteria in the Sydney Harbour estuary. The cyanobacterial community in the estuary appears to differ from the other reported estuarine environments in that the dominant lineages are of *Synechococcus* sub-cluster 5.1 rather than

the presence of both sub-clusters 5.1 and 5.2. This suggests that pelagic cyanobacteria in the estuary are strongly influenced by the adjacent coastal/oceanic waters, rather than from freshwater inputs. In order to confirm the dominance of sub-cluster 5.1 and to discount the possibility of biases (such as efficiency of DNA extraction and/or primer amplification) against other groups, techniques such as analytical flow cytometry and fluorescence *in-situ* hybridisation (FISH) with specific probes would help establish the seasonal cellular abundances of cyanobacteria in the estuary.

The study provides an account of the spatial and temporal partitioning of *Synechococcus* genotypes and the environmental parameters that influence these lineages within the estuary. There is a strong temporal partitioning with significant changes in the community composition between the two seasons with temperature, salinity and some nitrogen sources strongly influencing the observed differences. Along the length of the estuary, spatial partitioning of marine *Synechococcus* lineages is limited. One of the striking features in this data set is the reduction in the proportion of photosynthetic sequences in the universal 16S rDNA data in September. The drastic shift in the community structure from phototrophic to heterotrophic between the two seasons may be influenced by nutrient concentrations. The unexplained variation in abundance and community composition along the harbour is certainly influenced by important factors that were not assessed within the scope of this study, such as metal availability as well as biotic factors such as cyanophages and grazers. Further study of competition amongst microbial populations could also provide valuable insights. Additional work would also include isolation and genome sequencing to further investigate physiological characteristics and responses to anthropogenic stressors.

ACKNOWLEDGEMENTS

This work was supported by the Australian Research Council Discovery Grant DP110102718 to ITP and an International Postgraduate Research Scholarship to DV from the Australian government. We thank Justin Seymour and Mark Brown for the EAC sample collection and DNA; Justin Harrison and Thomas Jeffries for sample collection and nutrient analyses for the Sydney Harbour estuary samples.

The work on the Sydney Harbour estuary was supported by the Transfield Foundation Early Career Researcher Grant awarded to Thomas Jeffries, and by the Australian Research Council Discovery Grant Scheme (DP110103091 and DP120102764) to Justin Seymour. Sample collection for the Sydney Harbour estuary and physiochemical analyses were funded by Greater Sydney Local Land Services as part of the 'Sydney Harbour Water Quality Improvement Plan' and conducted at the Sydney Institute of Marine Science and Southern Cross University.

Supplementary information is provided in CD.

REFERENCES

Ahlgren N, Rocap G. (2012). Diversity and distribution of marine *Synechococcus*: multiple gene phylogenies for consensus classification and development of qPCR assays for sensitive measurement of clades in the ocean. Front Microbiol 3: 1-24.

Beck H, Birch G. (2012). Metals, nutrients and total suspended solids discharged during different flow conditions in highly urbanised catchments. Environ Monit Assess 184: 637-653.

Birch G, Drage D, Thompson K, Eaglesham G, Mueller J. (2015). Emerging contaminants (pharmaceuticals, personal care products, a food additive and pesticides) in waters of Sydney estuary, Australia. Mar Pollut Bull 97: 56-66.

Birch G, Lean J, Gunns T. (2015). Historic change in catchment land use and metal loading to Sydney estuary, Australia (1788–2010). Environmental Monitoring and Assessment 187:594.

Cai H, Wang K, Huang S, Jiao N, Chen F. (2010). Distinct patterns of picocyanobacterial communities in winter and summer in the Chesapeake Bay. Appl Environ Microbiol 76: 2955–2960.

Caporaso J, Kuczynski J, Stombaugh J, Bittinger K, Bushman F, Costello E, *et al*. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7: 335-336.

Caroppo C. (2015). Ecology and biodiversity of picoplanktonic cyanobacteria in coastal and brackish environments. Biodivers Conserv 24: 949-971.

Chen F, Wang K, Kan J, Suzuki M, Wommack E. (2006). Diverse and unique picocyanobacteria in Chesapeake Bay, revealed by 16S-23S rRNA Internal Transcribed Spacer sequences. Appl Environ Microbiol 72: 2239-2243.

Das P, Marchesiello P, Middleton J. (2000). Numerical modelling of tide-induced residual circulation in Sydney Harbour. Mar Freshwater Res 51: 97-112.

Edgar R. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26: 2460-2461.

Guindon S, Dufayard J, Lefort V, Anisimova M, Hordijk W, Gascuel O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59: 307-321.

Hatje V, Apte S, Hales L, Birch G. (2003). Dissolved trace metal distributions in Port Jackson estuary (Sydney Harbour), Australia. Mar Pollut Bull 46: 719-730.

Hedge L, Johnston E, Birch G, Booth D, Creese B, Doblin M, *et al*. (2013). Living harbour. Dynamic Science: A Systematic Review of the Science of Sydney Harbour The Sydney Institute of Marine Science, Sydney, New South Wales, Australia (ISBN: 978-0-646-91693-0, Publisher: The Sydney Institute of Marine Science, Sydney, New South Wales, Australia).

Hose G, Gordon G, McCullough F, Pulver N, Murray B. (2005). Spatial and rainfall related patterns of bacterial contamination in Sydney Harbour estuary. J Water Health 3: 349-358.

Huang S, Wilhelm S, Harvey H, Taylor K, Jiao N, Chen F. (2012). Novel lineages of *Prochlorococcus* and *Synechococcus* in the global oceans. ISME J 6: 285-297.

Humily F, Farrant G, Marie D, Partensky F, Mazard S, Perennou M, *et al*. (2014). Development of a targeted metagenomic approach to study a genomic region involved in light harvesting in marine *Synechococcus*. FEMS Microbiol Ecol 88: 231-249.

Jardillier L, Zubkov M, Pearman J, Scanlan D. (2010). Significant CO_2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. ISME J 4:1180–1192.

Jeffries T, Schmitz-Fontes M, Harrison D, van-Dongen-Vogels V, Eyre B, Ralph P, *et al*. (2016). Bacterioplankton dynamics within a large anthropogenically impacted urban estuary. Front Microbiol 6: A 1438.

Larsson J, Celepli N, Ininbergs K, Dupont C, Yooseph S, Bergman B *et al*. (2014). Picocyanobacteria containing a novel pigment gene cluster dominate the brackish water Baltic Sea. ISME J 8: 1892-1903.

Lee S, Birch G, Lemckert C. (2011). Field and modelling investigations of fresh-water plume behaviour in response to infrequent high-precipitation events, Sydney Estuary, Australia. Estuar Coast Shelf S 92: 389-402.

Li W. (1994). Primary productivity of prochlorophytes, cyanobacteria, and eucaryotic ultraphytoplankton: measurements from flow cytometric sorting. Limnol Oceanogr 39:169–175.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, *et al*. (2004). ARB: a software environment for sequence data. Nucl Acids Res 32: 1363-1371.

Mackey K, Buck K, Casey J, Cid A, Lomas M, Sohrin Y, *et al*. (2012). Phytoplankton responses to atmospheric metal deposition in the coastal and open-ocean Sargasso Sea. Front Microbiol 3: 359.

Mazard S, Ostrowski M, Partensky F, Scanlan D. (2012). Multi-locus sequence analysis, taxonomic resolution and biogeography of marine *Synechococcus*. Environ Microbiol 14: 372-386.

McCready S, Birch G, Taylor S. (2003). Extraction of heavy metals in Sydney Harbour sediments using 1M HCl and 0.05M EDTA and implications for sediment-quality guidelines. Aust J Earth Sci 50: 249-255.

McCready S, Slee D, Birch G, Taylor S. (2000). The distribution of polycyclic aromatic hydrocarbons in surficial sediments of Sydney Harbour, Australia. Mar Pollut Bull 40: 999-1006.

Murrell M, Lores E. (2004). Phytoplankton and zooplankton seasonal dynamics in a subtropical estuary: importance of cyanobacteria. J Plankton Res 26: 371-382.

Ning X, Cloern J, Cole B. (2000). Spatial and temporal variability of picocyanobacteria *Synechococcus* sp. in San Francisco Bay. Limnol Oceanogr 45: 695-702.

Oksanen J, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hare R, *et al*. (2015). vegan: Community Ecology Package. R package version 2.3-0. CRAN.R-project.org/

Ostrowski M, Mazard S, Pitt F, Holland R, Tarran G, Puxty R, *et al*. (*submitted*). Biogeographical provinces of the Atlantic Basin inferred from genetic diversity of a prokaryotic phototroph. Proc Natl Acad Sci USA.

Phlips E, Badylak S, Lynch T. (1999). Blooms of the picoplanktonic cyanobacterium *Synechococcus* in Florida Bay, a subtropical inner-shelf lagoon. Limnol Oceanogr 44: 1166-1175.

R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Saito M, Rocap G, Moffett J. (2005). Production of cobalt binding ligands in a *Synechococcus* feature at the Costa Rica upwelling dome. Limnol Oceanogr 50:279–290.

Scanlan D, Ostrowski M, Mazard S, Dufresne A, Garczarek L, Hess W, *et al*. (2009). Ecological genomics of marine picocyanobacteria. Microbiol Mol Biol Rev 73: 249-299.

Schaeffer A, Roughan M, Morris B. (2013). Cross-shelf dynamics in a Western Boundary Current regime: Implications for upwelling. J Phys Oceanogr 43: 1042-1059.

Seymour J, Doblin M, Jeffries T, Brown M, Newton K, Ralph P, *et al*. (2012). Contrasting microbial assemblages in adjacent water masses associated with the East Australian Current. Environ Microbiol Rep 4: 548-555.

Sohm J, Ahlgren N, Thomson Z, Williams C, Moffett J, Saito M *et al*. (2015). Co-occurring *Synechococcus* ecotypes occupy four major oceanic regimes defined by temperature, macronutrients and iron. ISME J 1-13.

Stuart R, Dupont C, Johnson A, Paulsen I, Palenik B. (2009). Coastal strains of marine *Synechococcus* species exhibit increased tolerance to copper shock and a distinctive transcriptional response relative to those of open-ocean strains. Appl Environ Microbiol 75(15): 5047-5057.

Tai V, Palenik B. (2009). Temporal variation of *Synechococcus* clades at a coastal Pacific Ocean monitoring site. ISME J 3: 903-915.

Wang K, Wommack E, Chen F. (2011). Abundance and distribution of *Synechococcus* spp. and cyanophages in the Chesapeake Bay. Appl Environ Microbiol 77: 7459-7468.

Yilmaz P, Parfrey L, Yarza P, Gerken J, Pruesse E, Quast C, *et al*. (2014). The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. Nucl Acids Res 42: D643-D648.

Zwirglmaier K, Jardillier L, Ostrowski M, Mazard S, Garczarek L, Vaulot D, *et al*. (2008). Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. Environ Microbiol 10: 147-161.

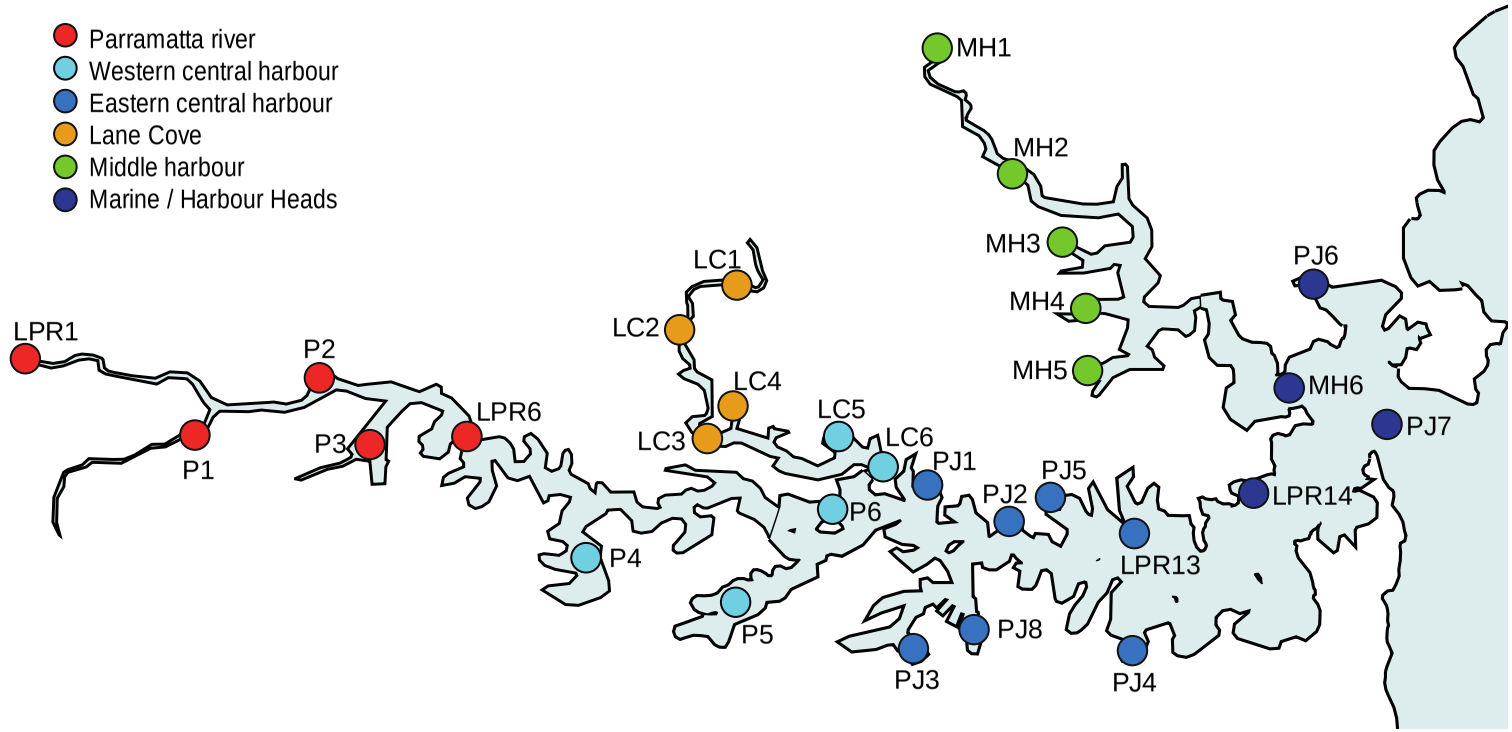


Figure 4.1: Map of the Sydney Harbour estuary depicting the sampling sites, color-coded by region: Parramatta River (red), Western Central Harbour (light blue), Eastern Central Harbour (blue), Lane Cove (orange), Middle Harbour (green) and Marine/Harbour Heads (dark blue).

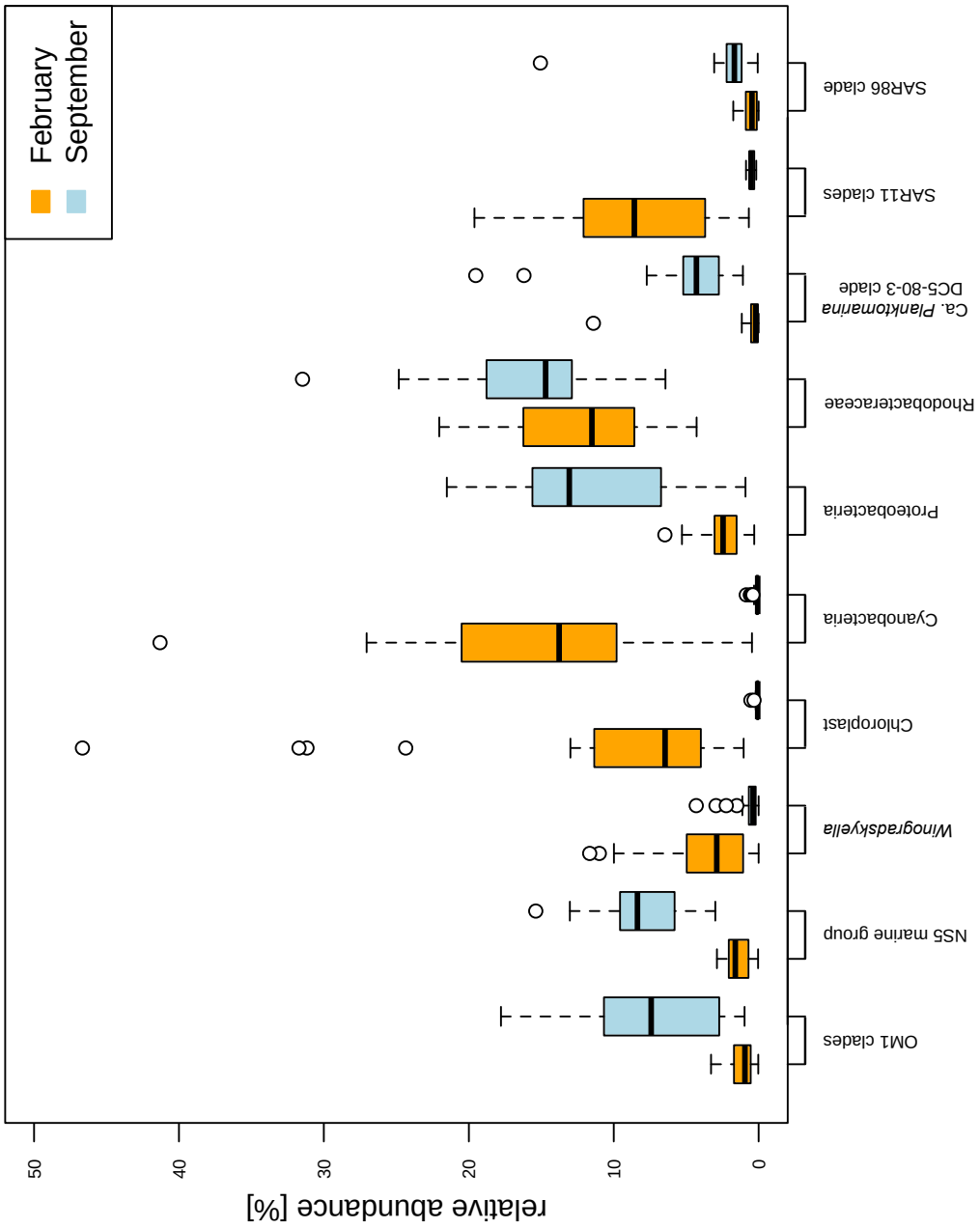


Figure 4.2: Box plot representation of the relative abundance (based on 16S rDNA amplicon sequencing, Jeffries et al, 2016) of selected microbial groups in the Sydney Harbour estuary in February (24.85°C) and September (18.29°C).

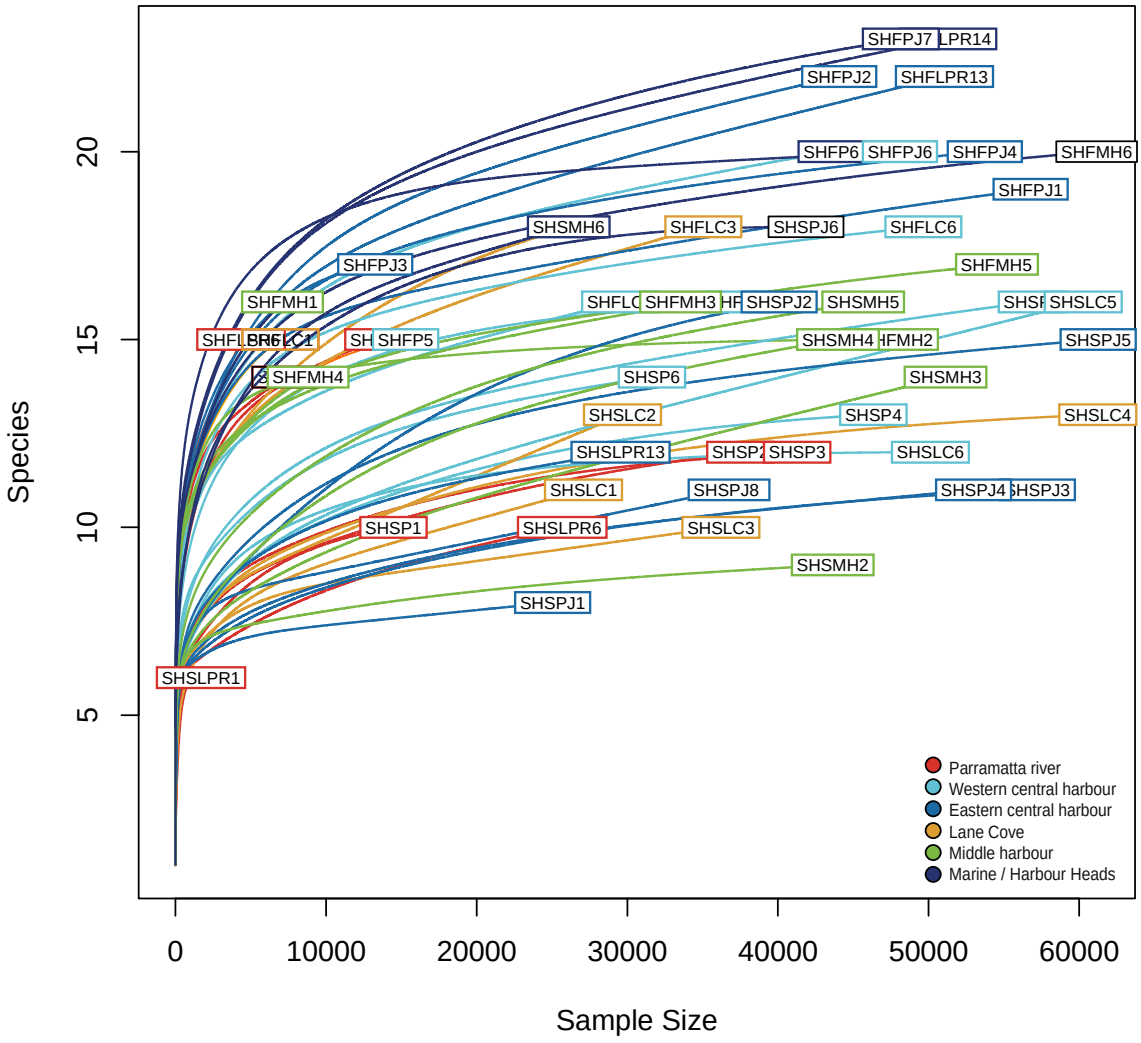
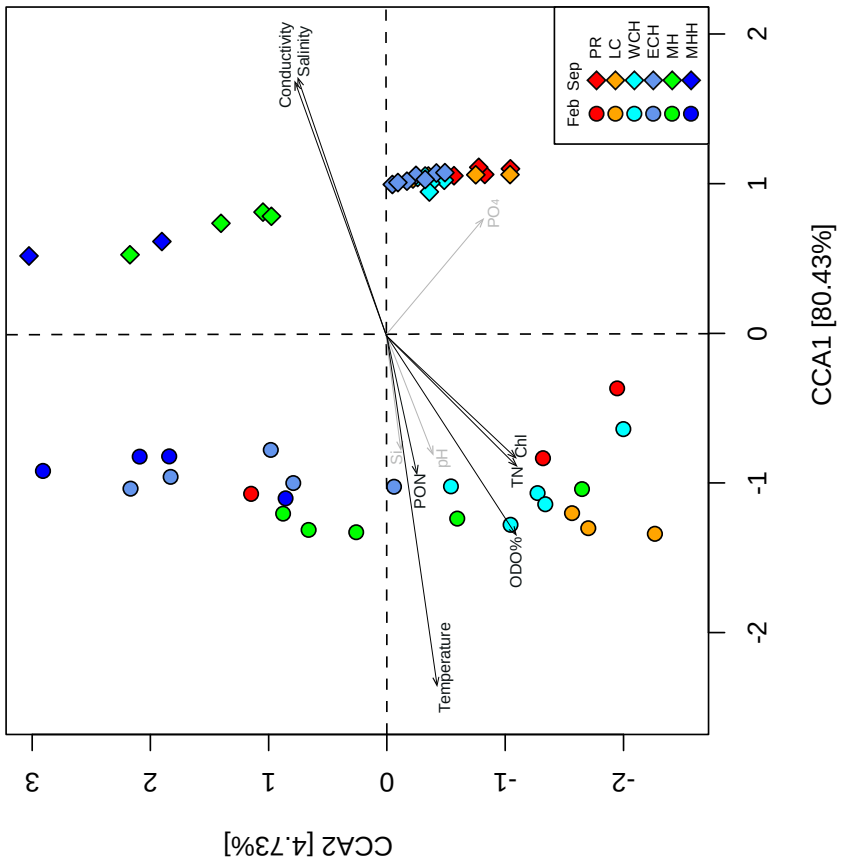


Figure 4.3: Rarefaction curves representing the *Synechococcus* sub-cluster 5.1 diversity (based on *petB* amplicon sequencing) in the Sydney Harbour estuary. Samples were color-coded by regions: Parramatta River (red), Western Central Harbour (light blue), Eastern Central Harbour (blue), Lane Cove (orange), Middle Harbour (green) and Marine/Harbour Heads (dark blue).


CCA1 [80.43%]
CCA2 [4.73%]
Feb Sep
PR
LC
WCH
ECH
MH
MHH
Conductivity
Salinity
PO4
Si
pH
PON
TN
Chl
ODO%
Temperature

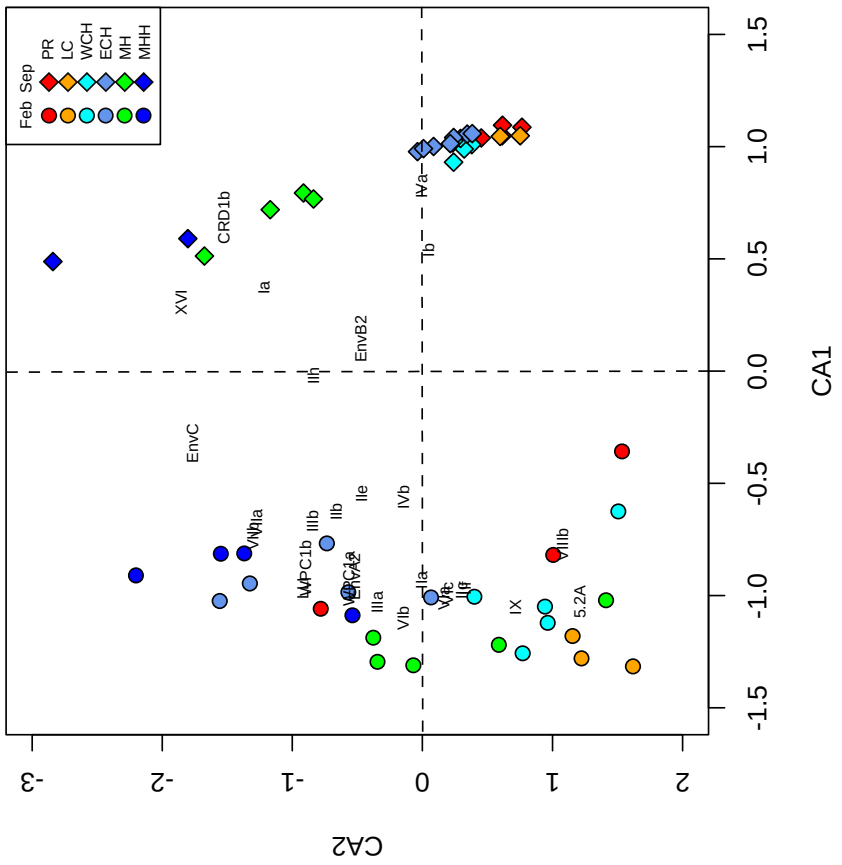
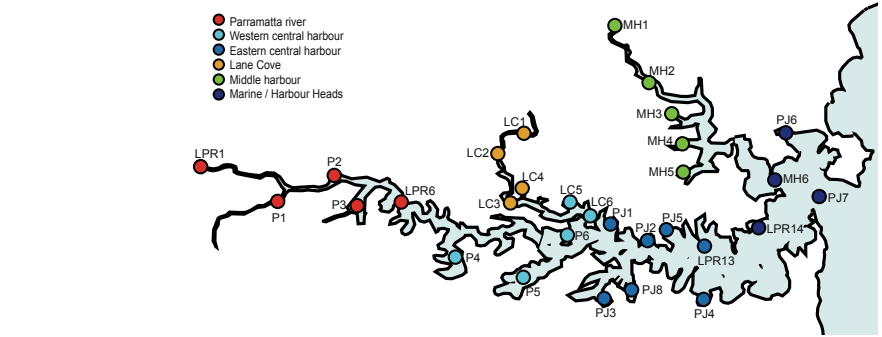
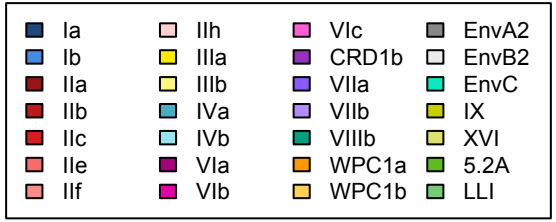
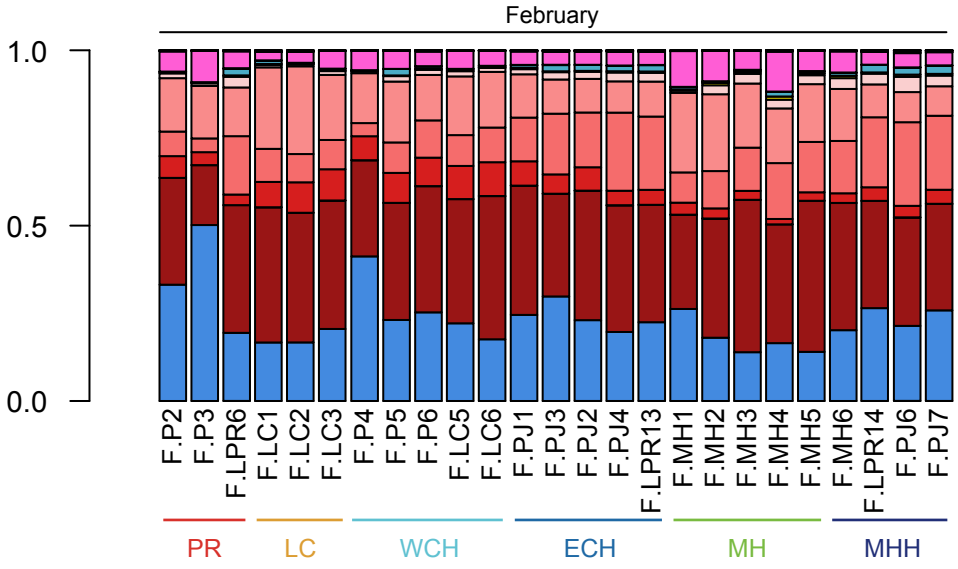

CA1
CA2
Feb Sep
PR
LC
WCH
ECH
MH
MHH
XVI
CRD1b
Ia
EnvB2
IIh
EnvC
IVa
Tb
IVb
IIe
IIb
IIIb
IVa
IIa
IIIa
VIb
IIa
IIf
IX
5.2A
VIIb

Figure 4.4: 2D spatial representation of correspondence analysis of the sites in the Sydney Harbour estuary according to the marine *Synechococcus* community composition (based on *petB* amplicon sequencing), a) excluding environmental parameters (CA) and b) including environmental parameters (CCA) with significant ones (p <0.001 and p <0.01 in black, p <0.05 in grey) indicated. Sites and marine *Synechococcus* lineages are represented on the two first dimensions CA1 and CA2. The percentage of the total inertia resolved from these dimensions is indicated as percentage next to each axis. Sites coloured according to the previously defined groups: Parramatta River (red), Western Central Harbour (light blue), Eastern Central Harbour (blue), Lane Cove (orange), Middle Harbour (green) and Marine/Harbour Heads (dark blue). Samples collected in February are indicated by dots, samples collected in September are indicated by diamonds.


Parramatta river
Western central harbour
Eastern central harbour
Lane Cove
Middle harbour
Marine / Harbour Heads
LPR1
P1
P2
P3
LPR6
P4
LC1
LC2
LC3
LC4
LC5
LC6
P6
P5
PJ3
PJ8
PJ1
PJ2
PJ5
PJ4
LPR13
LPR14
MH1
MH2
MH3
MH4
MH5
MH6
PJ6
PJ7


Ia
Ib
IIa
IIb
IIc
IIe
IIf
IIh
IIIa
IIIb
IVa
IVb
Vla
Vlb
Vlc
CRD1b
VIIa
VIIb
VIIIb
WPC1a
WPC1b
EnvA2
EnvB2
EnvC
IX
XVI
5.2A
LLI


February
1.0
0.5
0.0
F.P2
F.P3
F.LPR6
F.LC1
F.LC2
F.LC3
F.P4
F.P5
F.P6
F.LC5
F.LC6
F.PJ1
F.PJ3
F.PJ2
F.PJ4
F.LPR13
F.MH1
F.MH2
F.MH3
F.MH4
F.MH5
F.MH6
F.LPR14
F.PJ6
F.PJ7
PR
LC
WCH
ECH
MH
MHH

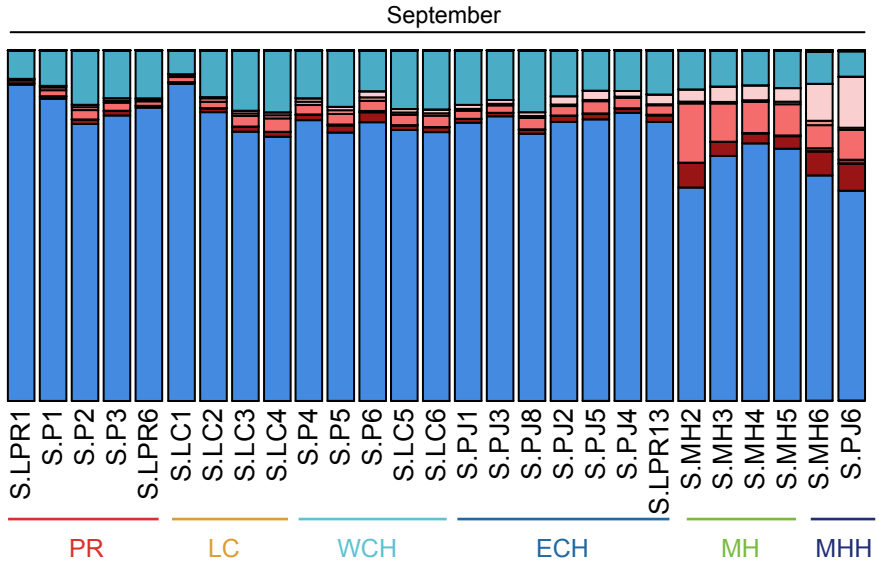

September
S.LPR1
S.P1
S.P2
S.P3
S.LPR6
S.LC1
S.LC2
S.LC3
S.LC4
S.P4
S.P5
S.P6
S.LC5
S.LC6
S.PJ1
S.PJ3
S.PJ8
S.PJ2
S.PJ5
S.PJ4
S.LPR13
S.MH2
S.MH3
S.MH4
S.MH5
S.MH6
S.PJ6
PR
LC
WCH
ECH
MH
MHH

Figure 4.5: Barchart representation of the marine *Synechococcus* component in the Sydney Harbour estuary in February and September. The community composition was derived from the *petB* gene (Mazard et al, 2012) through amplicon sequencing. Sites coloured according to the previously defined groups: Parramatta River (red), Western Central Harbour (light blue), Eastern Central Harbour (blue), Lane Cove (orange), Middle Harbour (green) and Marine/Harbour Heads (dark blue).

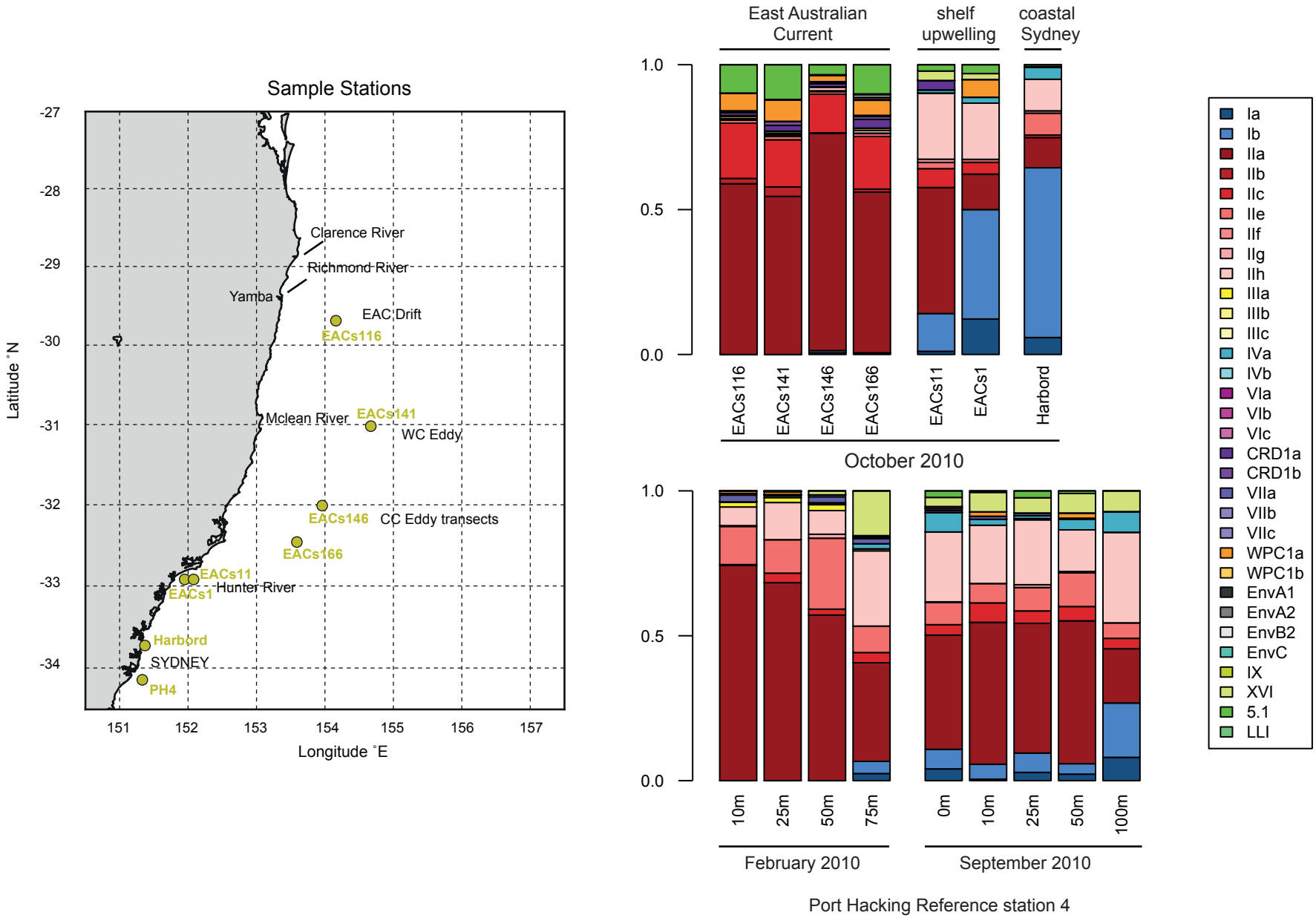

Sample Stations
Latitude ˚N
Longitude ˚E
-27
-28
-29
-30
-31
-32
-33
-34
151
152
153
154
155
156
157
Clarence River
Richmond River
Yamba
EAC Drift
EACs116
Mclean River
EACs141
WC Eddy
EACs146
CC Eddy transects
EACs166
EACs11
EACs1
Hunter River
Harbord
SYDNEY
PH4
East Australian Current
shelf upwelling
coastal Sydney
1.0
0.5
0.0
EACs116
EACs141
EACs146
EACs166
EACs11
EACs1
Harbord
October 2010
10m
25m
50m
75m
0m
10m
25m
50m
100m
February 2010
September 2010
Port Hacking Reference station 4
Ia
Ib
IIa
IIb
IIc
IIe
IIf
IIg
IIh
IIIa
IIIb
IIIc
IVa
IVb
Via
VIb
VIc
CRD1a
CRD1b
VIIa
VIIb
VIIc
WPC1a
WPC1b
EnvA1
EnvA2
EnvB2
EnvC
IX
XVI
5.1
LLI

Figure 4.6: Barchart representation of the marine *Synechococcus* component in oceanic regions surrounding Sydney Harbour: East Australian Current (EAC) and the Port Hacking (PH4) national reference station. The community composition was derived from the *petB* gene (Mazard et al, 2012) through amplicon sequencing. Sampling points are indicated on the insert reference map. Legend colour scheme is the same as in Figure 4.5.

Table 4.1: Environmental parameters and their influence on the *Synechococcus* community composition (based on *petB* gene) in the Sydney Harbour estuary

	Percentage of inertia	*p* value
Longitude	5.4	ns
Temperature	71.6	***
Conductivity	37.7	***
Salinity	38.1	***
pH	9.0	*
Turbidity	1.2	ns
Chlorophyll (post calibration)	10.1	**
Dissolved oxygen (ODO%)	24.9	***
Dissolved oxygen (ODO)	2.7	ns
Secci. Depth	6.5	ns
mono-nitrogen oxides (NO_x)	0.4	ns
phosphate (PO_4)	8.3	*
Ammonium (NH_4)	1.2	ns
nitrite (NO_2)	0.9	ns
Total dissolved phosphorus (TDP)	1.3	ns
Total dissolved nitrogen (TDN)	5.9	ns
Total phosphorus (TP)	1.4	ns
Total nitrogen (TN)	11.2	**
Particulate organic phosphorus (POP)	1.6	ns
Particulate organic nitrogen (PON)	11.7	***
Silicate	8.8	*
Total suspended solids (TSS)	6.4	ns

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns - not significant

Chapter 5:

Microbial primary producers of Oceania: molecular characterisation of prokaryotic and eukaryotic communities across a range of habitats

5.1 General introduction

Marine cyanobacteria occupy a wide range of environmental habitats. The various defined phylogenetic clades from the two main genera, *Prochlorococcus* and *Synechococcus*, were shown to have differential environmental distribution (Scanlan, 2012; Sohm et al, 2015; Zwirglmaier et al, 2008). One of the broadest parameters underlying this distribution was temperature. In order to complement the work with isolates on cellular acclimation to temperature shifts, we investigated the biogeography of marine cyanobacteria throughout different thermal niches in water bodies around Australia.

Marine cyanobacteria represent the most significant prokaryotic fraction of primary producers. However, the eukaryotic photosynthetic microorganisms have been shown to have equally significant impact on primary production in the marine environment (Li, 1994; Jardillier et al, 2010). Together, cyanobacteria and eukaryotic phytoplankton contribute up to 50% of the net primary production. In addition to carbon fixation, marine phytoplankton play key roles in global biogeochemical cycling (Flombaum et al, 2013). Therefore, in parallel to the studies conducted on marine cyanobacterial populations, we investigated the diversity and distribution of the marine photosynthetic eukaryotes at the same sites.

Australia's surrounding marine systems range from ~10˚S (tropical) to 45˚S (temperate), extending into the Antarctic and encompassing a broad amplitude of water bodies characterised by distinct sea surface temperature profiles. Sites were selected along a latitudinal gradient along the East coast of Australia to study marine *Synechococcus* and *Prochlorococcus* populations present within all of these distinct water-bodies.

Photosynthetic populations were also studied within the Sydney Harbour and supported by open ocean samples north and south of Sydney (Chapter 4 and section 5.3) as well as in Antarctic waters within two transects through the Southern Ocean to the sea ice edge (section 5.2). A study has been initiated and samples collected in waters off Maria Island (Tasmania) (section 5.4). In addition to these sites,

other regions are being analysed to complement the thermal range, and oceanographic features along the East coast of Australia (e.g. Darwin to Cairns transect, waters off Coffs Harbour) but do not form part of this body of work (Ostrowski et al, unpublished).

The study of the microbial communities in the Southern Ocean also explored the overall microbial plankton population including cyanobacteria, other prokaryotes and eukaryotes to gain insights into the community structure and composition as well as the potential interrelationships across the taxonomic groups.

The main questions underpinning this set of work were:

1- What are the relative proportions of photosynthetic microorganisms within the different thermal niches?
2- What is the distribution of the various groups, genera and clades and genotypes within the different thermal niches?
3- Are these differential environmental distributions correlated to specific environmental parameters?

The study of cyanobacterial community structure conducted on the Sydney Harbour / Port Hacking site was reported in Chapter 4. The following sections present the ongoing work at the other regions with varying temperature regimes.

5.2 Southern Ocean

5.2.1 *Introduction*

The Southern Ocean plays a major part in global primary production, nutrient transport and cycling as well as supporting the marine ecosystem. Though temperatures in the Southern Ocean remain below 10°C throughout the year, primary production and biomass is high particularly in summer (Wilkins et al, 2013). Shifts in the planktonic microbial communities along latitudinal transects across the Polar

Front have been previously reported using metagenomics (Wilkins et al, 2013). Distinct partitioning of community composition has been observed with cyanobacteria and eukaryotic phytoplankton as the dominant primary producers in the northern and southern waters of the Polar Front, respectively, based on metagenomics, pigment analysis and 16S rDNA clone libraries (Wilkins et al, 2013; Wilmotte et al, 2002).

Marine photosynthetic microorganisms are responsible for a large proportion of marine CO_2 fixation that accounts for over half of the Earth's global primary production (Field et al, 1998). The three main groups of photosynthetic microorganisms consists of the photosynthetic picoeukaryotes (PPEs), larger eukaryotes (such as diatoms) and the two sister genera of prokaryotes, *Prochlorococcus* and *Synechococcus.* The *Prochlorococcus* genus is mainly limited to the warmer oligotrophic oceanic regions within the 40˚N to 40˚S latitudes. This genus is not expected to play a significant role in the primary production of high latitude regions (Partensky et al, 1999), however, there are very few direct observations that chart the southerly extent of *Prochlorococcus* populations in the Southern hemisphere. The *Synechococcus* genus is a very diverse group that can be found throughout most aquatic environments from freshwater to hypersaline regions, with varying temperature and trophic regimes. Study of their phylogeny and genomes has identified an extremely high genetic variation (Dufresne et al, 2008; Scanlan et al, 2009). The marine *Synechococcus* is divided into sub-clusters with sub-cluster 5.1 being the predominant group found in oceanic regions. Microbial eukaryotes may be photosynthetic or heterotrophic (Caron et al, 2012). Photosynthetic eukaryotes are often referred to as algae or phytoplankton with PPE being the smallest size class (<2-3 μm). The eukaryotic component is considerably more diverse than prokaryotic phototrophs and composed of multiple taxa derived from a range of primary and secondary endosymbiosis, with virtually every class of algae being represented (Caron et al, 2012).

Due to the importance of microbes as the basis of the ecosystem and their involvement in global geochemical cycles, it has become crucial to understand the structure and dynamics of microbial communities, arguably even more so within the fragile polar regions, so as to monitor ecosystem

health and predict climate change associated impacts. Indeed, the IPCC 5th assessment report highlights that the various ecosystems in these regions have been shown to present accelerated alterations compared to the rest of the world. Arctic surface air temperatures have been warming at nearly twice the global rate for several decades, and there are already visible changes within the polar ecosystems (flora and fauna distribution) alongside a dramatic reduction of sea ice. A larger body of work has been conducted in Arctic regions due to easier access. In contrast very little is known at the molecular taxonomic level about the microbial communities, picocyanobacterial genotypes and photosynthetic eukaryotes inhabiting high latitude Southern oceanic waters. The objectives of this study were to explore the microbial community composition and structure in the Southern Ocean, more particularly in waters sampled during two latitudinal transects in early 2014 in areas with sea surface temperatures ranging from 10°C to -2°C, coupled to extensive sampling around the retreating sea ice edge around the Totten and Mertz Glaciers.

5.2.2 *Materials and methods*

5.2.2.1 Sample collection

The transit voyage (NBP 14-02) was conducted between 31 January and 10 March 2014 (Austral summer), representing two transects with a latitudinal range from 51˚S to 67˚S (see details in Figure 5.1 and Supplementary table 5.1). The inward transect directed into Mertz Glacier, the outward transect shipped out from Totten Glacier. Samples were collected by Dr Leanne Armand (Macquarie University) on-board the research vessel RV *Nathaniel B. Palmer,* as part of the U.S. Polar Expedition to East Antarctica. At each sampled station, 2 litres of seawater from the ship's underway seawater supply were passed through a Sterivex filter (0.22µm) without pre-fractionation using a peristaltic pump. The filters were then frozen at -80˚C until further experiments.

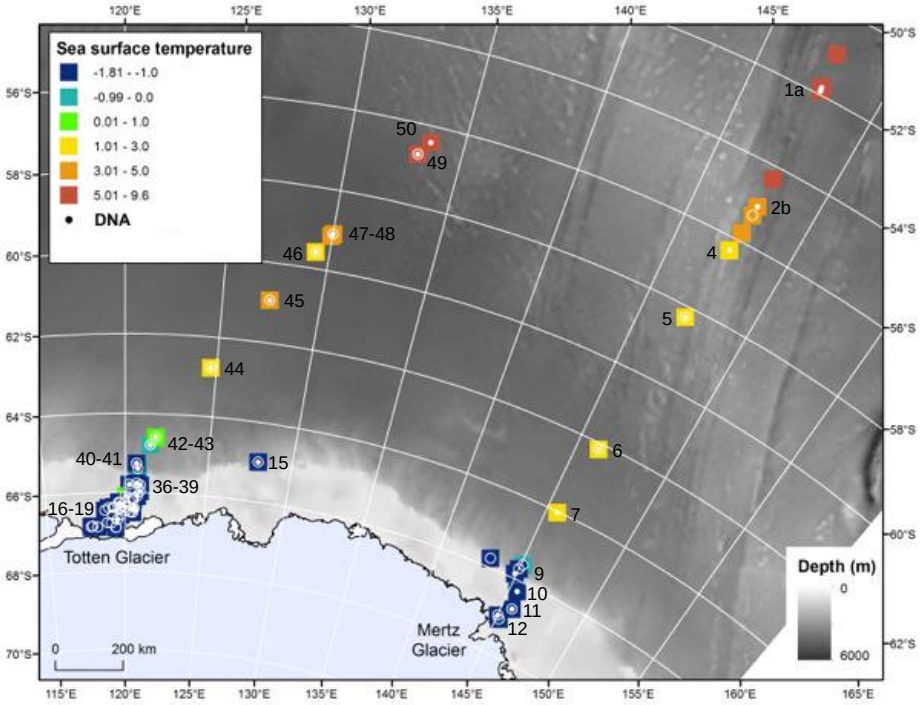


Figure 5.1: Map of the NBP 14-02 research voyage sampling track. The voyage initially travelled south-east towards the Mertz Glacier before crossing to Totten Glacier and travelling back south-east. The bathymetry is indicated as grey-scale. The sites sampled for DNA are indicated by a white dot and numbered. Sea surface temperature is indicated at each sampling point.

5.2.2.2 Nucleic acids extraction

DNA was extracted from all Sterivex filters using the PowerWater DNA isolation kit (MoBio, Australia). All sample processing was performed under quarantine regulation. The Australian Quarantine and Inspection Service (AQIS) required training, 'Quarantine Approved Premises for Accredited Persons Training', was undertaken prior to processing samples. Amplification of phylogenetic markers followed by next generation sequencing is a useful technique to assess community composition and abundance of different taxonomic groups. The molecular markers used were directed towards the general prokaryotic 16S rDNA (V1 to V3 region; primers 27F, 519R; Lane et al, 1985, 1991) and eukaryotic V9 region from 18S rDNA (primers 1380F, 1510R; Amaral-Zettler et al, 2009). In addition a molecular marker of higher phylogenetic resolution *petB*, targeting marine

Synechococcus sub-cluster 5.1, was used (Mazard et al, 2012). The PCR primers were modified for the preparation of amplicon libraries for Illumina sequencing with adaptors and 8mer barcodes (as per the 16S rDNA metagenomic sequencing library preparation guide from Illumina, Inc.) fused to the 5' of the specific PCR primers for a one step PCR amplification reaction (Supplementary table 5.2).

PCR amplifications on all DNA samples using the *petB* modified oligonucleotide primers were conducted as previously described in Chapter 4. Amplicon libraries for the 18S rRNA gene were prepared in 50 µl PCR reactions, each reaction contained 25 pmol each of forward and reverse Illumina primers, 0.8 mM dNTPs, 1X reaction buffer, 1 unit Taq polymerase (Qiagen, Australia) and 5 – 10 ng of template DNA. The PCR program comprised a denaturation step at 94°C for 5 min, followed by 30 cycles of denaturation at 94°C for 20s, annealing at 57°C for 20s, extension at 72°C for 30s. The final step was an extension for 6 min at 72°C. Amplicon libraries for the 16S rRNA gene were prepared in 25 µl PCR reactions, each reaction contained 10 pmol each of forward and reverse Illumina primers and 1X reaction buffer (5 PRIME HotMasterMix) containing 45 mM KCl, 2.5 mM Mg^{2+}, 200 µM of each dNTP and 0.5 unit Taq polymerase (VWR, Australia) and 5 – 10 ng of template DNA. The PCR program comprised a denaturation step at 94°C for 2 min, followed by 35 cycles of denaturation at 94°C for 15s, annealing at 55°C for 20s, extension at 68°C for 30s. The final step was an extension for 7 min at 68°C.

All amplicons were quantified using Quant-iT PicoGreen dsDNA assays following the manufacturer's instructions (Life Technologies, Australia). Amplicons from each site were pooled (15/10 ng DNA per site for 18S rDNA or *petB* respectively) and purified using Agencourt AMPure XP bead purification (Beckman Coulter Inc., Australia). For 16S rDNA, 50 ng of amplicons from each site were pooled and purified using a MinElute 96 UF PCR purification plate (Qiagen Australia). Samples were submitted for sequencing at the Ramaciotti Centre for Genomics (NSW, Australia). The amplicons were sequenced using an Illumina MiSeq instrument. The pooled *petB* amplicons and the pooled 16S rDNA amplicons were sequenced on a 300bp Paired-End run, the shorter 18S rDNA amplicons were sequenced on a 150bp Paired-End run.

5.2.2.3 Bioinformatic analyses

The paired-end sequences for the 16S rDNA markers were joined (>80% joined) using the flash algorithm (Magoc and Salzberg, 2011). The sequences were then cleaned and processed through an in-house bioinformatic analysis pipeline based on the USEARCH-64 program (Edgar, 2010). Briefly, sequences were processed through an initial dereplication, and then sorted into clusters at 97% identity. Groups with less than 2 sequences were discarded, chimeras removed and the original file was mapped against the de-novo defined OTUs (USEARCH global) to create the output result table. The taxonomy for each OTU was assigned against the Silva119 release_99% redundant database (Yilmaz et al, 2014) using mother classify.seqs command using Knn, numwanted =3 (Schloss et al, 2009). The total sequences per site was subsequently rarefied to the smallest number of sequences obtained. The same method was followed for 18S rDNA amplicon processing. The OTU matrix with taxonomic assignment is provided in Supplementary tables 5.3 (16S rDNA) and 5.4 (18S rDNA).

The sequence reads for the *petB* markers were processed using the USEARCH64 (Edgar, 2010) OTU clustering pipeline with some modifications. Sequencing adaptors and primers were removed using qiime split_libraies.py, including sequence trimming for length (280 bp), quality (19) and demultiplexing the sequencing libraries. The trimmed sequences were collected into a single fasta file, dereplicated (using usearch –derep) filtered to remove sequences with less than 4 representatives before OTU clustering. For *de novo* OTU picking, dereplicated sequences were clustered at 97% identity, Chimeras were removed at this stage, and the resultant OTUs were further clustered at 94%. A mapping file was created containing *de novo petB* OTUs and *petB* closed reference sequences [using an up to date chimera-free *petB* sequence database containing sequences from complete genomes, cloned bidirectional sequences from the Warwick and Roscoff culture collections, and bidirectional sequences from clone libraries of environmental amplicons (Mazard et al, 2012; Humily et al, 2014)]. The quality trimmed sequence reads were then searched against the created mapping file containing reference sequences and *de novo* OTUs to produce mapped (chimeric and non-chimeric) and non-mapped reads (usearch –usearch-global). The mapped sequences were used to generate an

OTU table using the script uc2otutab.py accompanying usearch. Representative sequences clustered at 94% and the closed reference sequences were used to generate a multiple sequence alignment and a phylogenetic tree using ARB, which was then used to assign taxonomy and examine the phylogenetic placement of *de novo* OTUs. The OTU matrix with taxonomic assignment is provided in Supplementary table 5.5.

Results were visualised using R software (R Core Team, 2015). Bar-plots were generated for normalised taxonomic composition of the major community components (for 16S rDNA and 18S rDNA, with data tables filtered for taxonomic groups representing more than a total of 0.01 for the sum of abundance at all normalised sites).

In barplots, the colour scheme used for the sites is based on transects and distance from ice edge.

5.2.2.4 Statistical analyses

Ordination analyses (correspondence analysis, CA, and canonical correspondence analyses, CCA) and estimation of diversity indices were performed in R (R Core Team, 2015) using the Vegan package version 2.3-0 (Oksanen et al, 2015). The community composition in terms of taxonomic unit proportions were analysed with the available environmental parameter profiles at each site (temperature). The statistical significance of influence on the eukaryotic community structure, for the environmental parameter, was determined by ANOVA of the single factor CCA. The Bray-Curtis dissimilarity matrix was calculated for each sample of the 16S and 18S rDNA datasets through the vegdist and spantree algorithms from the Vegan package and rarefaction curves produced for all datasets. All results were visualised using R software.

5.2.3 *Results and discussion*

To explore the microbial community composition in the Southern Ocean, water samples collected along two latitudinal transects to the Totten and Mertz Glaciers were used to amplify phylogenetic markers from extracted DNA. Amplicons were obtained for all samples for the 16S and 18S rDNA

phylogenetic markers, whereas the *petB* phylogenetic marker retrieved amplicons only for samples Transit 1a to 11 and Transit 47 to 50 (Figure 5.1). The microbial diversity was analysed using Bray-Curtis dissimilarity and rarefaction plots with the non-aggregated OTUs defined at a cut-off of 97% as well as for aggregated OTUs for 18S (i.e. OTUs that were assigned the same taxonomy were collapsed into one group).

5.2.3.1 General prokaryotic communities

The prokaryotic fraction of the microbial plankton community was analysed using 16S rDNA amplicon libraries. The microbial community profiles were distinctly different in diversity and composition between the Mertz Glacier and Totten Glacier transects. The overall microbial diversity was higher in several of the northward Totten Glacier transect sites in comparison to the poleward Mertz Glacier transect (Figure 5.2a). The sites closest to the Totten Glacier (0.65°C to -1.8°C) appeared to be the most diverse with a decrease in diversity towards lower latitudes (2.9°C – 5.7°C) along the Totten Glacier transect. The opposite trend was observed for the Mertz Glacier transect with higher diversity at lower latitudes. Community composition profile of the Totten Glacier transect based on Bray-Curtis dissimilarity showed a clear distinction between most sites at the Totten Glacier and at warmer latitudes below 62°S (Figure 5.2b). Two of the warmest sites (1a and 2b) of the Mertz Glacier transect (>4°C) clustered with the warmer sites of the Totten Glacier transect. The remaining sites of the Mertz Glacier transect showed more local variation in community composition as evident from the several clusters. For several sites, the sampling level does not represent the maximum diversity (Figure 5.2a).

Distinct differences in the microbial community composition are observed along the two transects (Figure 5.3). The Alphaproteobacteria SAR11 is known to be the most abundant bacterioplankton in the marine environment (Morris et al, 2002). SAR11 was detected at all sites but was most prevalent in the lower latitude sites with warmer temperatures, particularly in the Totten Glacier transect (up to

48.5% relative abundance of 16S rDNA sequences). Wilkins et al (2013) reported that this group occurs at higher abundance north of the Antarctic Circumpolar Current.

The class Flavobacteria belonging to the order Bacteroidetes are a major component of the SO bacterioplankton (Abell and Bowman, 2005; Wilkins et al, 2013). In this study, the abundance of the class Flavobacteria was >10% in the majority of sites. The highest abundance (>50%) was at the three sites closest to the Mertz Glacier where Flavobacteriaceae was the dominant family. Genera of Flavobacteriaceae showed partitioning that may be temperature related. The genus *Polaribacter*, of Flavobacteriaceae, showed higher abundance (2.3-8%) at temperatures below 2.5°C and was mostly <1% in the remaining sites. In contrast to the occurrence of *Polaribacter*, another group of Flavobacteriaceae *NS4* (uncultured marine group) appears to be restricted to waters of temperatures higher than 2°C.

At the majority of sites of the Totten Glacier transect, Gammaproteobacteria accounted for >14.5% of the community sequences whilst in the Mertz Glacier transect, the majority of sites had <10%. In addition to temperature, other factors appear to influence their distribution since similar temperature regimes differ in taxonomic groups of Gammaproteobacteria. Classes Alteromonadales and Oceanospirillales compose the majority of Gammaproteobacteria. Alteromonadales occurred in a wide range of proportions (0.9 – 60.9%) in the Totten Glacier transect. The genus *Alteromonas*, the main group of Alteromonadales, displayed varied abundance with seemingly limited temperature influence. At site transit 36, the genus *Alteromonas* was the most abundant group (59.8% of the total community) which suggests a potential bloom. The occurrence of other genera of Alteromonadales suggests temperature-related spatial partitioning with more prevalence of *Glaciecola* at colder sites (< 0°C) whilst SAR92 and *Colwellia* at warmer sites (> 2°C). In the Mertz Glacier transect, Alteromonadales occurred at very low abundance (0.3-1.7%) with *Glaciecola* accounting for 0.5-1.1% at the colder sites. SAR92 which was found to be one of the prominent prokaryotes in sea ice edge community at McMurdo Sound in January 2013 (Bertrand et al, 2015), never occurred at more than 1.6% of the total community in either transect. Potential temperature-influenced partitioning is observed in another

class of Gammaproteobacteria, Oceanospirillales. This class comprising SAR86 as the major proportion was more abundant (2.9-8.5%) in warmer waters above 2.9˚C and <1.5% in the remaining sites in both transects.

At the majority of sites below 2.5°C, in both transects, a high abundance of 16S rDNA sequences were assigned to chloroplast (>30%) and mitochondria (>5%). The abundance changes of chloroplast and mitochondrial sequences mostly followed similar trends. In the Totten transect, higher abundance of chloroplast and mitochondrial sequences at the higher latitudes (i.e. 16–54% below and <16% above 63°S) suggests an increased presence of eukaryotic phytoplankton at colder high latitudes.

Temperature appears to play a significant role in variation in the prokaryotic community composition and diversity both within each transect and between the two Southern Ocean transects. Heterogeneity in the composition despite similar temperature conditions suggests the involvement of other environmental parameters in shaping the community.

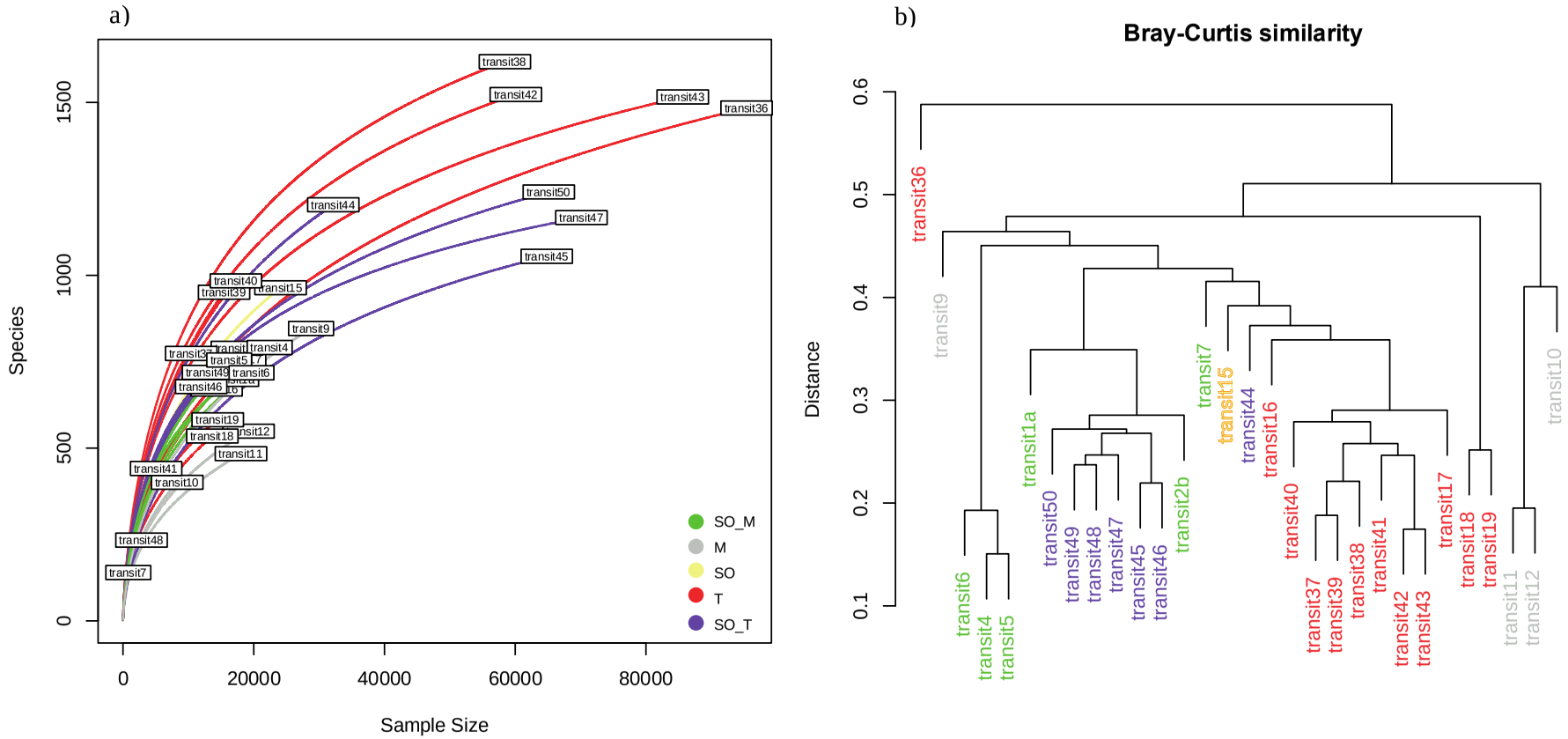
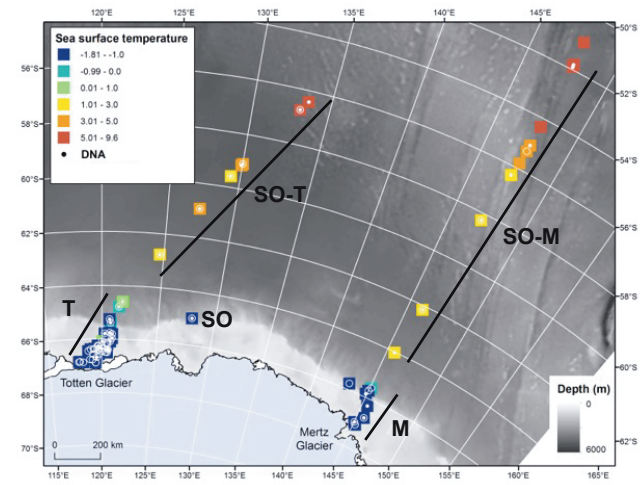


Figure 5.2: Representation of the microbial community diversity (16S rDNA) in Antarctic waters, a) rarefaction curves and b) Bray-Curtis dissimilarity. Samples were color-coded by regions: SO-M Southern ocean transect towards Mertz Glacier (green), M Mertz Glacier sea shelf (grey), SO transfer station at the edge of the ice shelf in the Southern Ocean towards Totten Glacier (orange), T Totten Glacier sea shelf (red), SO-T Southern Ocean transect from Totten Glacier (purple). The colour scheme used here is not related to that in Figure 5.1.



Bacteria
Bacteroidetes
Bacteroidetes
Flavobacteriales
Flavobacteriales
Cryomorphaceae;Brumimicrobium
Cryomorphaceae;Fluviicola
Cryomorphaceae;Owenweeksia
Cryomorphaceae;Owenweeksia;UC bact.
Cryomorphaceae;Owenweeksia;UC marine bact.
Flavobacteriaceae
Flavobacteriaceae;Flavobacterium
Flavobacteriaceae;Formosa
Flavobacteriaceae;Maribacter
Flavobacteriaceae;NS2b marine group;UC marine bact.
Flavobacteriaceae;NS4 marine group
Flavobacteriaceae;NS4 marine group;UC bact.
Flavobacteriaceae;NS5 marine group
Flavobacteriaceae;Polaribacter
Flavobacteriaceae;Psychroserpens
Flavobacteriaceae;Ulvibacter
Flavobacteriaceae;uncultured
Flavobacteriaceae;Winogradskyella;UC bact.
NS7 marine group
NS9 marine group
Sphingobacteriia;Sphingobacteriales;NS11-12 marine group
Sphingobacteriia;Sphingobacteriales;Saprospiraceae
Sphingobacteriia;Sphingobacteriales;Saprospiraceae;Lewinella
Sphingobacteriia;Sphingobacteriales;Saprospiraceae;uncultured
Candidate division OD1;uncultured bacterium
Cyanobacteria
Cyanobacteria;Chloroplast
Cyanobacteria;Chloroplast;uncultured bacterium
Proteobacteria
Proteobacteria
Alphaproteobacteria
Alpha
Rhodobacterales;Rhodobacteraceae
Rhodobacterales;Rhodobacteraceae;Candidatus Planktomarina DC5-80-3
Rhodobacterales;Rhodobacteraceae;Loktanella
Rhodobacterales;Rhodobacteraceae;Octadecabacter
Rhodobacterales;Rhodobacteraceae;Roseobacter clade NAC11-7 lineage
Rhodobacterales;Rhodobacteraceae;Sulfitobacter
Proteobacteria
Alpha
Rhodobacterales;Rhodobacteraceae;Sulfitobacter;UC marine bact.
Rhodobacterales;Rhodobacteraceae;uncultured
Rhodospirillales;Rhodospirillaceae;AEGEAN-169 marine group
Rickettsiales
Rickettsiales;mitochondria
Rickettsiales;mitochondria;UC bact.
Rickettsiales;SAR116 clade
SAR11 clade
SAR11 clade;Chesapeake-Delaware Bay
SAR11 clade;Surface 1
SAR11 clade;Surface 1;UC marine bact.
SAR11 clade;Surface 4;UC alpha proteobact.
Sphingomonadales;Sphingomonadaceae
Betaproteobacteria
Betaproteobacteria;Methylophilales;Methylophilaceae;Methylotenera
Epsilonproteobacteria;Campylobacterales;Helicobacteraceae;Sulfurovum
Gamma
Gammaproteobacteria
Alteromonadales
Alteromonadales;Alteromonadaceae
Alteromonadales;Alteromonadaceae;Alteromonas
Alteromonadales;Alteromonadaceae;Glaciecola
Alteromonadales;Alteromonadaceae;OM60(NOR5) clade;UC marine bact.
Alteromonadales;Alteromonadaceae;SAR92 clade
Alteromonadales;Colwelliaceae;Colwellia
Alteromonadales;Colwelliaceae;Colwellia;UC bact.
Alteromonadales;Moritellaceae;Moritella
BD7-8 marine group
Gammaproteobacteria Incertae Sedis;Unknown Family;Arenicella
KI89A clade
Oceanospirillales
Oceanospirillales;JL-ETNP-Y6
Oceanospirillales;Oceanospirillaceae;Neptunomonas
Oceanospirillales;SAR86 clade
Oceanospirillales;SAR86 clade;UC marine bact.
Pseudomonadales;Pseudomonadaceae;Pseudomonas
Thiotrichales;Piscirickettsiaceae;Thiomicrospira
Thiotrichales;Piscirickettsiaceae;uncultured;UC bact.
Thiotrichales;Thiotrichaceae;Leucothrix;UC bact.

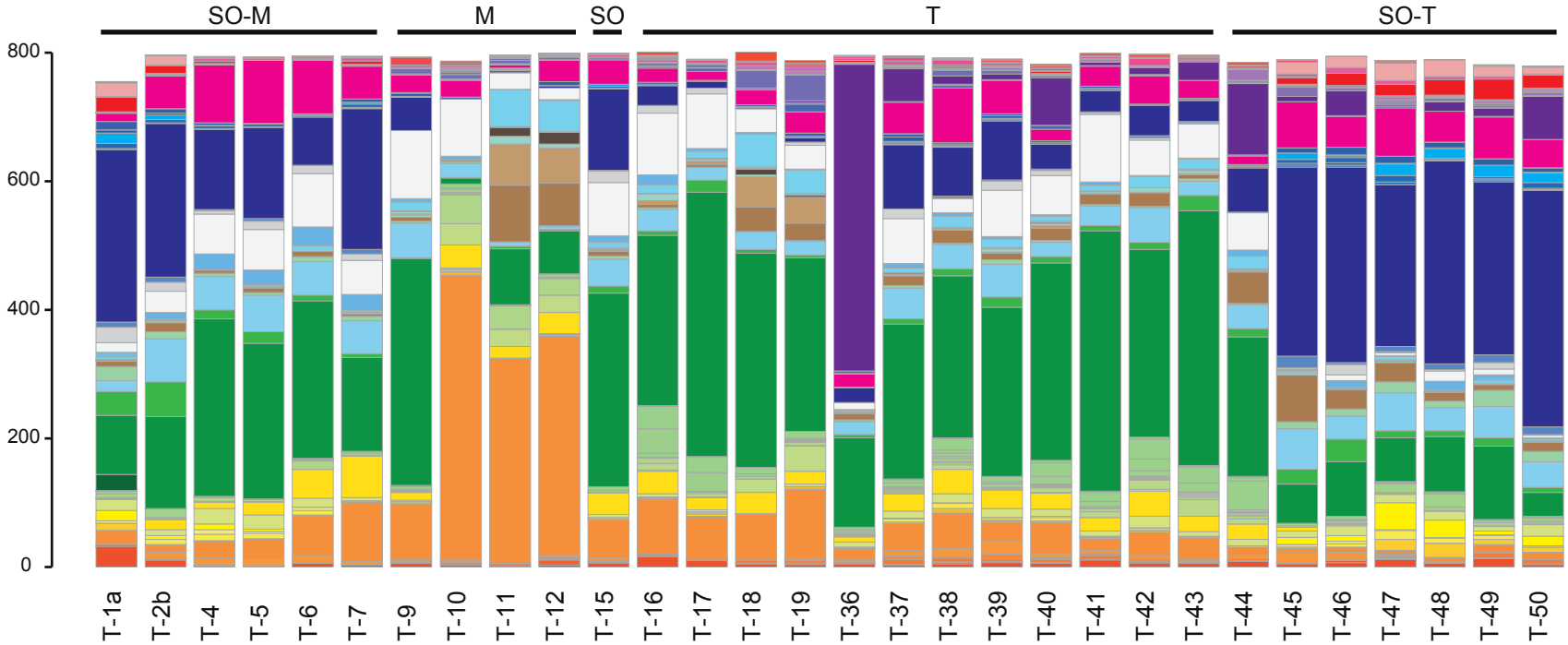


Figure 5.3: Barchart representation of the overall planktonic microbial component in Antarctic waters. The community composition was derived from the V1-V3 region of the 16S rDNA (Lane et al, 1985, 1991) through amplicon sequencing. Only the major groups with proportions greater than 1% of the total community are shown. As indicated on the map, sampling points were grouped by regions: SO-M Southern Ocean transect towards Mertz Glacier (sites 1a to 7), M Mertz Glacier sea shelf (sites 9 to 12), SO transfer station at the edge of the ice shelf in the Southern Ocean towards Totten Glacier (site 15), T Totten Glacier sea shelf (sites16 to 43), SO-T Southern Ocean transect from Totten Glacier (sites 44 to 50).

5.2.3.2 Photosynthetic eukaryotes

Amplicon libraries of the V9 region of the 18S rDNA phylogenetic marker (Amaral-Zettler et al, 2009) were obtained for all samples. All sites present a large diversity of eukaryotic microbes. Warmer sites have nearly twice the total estimated number of assigned species than colder sites (~>300 and ~150-200 respectively; Figure 5.4a). Unexpectedly, the two transects present different general diversity characteristics, with the northward Totten Glacier transect sites presenting an average of 50 more assigned species than the downwards Mertz Glacier transect samples. There appears to be no global change of diversity within each of the distinct transect sites, below a threshold of sea surface temperature of approximately 4˚C.

Bray Curtis dissimilarity of the eukaryotic microbial community composition was used to visualise the relations between sites (Figure 5.4b). There is a clear separation of the communities between the northern sites of each transects and the southern sites (~2.5˚C). Communities from Mertz Glacier waters cluster separately and, hence, appear distinct from the ones from Totten Glacier despite the relative proximity of the sites. Interestingly, some of the sites close or just above the ice edge appear to have more similar communities (transit 7, 15 and transit 44, 36) and cluster together between the two clusters of sites from the glaciers. These might represent transitional communities undergoing changes under variable conditions created by the melting of the ice due to the retreat of the ice edge. The variations in the clustering of sites within the Southern Ocean waters could be the reflection of the strong currents and fronts around the Antarctic (e.g. Antarctic Circumpolar Current, Polar Front and Antarctic divergence).

Analysis of the eukaryotic community profile at a low taxonomic resolution (family-level or higher) shows multiple classes of organisms in the warmest sites from both transects (1a/2b and 49/50; Figure 5.5), including Prymnesiophyceae, Stramenopiles such as Pelagophyceae, Picobiliphyta, Syndiniales and other unclassified Dinophyta. Overall the cold sites show a higher abundance of Dinophyceae (up to 36.9% at transit 12) and heterotrophic Ciliophora (Litostomatea, Phyllopharyngea, Spirotrichea; up

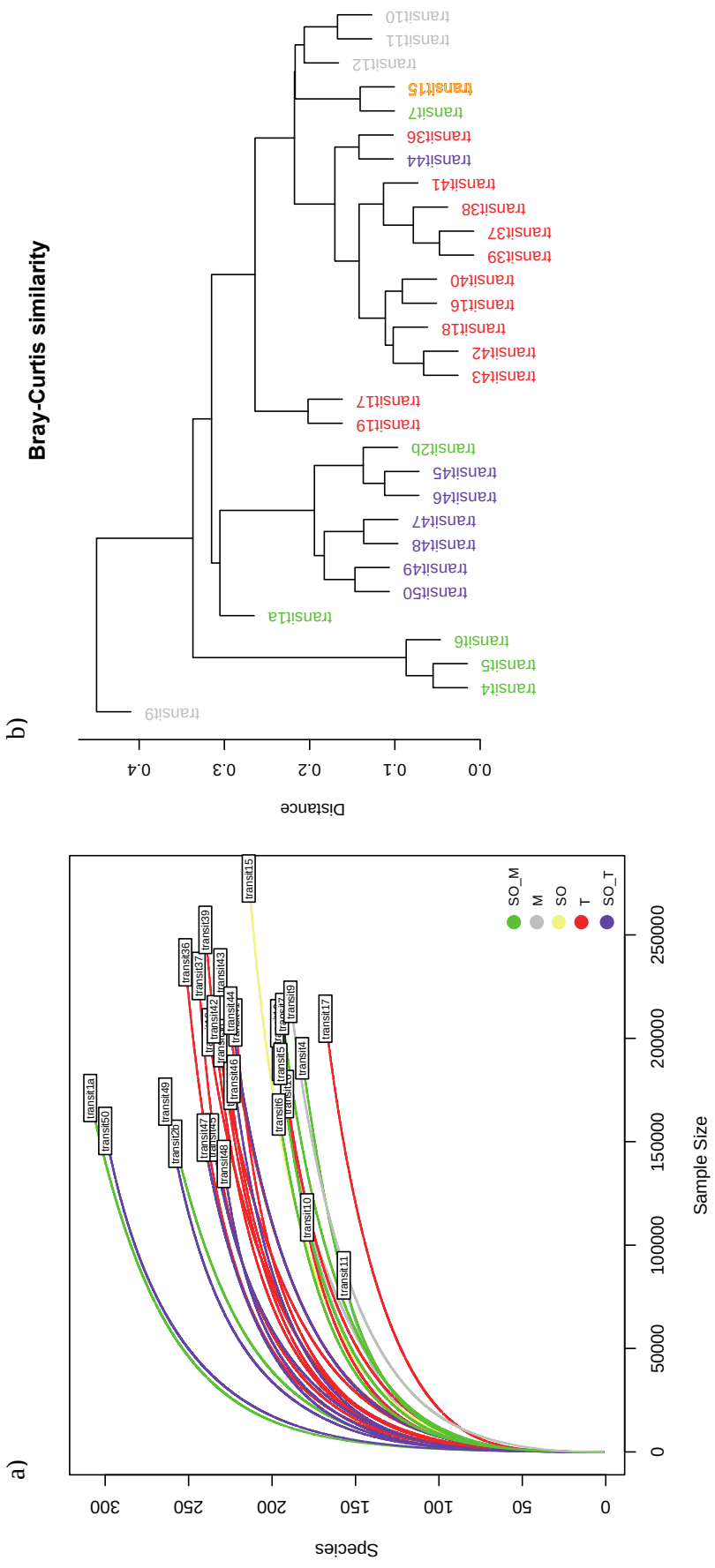
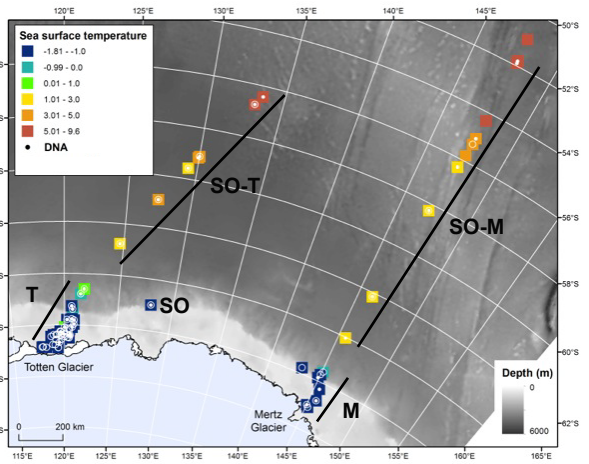


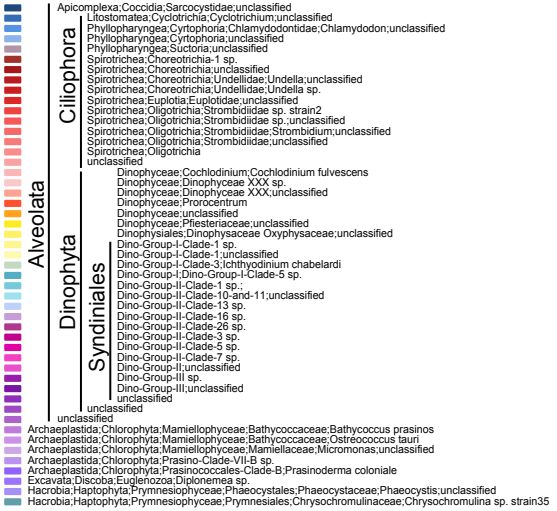
Figure 5.4: Representation of the eukaryotic microbial diversity (18S rDNA) in Antarctic waters, a) rarefaction curves and b) Bray-Curtis dissimilarity. Samples were color-coded by regions: SO-M Southern Ocean transect towards Mertz Glacier (green), M Mertz Glacier sea shelf (grey), SO transfer station at the edge of the ice shelf in the Southern Ocean towards Totten Glacier (orange), T Totten Glacier sea shelf (red), SO-T Southern Ocean transect from Totten Glacier (purple). The colour scheme used here is not related to that in Figure 5.1.

to 15.6% at transit 19). However, a clear difference can be visible between the colder sites from the Mertz Glacier transect dominated by diatoms (Bacillariophyta, up to 62.3% transit 4), whereas the colder sites from the Totten Glacier transect present a larger proportion of unclassified Alveolata (up to 52.4% at transit 43).

The community composition at a higher resolution (family-level or lower) reveals some finer local speciation and distinct variation in eukaryotes (Figure 5.5). Along the Mertz Glacier transect, a large proportion of centric diatoms (unclassified Mediophyceae) can be observed in sites 4-6 (up to 47.7%), while a bloom of unclassified Bacillariophyta appears in site 9 (51.5%). A higher ratio of centric to pennate diatoms is observed at the transitional sites of the Totten Glacier to Southern Ocean transect (sites 41 to 45). This high centric diatom abundance contrasts to other sites throughout the transects where equal proportion of centric:pennate diatoms is found. The various species of centric diatoms show differing spatial distribution with *Chaetoceros* restricted to the cold sites of both Mertz and Totten Glaciers (up to 10.2% at transit 41). *Thalassiosira*, however, is found in warmer but still intermediate sites. It is also found in some cold sites, e.g. transit 17 and 19. At these sites where *Thalassiosira* increase, the proportion of *Chaetoceros* is lower. The variations in the microbial community from site to site reflect the heterogeneity of the region, mainly in the glacier areas, enhanced by the retreat of the sea ice at the time of sampling.

The warmer sites present distinct eukaryotic populations with groups of eukaryotic microbes not found at the colder sites. The Syndiniales dino group II and to a lesser extend group I can be seen mainly in sites 1a/2b and 45-50, up to 8.7% and 3.3% for group II and I respectively, whilst at < 1% in the other samples. Dino group III though at much lower abundance (0.5%) is also only detected at these sites. Dino group II clade 16 is only seen at the warmer site transit 1a (8.1˚C, 2.3% of the total community) and is negligible at all other sites, indicating that this clade might have a thermal niche lower limit between 6-8˚C. The other groups restricted to the warmer sites are the Pelagophyceae (*Aureococcus anophagefferens*, 7.4% at transit 1a) and the Prymnesiophyceae (*Phaeocystis*, 13.15% at transit 49).


Sea surface temperature
-1.81 - -1.0
-0.99 - 0.0
0.01 - 1.0
1.01 - 3.0
3.01 - 5.0
5.01 - 9.6
DNA
SO-T
SO-M
T
SO
M
Totten Glacier
Mertz Glacier
Depth (m)
0
6000
0
200 km


Alveolata
Ciliophora
Apicomplexa;Coccidia;Sarcocystidae;unclassified
Litostomatea;Cyclotrichia;Cyclotrichium;unclassified
Phyllopharyngea;Cyrtophoria;Chlamydodontidae;Chlamydodon;unclassified
Phyllopharyngea;Cyrtophoria;unclassified
Phyllopharyngea;Suctoria;unclassified
Spirotrichea;Choreotrichia-1 sp.
Spirotrichea;Choreotrichia;unclassified
Spirotrichea;Choreotrichia;Undellidae;Undella;unclassified
Spirotrichea;Choreotrichia;Undellidae;Undella sp.
Spirotrichea;Euplotia;Euplotidae;unclassified
Spirotrichea;Oligotrichia;Strombidiidae sp. strain2
Spirotrichea;Oligotrichia;Strombidiidae sp.;unclassified
Spirotrichea;Oligotrichia;Strombidiidae;Strombidium;unclassified
Spirotrichea;Oligotrichia;Strombidiidae;unclassified
Spirotrichea;Oligotrichia
unclassified
Dinophyta
Syndiniales
Dinophyceae;Cochlodinium;Cochlodinium fulvescens
Dinophyceae;Dinophyceae XXX sp.
Dinophyceae;Dinophyceae XXX;unclassified
Dinophyceae;Prorocentrum
Dinophyceae;unclassified
Dinophyceae;Pfiesteriaceae;unclassified
Dinophysiales;Dinophysaceae Oxyphysaceae;unclassified
Dino-Group-I-Clade-1 sp.
Dino-Group-I-Clade-1;unclassified
Dino-Group-I-Clade-3;Ichthyodinium chabelardi
Dino-Group-I;Dino-Group-I-Clade-5 sp.
Dino-Group-II-Clade-1 sp.;
Dino-Group-II-Clade-10-and-11;unclassified
Dino-Group-II-Clade-13 sp.
Dino-Group-II-Clade-16 sp.
Dino-Group-II-Clade-26 sp.
Dino-Group-II-Clade-3 sp.
Dino-Group-II-Clade-5 sp.
Dino-Group-II-Clade-7 sp.
Dino-Group-II;unclassified
Dino-Group-III sp.
Dino-Group-III;unclassified
unclassified
unclassified
Archaeplastida;Chlorophyta;Mamiellophyceae;Bathycoccaceae;Bathycoccus prasinos
Archaeplastida;Chlorophyta;Mamiellophyceae;Bathycoccaceae;Ostreococcus tauri
Archaeplastida;Chlorophyta;Mamiellophyceae;Mamiellaceae;Micromonas;unclassified
Archaeplastida;Chlorophyta;Prasino-Clade-VII-B sp.
Archaeplastida;Chlorophyta;Prasinococcales-Clade-B;Prasinoderma coloniale
Excavata;Discoba;Euglenozoa;Diplonemea sp.
Hacrobia;Haptophyta;Prymnesiophyceae;Phaeocystales;Phaeocystaceae;Phaeocystis;unclassified
Hacrobia;Haptophyta;Prymnesiophyceae;Prymnesiales;Chrysochromulinaceae;Chrysochromulina sp. strain35
Hacrobia;Haptophyta;Prymnesiophyceae;Prymnesiales;Chrysochromulinaceae;Chrysochromulina sp.
Hacrobia;Haptophyta;Prymnesiophyceae;Prymnesiales;Chrysochromulinaceae;Chrysochromulina;unclassified
Hacrobia;Haptophyta;Prymnesiophyceae;unclassified
Hacrobia;Katablepharidophyta;Katablepharidaceae;Katablepharidales;unclassified
Hacrobia;Picobiliphyta sp.
Hacrobia;Picobiliphyta;unclassified
Hacrobia;Telonemia-Group-1 X sp.
Hacrobia;Telonemia-Group-1;unclassified
Hacrobia;Telonemia-Group-2 X sp.
Hacrobia;unclassified
Rhizaria
Cercozoa;Filosa-Thecofilosea;Cryomonadida;Protaspa-lineage X sp.
Cercozoa;Filosa-Thecofilosea;Cryomonadida;Protaspa-lineage;Protaspa;unclassified
Cercozoa;Filosa-Thecofilosea;Cryomonadida;unclassified
Cercozoa;Filosa-Thecofilosea;Ebriida;TAGIRI1-lineage sp.
Cercozoa;Filosa-Thecofilosea;unclassified
Cercozoa;unclassified
Radiolaria;Acantharea;Arthracanthida-Symphyacanthida;unclassified
Radiolaria;RAD-B-Group-IV sp.
Stramenopiles
Ochrophyta
Bacillariophyta
Araphid-pennate;unclassified
Polar-centric-Mediophyceae;Chaetoceros;unclassified
Polar-centric-Mediophyceae;Cyclotella;unclassified
Polar-centric-Mediophyceae;Eucampia;unclassified
Polar-centric-Mediophyceae;Thalassiosira;unclassified
Polar-centric-Mediophyceae;unclassified
Radial-centric-basal-Coscinodiscophyceae;Stellarima microtrias
Radial-centric-basal-Coscinodiscophyceae;unclassified
Raphid-pennate;Pseudo-nitzschia;unclassified
Raphid-pennate sp.
Raphid-pennate;unclassified
unclassified
Chrysophyceae-Synurophyceae;Clade-F;Paraphysomonas;unclassified
Chrysophyceae-Synurophyceae;Clade-H sp.
Dictyochophyceae;Dictyochales;Dictyocha speculum
Dictyochophyceae;Florenciellales;Pseudochattonella verruculosa
Pelagophyceae;Aureococcus anophagefferens
Pelagophyceae;Pelagomonas calceolata
Pelagophyceae;unclassified
unclassified
Stramenopiles X;MAST;MAST-1;MAST-1A sp.
Stramenopiles X;MAST;MAST-3;MAST-3E sp.
Stramenopiles X;MAST;MAST-3;unclassified
Stramenopiles X;MAST;MAST-7;MAST-7A sp.
Stramenopiles X;MAST;MAST-7;unclassified
Stramenopiles X;MAST;MAST-8;unclassified
Stramenopiles X;MAST;unclassified
Stramenopiles X;unclassified
unclassified
unclassified

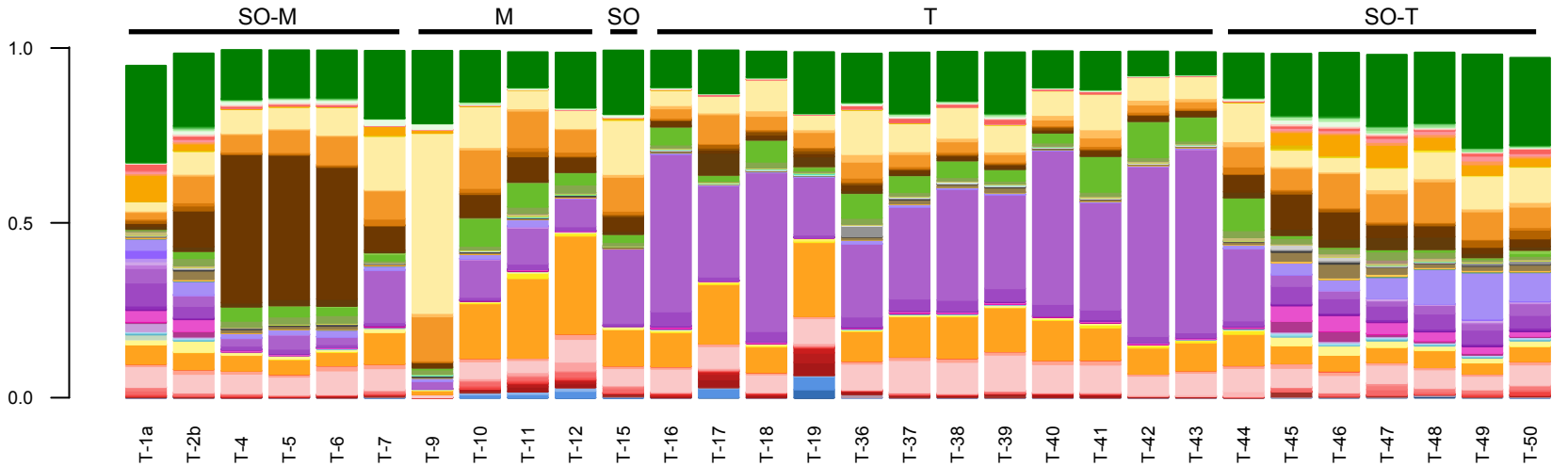

SO-M
M
SO
T
SO-T
1.0
0.5
0.0
T-1a
T-2b
T-4
T-5
T-6
T-7
T-9
T-10
T-11
T-12
T-15
T-16
T-17
T-18
T-19
T-36
T-37
T-38
T-39
T-40
T-41
T-42
T-43
T-44
T-45
T-46
T-47
T-48
T-49
T-50

Figure 5.5: Barchart representation of the eukaryotic microbial component in Antarctic waters. The community composition was derived from the V9 region of the 18S rDNA (Amaral-Zettler et al, 2009) through amplicon sequencing. Only the major groups with proportions greater than 1% of the total community are shown. As indicated on the map, sampling points were grouped by regions: SO-M Southern Ocean transect towards Mertz Glacier (sites 1a to 7), M Mertz Glacier sea shelf (sites 9 to 12), SO transfer station at the edge of the ice shelf in the Southern Ocean towards Totten Glacier (site 15), T Totten Glacier sea shelf (sites16 to 43), SO-T Southern Ocean transect from Totten Glacier (sites 44 to 50).

Interestingly, a small proportion of *Pseudochattonella verruculosa* was detected at transit 45. *P. verruculosa* is an ichthyotoxic heterotrophic flagellate associated with mixed harmful algal blooms (Hosoi Tanabe et al, 2007; Hoe Chang et al, 2014). There is also a higher proportion of the some of the bacterivorous MAST groups in the warmer sites despite still in relatively low representation, though MAST1A being detected at >1% in the Totten Glacier sites 37/39.

Other groups of Syndiniales, such as the dino group II clades 26 (up to 3%) as well as Picobiliphyta (up to 4.5%) were detected but restricted to the intermediate region between cold and warm waters possibly delimitating the polar front (transit 45-46, 2b).

A recent report by Bertrand et al (2015) presented an overview of the eukaryotic community composition in Southern ocean sea ice edge in 2013 at McMurdo Sound (165˚E), situated to the east of our study sites. Totten Glacier and McMurdo Sound are equidistant at either side from the Mertz Glacier (~1000 km). There pennate diatoms dominated the community representing close to half of the population, centric diatoms, dinophyta and ciliophora composing most of the rest while other eukaryotes represented just about 5% of the total community. The differences seen could be due to the influence of the Ross Sea gyre (westward) on the McMurdo Sound communities, against the Antarctic circumpolar current (eastward) for our sites.

Ordination analyses were performed on the eukaryotic profiles alone (Figure 5.6) as well as using a very restricted set of environmental factors (data not shown). The different samples separated along the two main dimensions and grouped in similar clusters as obtained from Bray Curtis distance. The sea surface temperature factor was strongly significant (p <0.001), and alone explained 30.9% of the community composition. However, a large amount of the distribution remains unexplained and further biotic and abiotic factors need to be assessed.

While temperature plays a clear and significant role in the diversity and composition of eukaryotic microbes along the two transects, it is clear that other factors influence the community composition. Indeed, the marked difference observed between the two geographically close transects, particularly at

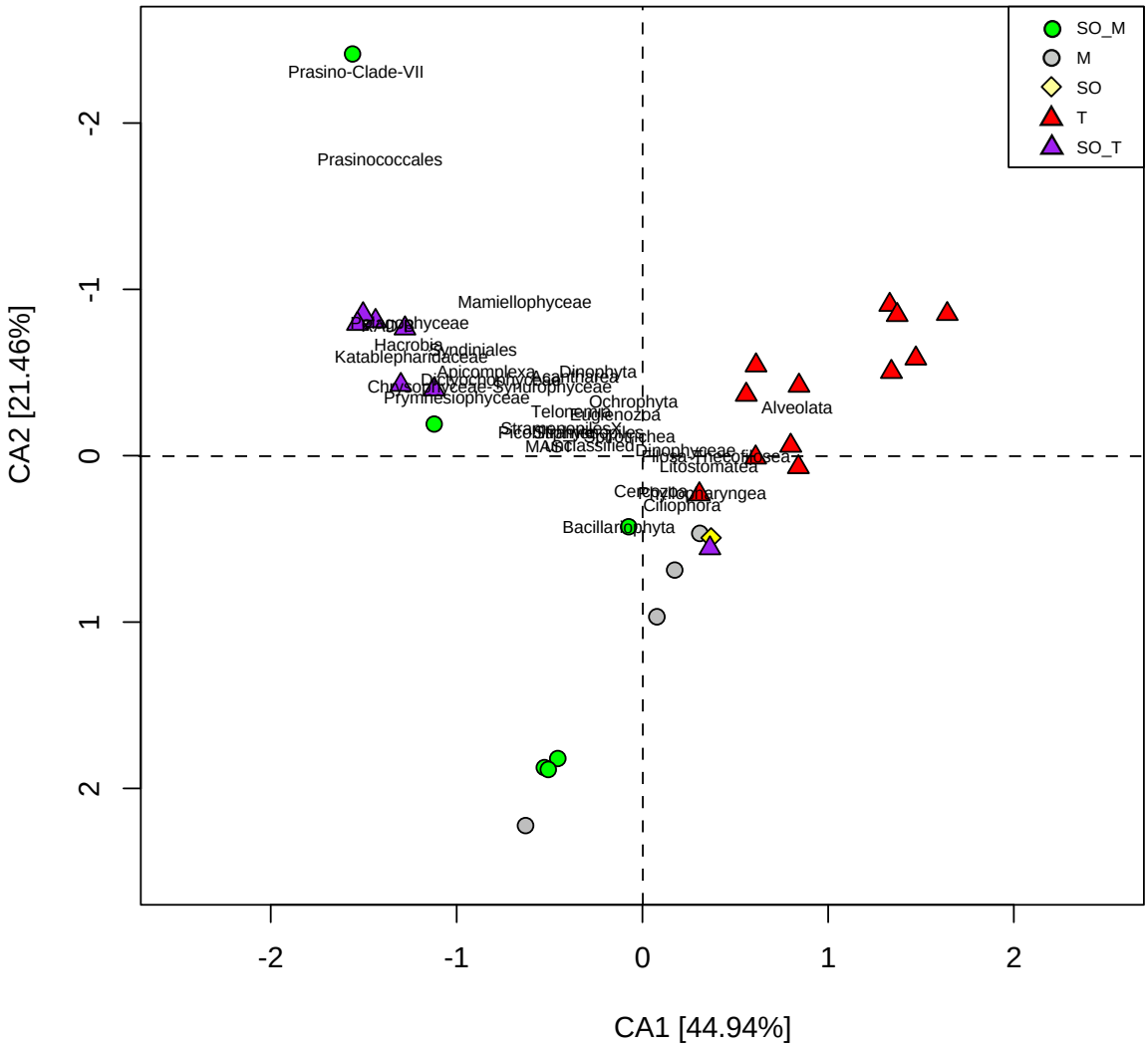


Figure 5.6: 2D representation of correspondence analysis of the sites according to their eukaryotic microbial community composition (groups >1% of the total community). Sites and eukaryotic component to the level of the class are represented on the two first dimensions CA1 and CA2. The percentage of the total inertia resolved from these dimensions is indicated as percentage next to each axis. Sites coloured according to defined groups: SO-M (green dots) Southern Ocean transect towards Mertz Glacier (sites 1a to 7), M (grey dots) Mertz Glacier sea shelf (sites 9 to 12), SO (yellow diamond) transfer station at the edge of the ice shelf in the Southern Ocean towards Totten Glacier (site 15), T (red triangles) Totten Glacier sea shelf (sites16 to 43) , SO-T (purple triangles) Southern Ocean transect from Totten Glacier (sites 44 to 50).

the colder sites indicates local heterogeneity of the sites. Furthermore, another close site reported by Bertrand et al (2015) also presents variations from the two transects presented here. There could be a strong influence of the annual sea ice cover and residence. It is possible to argue that at these temperatures, the biology, i.e. community structures dynamics, could be slow and the community changes delayed from the retreat of the ice edge to open Southern Ocean waters. Indeed some of the sites, e.g. transit 44, present communities that have the characteristics of both cold and warm waters, though the physical parameters would group these sites with the "warmer" regions sampled.

The observed shift in the community structure, from diatoms in the Southern Ocean waters to Alveolate prominence in areas below 1˚C with occurrence of sea ice, e.g. in the colder sites of the Totten Glacier – Southern Ocean transect, could indicate a different trophic regime with various levels of trophic links to higher levels in the food chain between the different photosynthetic eukaryote groups (Massana, 2011). The cold conditions with the extended period of darkness encountered in polar regions, are known to favour mixotrophic species that take advantage of their photosynthetic capacities in summer while surviving as heterotrophs during the winter where light is unavailable (Laybourn-Parry, 2002). All major lineages of photosynthetic eukaryotes have shown capacity for mixotrophy with the exception of diatoms (Sherr and Sherr, 2002) supporting the observations of this study.

5.2.3.3 Marine Synechococcus sub-cluster 5.1

It has been suggested that cyanobacteria do not contribute significantly to primary production in Antarctic marine ecosystems. Low number of cells of the picocyanobacteria, *Synechococcus*, have been detected in high latitudes down to 69°S (Koh et al, 2012; Wilmotte et al, 2002) whilst there are unpublished reports of *Prochlorococcus* in the Antarctic (Post, personal communication). In this study, using 16S rDNA, cyanobacteria were detected at site 1a (2.3% of the total microbial community) which had the highest temperature (8.1°C) compared to other sites. At other sites, the percentage was very low (<0.08%).

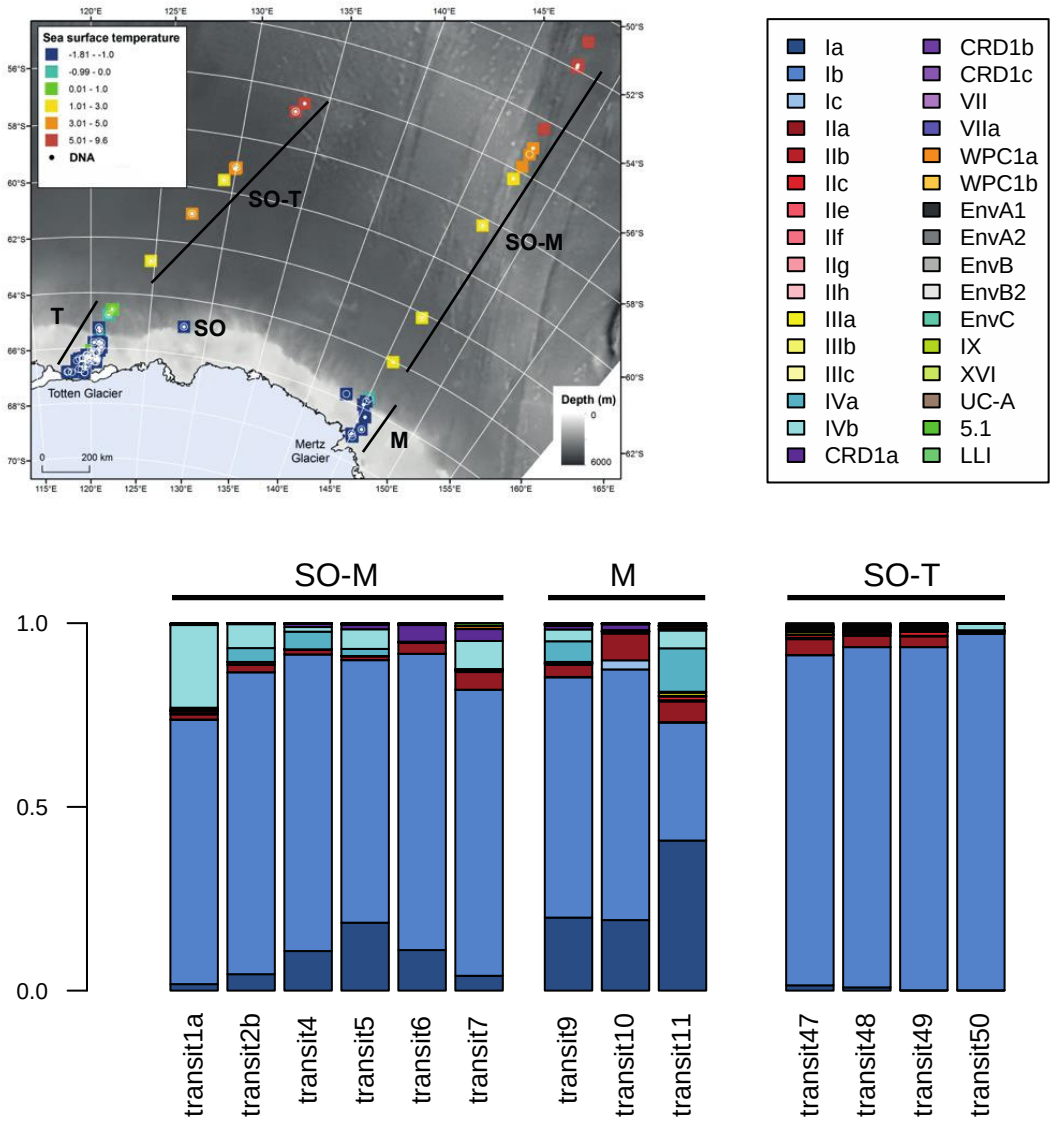


Figure 5.7: Barchart representation of the marine *Synechococcus* component in Antarctic waters. The community composition was derived from the *petB* gene (Mazard et al, 2012) through amplicon sequencing. As indicated on the map (see fig. 1 for further details), sampling points were grouped by regions: SO-M Southern Ocean transect towards Mertz Glacier (sites 1a to 7), M Mertz Glacier sea shelf (sites 9 to 12), SO-T Southern Ocean transect from Totten Glacier (sites 44 to 50).

Due to its specificity for *Synechococcus* sub-cluster 5.1, the *petB* phylogenetic marker would enable their detection even when present at very low numbers. Following two rounds of PCR amplification (total number of cycles of 50), *petB* amplicons were obtained for 13 of the 30 samples. The lack of

amplified product in some samples suggests very low abundance beyond the detection limit or the total absence of picocyanobacteria in those regions. The cells may also be present as a result of transport through air or water currents rather than residence.

Among these sites Clade Ib was the most prevalent *Synechococcus* lineage detected along both transects in the Southern Ocean (32.1-97.3%; Figure 5.7). The relative abundance of clade Ia mostly increased down the transect towards Mertz Glacier (1.8-40.8%). Clades IVa and IVb were prevalent (up to 11.8% and 22.5%, respectively) in the Mertz Glacier transect but was very low or not detected in the Totten Glacier transect. The detection of clades I and IV over other clades is not surprising since their higher relative abundance in high latitude niches and preference for lower temperatures below 20°C has been well described in other oceanic environments (Sohm et al, 2015; Zwirglmaier et al, 2008).

The detection of clade II in the Southern Ocean at temperatures below 10°C is unexpected since they are mostly found at lower latitudes and warmer temperature conditions. However, it remains to be determined if they are part of the resident population or were transported.

5.2.4 *Conclusions and ongoing work*

The observed microbial communities show quite high heterogeneity even within two close geographical transects, which reflect the dynamics of the regions studied mainly in relation to the annual variation of sea ice cover. Bertrand and co-workers (2015) reported interesting possible links between the various microbial populations and the nature of the trophic links needs to be investigated further in our study.

In agreement with previous observations of the thermal limits of photosynthetic prokaryotic groups (Koh et al, 2012; Wilmotte et al, 2002), picocyanobacteria were only present in warmer samples and do not play a major role in the primary production occurring in polar regions. Instead large populations of eukaryotic microbes, mostly diatoms and dinoflagellates, appear to be the dominant

primary producers. Furthermore, the shift of eukaryotic communities observed from diatoms to alveolate prominence in the cold sites correlates with the idea that mixotrophic communities are at an advantage in the extreme polar conditions (photosynthesis during the lit summer months while heterotrophic in the extended darkness of the austral winter).

The differences of microbial communities seen between the present study and the recent report from Bertrand et al (2015) could be due to the influence of the Ross Sea gyre (westward) on the McMurdo Sound communities, against the Antarctic circumpolar current (eastward) for our sites.

Temperature has a large influence in the community composition (~30%), however, it is evident that other factors play a major part in explaining the spatial distribution of microbial groups within these waters. Further analyses of these samples are currently on-going within the Paulsen group. Other parameters such as nutrients measured during both transects and features such as ice cover prior to sampling will be analysed and further work will be performed with the datasets (e.g. assignment of the chloroplasts and mitochondria from the 16S rDNA study). Flow cytometry will be applied to obtain cell counts and cells will be sorted for metagenomics sequencing and/ or single cell genomics of the eukaryotic phytoplankton. Possible links between the phytoplankton composition and the production of biochemical markers of sea-ice extent will be explored.

5.3 Sydney Harbour PPEs

5.3.1 *Introduction*

In conjunction with the study of the prokaryotic communities within the Sydney Harbour estuary (Chapter 4), the eukaryotic microbial communities were investigated. The main aim of this work is to characterise the major microbial populations involved in the primary production within Sydney Harbour and their seasonal variation.

5.3.2 *Materials and methods*

Sampling and preparation of the DNA is as described in the manuscript from Chapter 4. Amplification of the material with Illumina tagged primers targeting the V9 region of the 18S rDNA phylogenetic marker as well as the bioinformatics pipeline used and statistical analyses performed were as described in section 5.2. The OTU matrix with taxonomic assignment is provided in Supplementary table 5.6.

5.3.3 *Results and discussion*

5.3.3.1 General diversity

The overall community composition as analysed by Bray Curtis distance shows that sites cluster together by season (Figure 5.8b). The most inland sites possibly more influenced by freshwater input are the most dissimilar compared to the bulk SH sites (F: MH1, P2/3; S: LC1 LPR1 and P1). Within the cluster for each season, there is an approximate gradient of inland sites to marine sites, which is more marked in summer (February) than late winter. Rarefaction curves show that there is higher species richness in the marine sites, irrespective of season, compared to the more inland sites (Figure 5.8a).

5.3.3.2 Phylogenetic groupings

The eukaryotic community composition profile along the harbour shows dramatic changes both across the various sampling sites and between seasons (Figure 5.9). Overall, there is a higher proportion of general unassigned groups towards marine end of the harbour. In February, the most western sites (Parramatta River and Lane Cove) present a large proportion of centric diatoms: *Thalassiosira*, *Skeletonema costatum*, *Cyclotella*, as well as other unclassified polar-centric Mediophyceae. These communities disappear in the western part of the Harbour, towards the marine end. *Skeletonema* is often a major species in Atlantic coastal waters and characteristic diatom of coastal eutrophication (Abdalla et al, 1995) often dominant in waters with salinity fluctuation (Rijstenbil, 1987). While

centric diatoms are still present in the central harbour, these are represented by different groups, e.g. *Minutocellus* and other unclassified *Skeletonema*. In contrast, there is a larger proportion of non-photosynthetic Alveolata representatives (Cilliophora and Syndiniales). All sites, with the exception of the furthest inland sites present a significant proportion of Syndiniales dino group 1 of the clade 1. Interestingly, the proportion of various Dinophyta (Syndiniales dino group 1 and II and Dinophyceae) resembles more sediment environments than planktonic water column (Guillou et al, 2008).

Within the main harbour, eastwards to LC5, there is an occurrence of Ochrophyta, mainly *Aureococcus* and *Heterosigma* genera (up to a total of 16.5% in LPR13 and 17.8% in MH2). These are known to be involved in the formation of harmful algal blooms through surface aggregations also known as brown tides (Worden and Not, 2008).

The Archaeplastida and Hacrobia fraction of the photosynthetic picoeukaryotes (some Hacrobia are known to be bacterivorous) represent a significant proportion of eukaryotes in the Eastern sites and more particularly at the Middle Harbour sites (average of 23.5% of the site community throughout Middle Harbour and up to 33.4% at MH2). The Middle Harbour sites present very different community profile than rest of the harbour. There is an increased proportion of Chlorophyta (up to 26.7% of the total community), and more particularly *Micromonas* spp. (MH2 to MH6 ranging from 5.1 to 12.9%). Cryptophyceae are detected in the inland sites of that branch, mainly MH1 (13.9%) and ~MH2 (5.8%), this group has previously been reported in several studies to be more abundant in coastal systems than offshore (Massana, 2011). There is also a larger proportion of Cercozoa (Filosa-imbricatea novel clade 2, *Ebria*, up to 11.8% of the total community) in the sites MH2 to MH6 whereas these are negligible within the rest of the harbour.

In the September samples, a large proportion of raphid pennate can be observed in the Parramatta River sites (68.8% in LPR1, decreasing rapidly to 3.6% in LPR6). Local high levels of polar centric diatom from unclassified *Skeletonema* are distributed along the harbour (LC2, P5, PJ3, MH3-5).

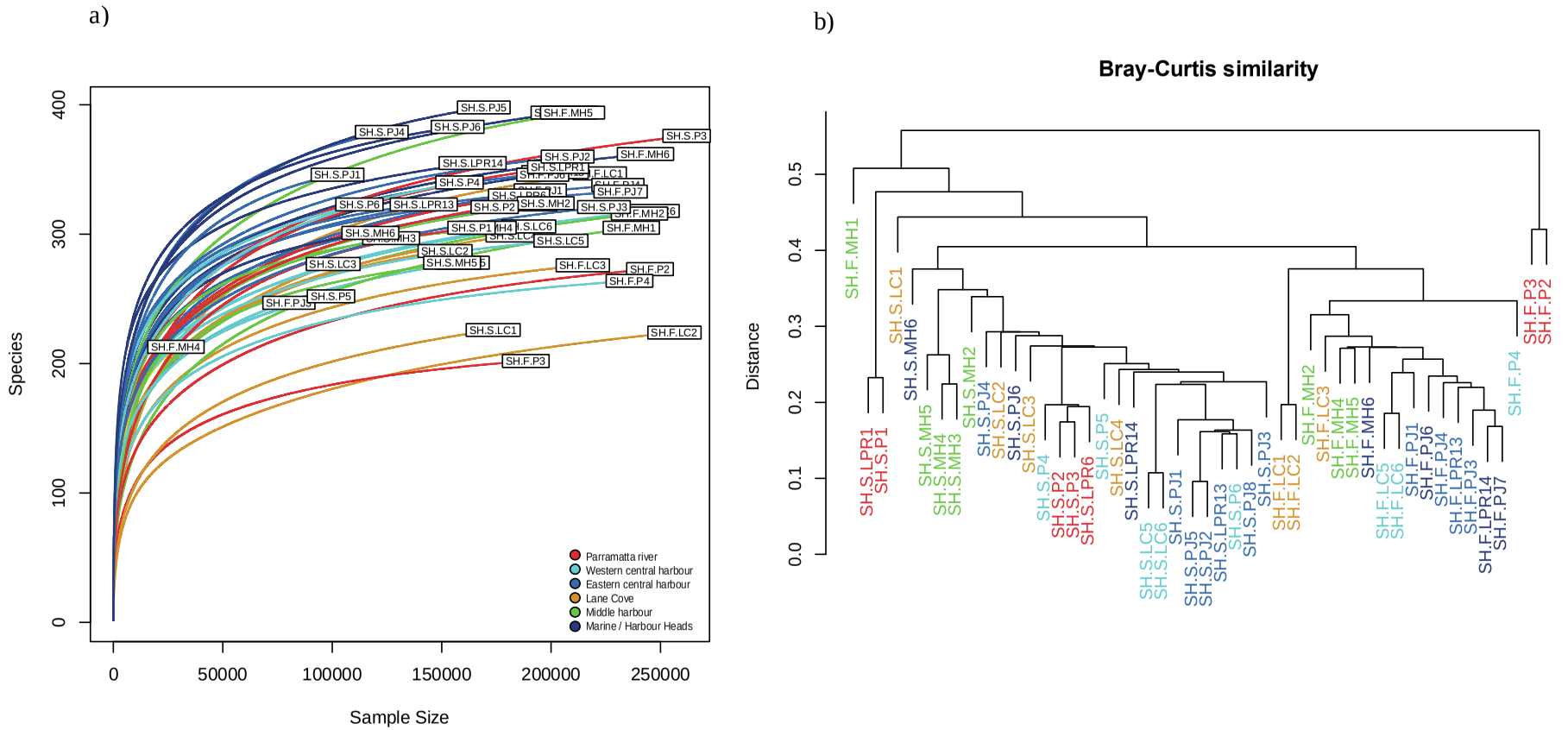


Figure 5.8: Representation of the eukaryotic microbial diversity (18S rDNA) in the Sydney Harbour estuary. a) rarefaction curves and b) Bray-Curtis dissimilarity. Samples were color-coded by regions: Parramatta River (red), Western Central Harbour (light blue), Eastern Central Harbour (blue), Lane Cove (orange), Middle Harbour (green) and Marine/Harbour Heads (dark blue).

The heterotroph MAST (marine Alveolates) known to be groups of parasitic/bacterivorous non photosynthetic alveolates are showing higher proportions of the total microbial eukaryotic community within the central harbour (P2-LPR6, LC3-LC6, up to 7.5% of the population in P5) and also to a lesser extend in some of the eastern sites of the central harbour, however, they are negligible in the inland sites LPR1 and LC1 as well as in all of the Middle Harbour sites. The LC1 site, however, presents high levels of the heterotrophic Ciliophora (23.6%) while there are increasing proportions of Dinophyceae in the eastern sites (up to 24.3% in PJ6 and 47.2% in MH6).

Pseudo-nitzschia, a marine neurotoxin (domoic acid) producing diatom known to form harmful algal blooms (Ajani et al, 2011) was detected in the easternmost sites LPR14 (1.7%) and PJ6 (6.5%) as well as in the middle harbour branch (1.6%-5.4%). *Alexandrium* was detected from eastward from PJ4 and in the Middle Harbour sites, more particularly the *catenella* group (8.3% in PJ4) known to produce STX toxins (Murray et al, 2011). *Alexandrium* was previously reported to have caused harmful algal blooms with outbreaks in the Parramatta River during October-November time (Ajani et al, 2001). These populations present in September could be seeding populations that have the potential to bloom in the harbour during spring due to higher nutrient concentrations.

Interestingly, and in contrast with the 16S rDNA data where cyanobacteria and chloroplast 16S rDNA sequences fell below 1.0% of the total community sequences, in September a larger proportion of the eukaryotic population appeared to be composed of photosynthetic groups compared to the February community. Members of the division Chlorophyta are distributed throughout the harbour: *Ostreococcus* are found at the western end in the Parramatta river, Lane Cove and western central harbour sites; *Bathycoccus* are present at similar sites but more abundant in the central harbour; and the Chlorellales (Trebouxiophyceae) are found mainly in semi-inland sites such as P1, LC1-2, PJ4 and MH2. Representatives of the Cryptophyceae equally show a higher abundance in more central sites of the harbour while they are negligible in inland and the Middle Harbour sites.

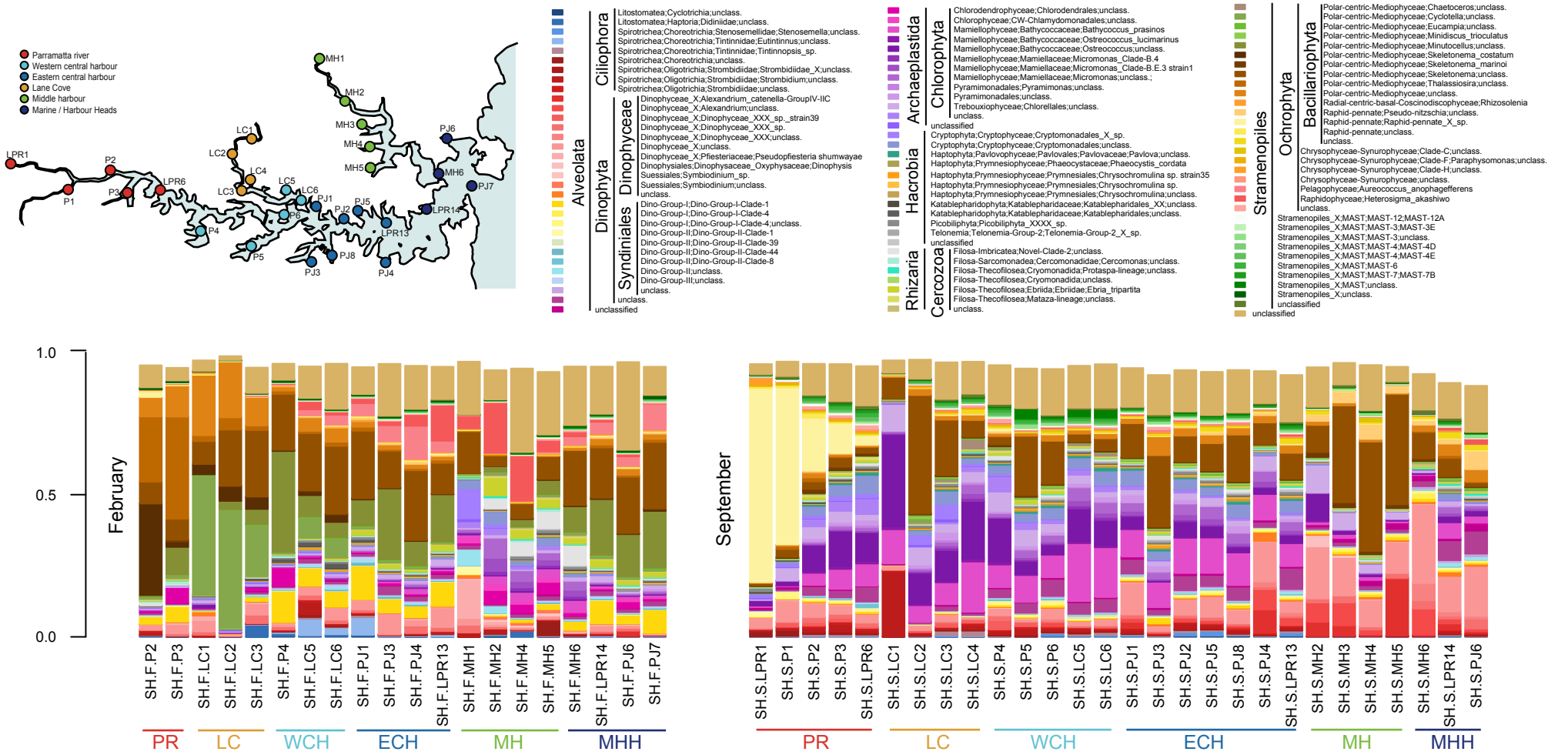

Parramatta river
Western central harbour
Eastern central harbour
Lane Cove
Middle harbour
Marine / Harbour Heads
LPR1
P1
P2
P3
LPR6
P4
P5
P6
LC1
LC2
LC3
LC4
LC5
LC6
PJ1
PJ2
PJ3
PJ4
PJ5
PJ6
PJ7
PJ8
LPR13
LPR14
MH1
MH2
MH3
MH4
MH5
MH6
Alveolata
Ciliophora
Litostomatea;Cyclotrichia;unclass.
Litostomatea;Haptoria;Didiniidae;unclass.
Spirotrichea;Choreotrichia;Stenosemellidae;Stenosemella;unclass.
Spirotrichea;Choreotrichia;Tintinnidae;Eutintinnus;unclass.
Spirotrichea;Choreotrichia;Tintinnidae;Tintinnopsis_sp.
Spirotrichea;Choreotrichia;unclass.
Spirotrichea;Oligotrichia;Strombidiidae;Strombidiidae_X;unclass.
Spirotrichea;Oligotrichia;Strombidiidae;Strombidium;unclass.
Spirotrichea;Oligotrichia;Strombidiidae;unclass.
Dinophyta
Dinophyceae
Dinophyceae_X;Alexandrium_catenella-GroupIV-IIC
Dinophyceae_X;Alexandrium;unclass.
Dinophyceae_X;Dinophyceae_XXX_sp._strain39
Dinophyceae_X;Dinophyceae_XXX_sp.
Dinophyceae_X;Dinophyceae_XXX;unclass.
Dinophyceae_X;unclass.
Dinophyceae_X;Pfiesteriaceae;Pseudopfiesteria shumwayae
Dinophysiales;Dinophysaceae_Oxyphysaceae;Dinophysis
Suessiales;Symbiodinium_sp.
Suessiales;Symbiodinium;unclass.
unclass.
Syndiniales
Dino-Group-I;Dino-Group-I-Clade-1
Dino-Group-I;Dino-Group-I-Clade-4
Dino-Group-I;Dino-Group-I-Clade-4;unclass.
Dino-Group-II;Dino-Group-II-Clade-1
Dino-Group-II;Dino-Group-II-Clade-39
Dino-Group-II;Dino-Group-II-Clade-44
Dino-Group-II;Dino-Group-II-Clade-8
Dino-Group-II;unclass.
Dino-Group-III;unclass.
unclass.
unclassified
Archaeplastida
Chlorophyta
Chlorodendrophyceae;Chlorodendrales;unclass.
Chlorophyceae;CW-Chlamydomonadales;unclass.
Mamiellophyceae;Bathycoccaceae;Bathycoccus_prasinos
Mamiellophyceae;Bathycoccaceae;Ostreococcus_lucimarinus
Mamiellophyceae;Bathycoccaceae;Ostreococcus;unclass.
Mamiellophyceae;Mamiellaceae;Micromonas_Clade-B.4
Mamiellophyceae;Mamiellaceae;Micromonas_Clade-B.E.3 strain1
Mamiellophyceae;Mamiellaceae;Micromonas;unclass.;
Pyramimonadales;Pyramimonas;unclass.
Pyramimonadales;unclass.
Trebouxiophyceae;Chlorellales;unclass.
unclass.
unclassified
Hacrobia
Cryptophyta;Cryptophyceae;Cryptomonadales_X_sp.
Cryptophyta;Cryptophyceae;Cryptomonadales;unclass.
Haptophyta;Pavlovophyceae;Pavlovales;Pavlovaceae;Pavlova;unclass.
Haptophyta;Prymnesiophyceae;Phaeocystaceae;Phaeocystis_cordata
Haptophyta;Prymnesiophyceae;Prymnesiales;Chrysochromulina sp. strain35
Haptophyta;Prymnesiophyceae;Prymnesiales;Chrysochromulina sp.
Haptophyta;Prymnesiophyceae;Prymnesiales;Chrysochromulina;unclass.
Katablepharidophyta;Katablepharidaceae;Katablepharidales_XX;unclass.
Katablepharidophyta;Katablepharidaceae;Katablepharidales;unclass.
Picobiliphyta;Picobiliphyta_XXXX_sp.
Telonemia;Telonemia-Group-2;Telonemia-Group-2_X_sp.
unclassified
Rhizaria
Cercozoa
Filosa-Imbricatea;Novel-Clade-2;unclass.
Filosa-Sarcomonadea;Cercomonadidae;Cercomonas;unclass.
Filosa-Thecofilosea;Cryomonadida;Protaspa-lineage;unclass.
Filosa-Thecofilosea;Cryomonadida;unclass.
Filosa-Thecofilosea;Ebriida;Ebriidae;Ebria_tripartita
Filosa-Thecofilosea;Mataza-lineage;unclass.
unclass.
Stramenopiles
Ochrophyta
Bacillariophyta
Polar-centric-Mediophyceae;Chaetoceros;unclass.
Polar-centric-Mediophyceae;Cyclotella;unclass.
Polar-centric-Mediophyceae;Eucampia;unclass.
Polar-centric-Mediophyceae;Minidiscus_trioculatus
Polar-centric-Mediophyceae;Minutocellus;unclass.
Polar-centric-Mediophyceae;Skeletonema_costatum
Polar-centric-Mediophyceae;Skeletonema_marinoi
Polar-centric-Mediophyceae;Skeletonema;unclass.
Polar-centric-Mediophyceae;Thalassiosira;unclass.
Polar-centric-Mediophyceae;unclass.
Radial-centric-basal-Coscinodiscophyceae;Rhizosolenia
Raphid-pennate;Pseudo-nitzschia;unclass.
Raphid-pennate;Raphid-pennate_X_sp.
Raphid-pennate;unclass.
unclass.
Chrysophyceae-Synurophyceae;Clade-C;unclass.
Chrysophyceae-Synurophyceae;Clade-F;Paraphysomonas;unclass.
Chrysophyceae-Synurophyceae;Clade-H;unclass.
Chrysophyceae-Synurophyceae;unclass.
Pelagophyceae;Aureococcus_anophagefferens
Raphidophyceae;Heterosigma_akashiwo
unclass.
Stramenopiles_X;MAST;MAST-12;MAST-12A
Stramenopiles_X;MAST;MAST-3;MAST-3E
Stramenopiles_X;MAST;MAST-3;unclass.
Stramenopiles_X;MAST;MAST-4;MAST-4D
Stramenopiles_X;MAST;MAST-4;MAST-4E
Stramenopiles_X;MAST;MAST-6
Stramenopiles_X;MAST;MAST-7;MAST-7B
Stramenopiles_X;MAST;unclass.
Stramenopiles_X;unclass.
unclassified
unclass.
1.0
0.5
0.0
February
SH.F.P2
SH.F.P3
SH.F.LC1
SH.F.LC2
SH.F.LC3
SH.F.P4
SH.F.LC5
SH.F.LC6
SH.F.PJ1
SH.F.PJ3
SH.F.PJ4
SH.F.LPR13
SH.F.MH1
SH.F.MH2
SH.F.MH4
SH.F.MH5
SH.F.MH6
SH.F.LPR14
SH.F.PJ6
SH.F.PJ7
PR
LC
WCH
ECH
MH
MHH
September
SH.S.LPR1
SH.S.P1
SH.S.P2
SH.S.P3
SH.S.LPR6
SH.S.LC1
SH.S.LC2
SH.S.LC3
SH.S.LC4
SH.S.P4
SH.S.P5
SH.S.P6
SH.S.LC5
SH.S.LC6
SH.S.PJ1
SH.S.PJ3
SH.S.PJ2
SH.S.PJ5
SH.S.PJ8
SH.S.PJ4
SH.S.LPR13
SH.S.MH2
SH.S.MH3
SH.S.MH4
SH.S.MH5
SH.S.MH6
SH.S.LPR14
SH.S.PJ6
PR
LC
WCH
ECH
MH
MHH

Figure 5.9: Barchart representation of the eukaryotic microbial component in the Sydney Harbour estuary. The community composition was derived from the V9 region of the 18S rDNA (Amaral-Zettler et al, 2009) through amplicon sequencing. Only the major groups with proportion superior to 1% of the total community are shown. As indicated on the map, sampling points were grouped by regions: Parramatta River (red), Western Central Harbour (light blue), Eastern Central Harbour (blue), Lane Cove (orange), Middle Harbour (green) and Marine/Harbour Heads (dark blue).

5.3.3.3 Statistical analyses

Ordination analyses were carried out with the community profiles and environmental factors measured, e.g. physical: temperature, turbidity, Secci depth; chemical: pH, chlorophyll, nutrients such as PO_4 and total (dissolved-) phosphate, various nitrogen compounds such as ammonium, nitrate as well as total (dissolved-) N (Figure 5.10). The two first dimensions of ordination explained 42.8% of the community structure (28.52% and 14.32%, for the first and second dimension respectively). The spatial partitioning of the various sites and phylogenetic groups highlights a strong seasonal separation between the sites as seen in earlier diversity analyses. This seasonal separation is mostly driven by temperature and salinity. These two significant factors (p <0.001) have opposite influences and explain 18.8% and 13.3% of the distribution, respectively.

The other environmental factors available had lesser influence on the eukaryotic composition profiles. The sites from Middle Harbour show the most separation between the seasons, with less nutrient load in February than in September.

The Parramatta River subset is influenced positively with nutrient/turbidity in both seasons for the more inland sites while part of the September samples show similar variation as in the western central harbour. This influence could be due to naval traffic increasing turbidity in the Parramatta River by disturbing the river bed. Whereas LC is strongly influenced by global nutrient loads in February but less in September with possibly more influence of inorganic phosphate. The other subsets of sites appear to have a much weaker correlation with the measured nutrients. This weak correlation gradually decreases from the central harbour to the marine/harbour head.

The Secci depth of the sites correlated with the community composition (8.1%, p <0.001) with near complete opposite influence to turbidity and total suspended solids. Biotic linked factors such as chlorophyll (11.5%, p <0.001) and ODO% (13.2%, p <0.001) influenced the community structure together with the nutrient load in terms of total overall and dissolved N and P: TN (10.7%, p <0.001), TDP (9.4%, p <0.001), TP (9.5%, p <0.001), TDN (7.2%, p <0.005), POP (8.1%, p <0.005). Inorganic

phosphate (PO_4) possibly had a different influence on the community structures than overall loads (6.9%, p <0.005) and might be contributing to the variation observed in the middle harbour sites.

5.3.4 *Conclusions*

It was observed in the 16S rDNA study of the Sydney Harbour estuary (Chapter 4) that the abundance of photosynthetic microbes decreased dramatically in late winter compared to summer (average of all sites for summer to winter for cyanobacteria 14.75% to 0.15%, chloroplasts 10.1% to 0.09%). This suggested a shift from a productive ecosystem in terms of primary production to a more consuming system in winter with higher heterotrophic microbes. The proportion of eukaryotes seen in September might only represent a very minor part of the total microbial communities in Sydney Harbour, hence, carry out a minimal ecosystem service. There is a strong seasonal influence on the community structure partly explained through the changes in temperature and salinity, however, other factors, not measured during this study, strongly drive the community structure within the harbour. Sydney Harbour is known to have strong heavy metal and other toxicant load within its sediment (Hedge et al, 2013). A possible hypothesis can be that winter mixing increases the release of these toxicants (e.g. Pb, Cu, Zn etc.) within the water column. Phototrophic communities are known to be sensitive to some of these metals such as Cu (Mackey et al, 2012; Stuart et al, 2009).

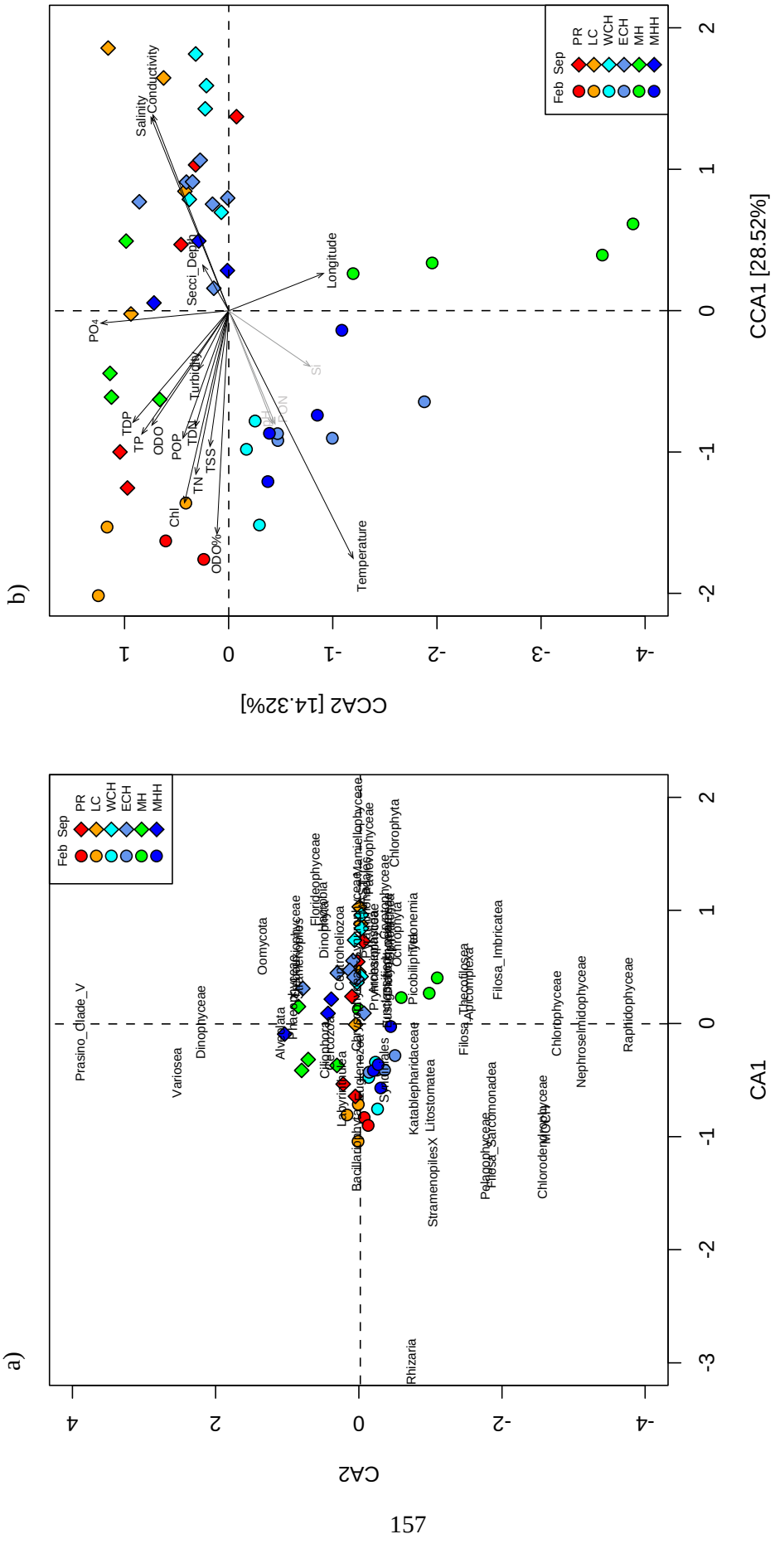

a)
CA2
CA1
Feb Sep
PR
LC
WCH
ECH
MH
MHH
Prasino_Clade_V
Variosea
Dinophyceae
Oomycota
Alveolata
Dinophyta
Florideophyceae
Cercozoa
Ochrophyta
Chlorophyta
Picobiliphyta
Rhizaria
Katablepharidaceae
Litostomatea
StramenopilesX
Filosa_Thecofilosea
Filosa_Imbricatea
Pelagophyceae
Filosa_Sarcomonadea
Chlorodendrophyceae
Mamiellophyceae
Chlorophyceae
Nephroselmidophyceae
Raphidophyceae
b)
CCA2 [14.32%]
CCA1 [28.52%]
Salinity
Conductivity
Secci_Depth
Longitude
PO4
Turbidity
TDP
TP
ODO
POP
TDN
TSS
TN
Chl
ODO%
Si
Temperature

Figure 5.10: 2D representation of correspondence analysis of the sites in the Sydney Harbour estuary according to their eukaryotic microbial community composition (groups >1% of the total community), a) excluding environmental parameters (CA) and b) including environmental parameters (CCA). Sites and eukaryotic component to the level of the class are represented on the two first dimensions CA1 and CA2. The percentage of the total inertia resolved from these dimensions is indicated as percentage next to each axis. Sites coloured according to defined groups: Parramatta River (red), Lane Cove (orange), Western Central Harbour (light blue), Eastern Central Harbour (blue), Middle Harbour (green) and Marine/Harbour Heads (dark blue). Samples in the month of February are denoted by circles and September by diamonds.

5.4 Maria Island

5.4.1 *Introduction*

Maria Island, off the coast of Tasmania is a region of complex oceanography with the influence of multiple ocean currents, the East Australian Current (EAC) from the North, the Leeuwin current (Zeehan current) from the West, and the Antarctic Circumpolar Current from South-West. The region is also a connection zone between major basins of the Pacific and Indian Oceans (Cresswell, 2000). Maria Island has a long running monitoring station (since 1944), National Reference Station (NRS) collecting oceanographic data (Hill et al, 2008). This region has seen increasing temperature conditions attributed to the strengthening of the EAC and warming of the oceans. Shifts in species community and structure have been attributed to these temperature changes (Poloczanska et al, 2007). These shifts have been observed in larger eukaryotic species whereas the microbial community which drive ecosystem function have been less explored. Preliminary metagenomic data from Maria Island (Bioplatforms Australia and Integrated Marine Observing System partnership [Bodrossy, Ostrowski, Brown, Seymour]) shows the changes in picocyanobacterial community over an annual period (2013-2014). The *petB* phylogenetic marker, recruited from metagenomic reads, was used to acquire preliminary insights as to the genotypes present within these waters (data not shown). From December to June, an increased abundance of picocyanobacterial sequences were detected including *Prochlorococcus* ecotypes HLI and HLII as well as *Synechococcus* sub-cluster 5.3 and 5.1 clade II sub-lineages. Clade IV of *Synechococcus* sub-cluster 5.1 was present year round. These genotypes are typical of the EAC and the eddy field further north suggesting an influence of the EAC currents on the local population, and/or significant warming of the water in that area towards the end of Summer and early Autumn.

The RV *Investigator's* voyage INV_E03 (10-17 November 2014) to Maria Island was an excellent opportunity to explore the structure and composition of microbial communities around Maria Island in regions at the tail of the EAC and the East Tasmanian shelf. In addition to equipment testing and

sample collection, one of the objectives of the trial voyage was to train early career researchers. This provided me with the opportunity to be a member of the microbial plankton research team and was provided training in the application of diverse sampling equipment and techniques.

5.4.2 *Research voyage and sampling*

Since the RV *Investigator* is a newly constructed scientific research ship, the trial voyage aimed to test all scientific equipment, laboratories and facilities on board the research vessel, develop safe working procedures for all operations and identify potential issues. As part of the training, samples and data were collected as for a research voyage. Sampling equipment included continuous plankton recorders (CPR), TRIAXUS, Conductivity, Temperature and Depth (CTD) and a variety of nets (EZ, Bongo, Neuston and trawl nets).

Samples were obtained for multiple analyses including flow cytometry, amplicon sequencing, single cell genomics and functional assays including transcriptomics and proteomics. These complementary techniques will help decipher the microbial community structure. Samples were obtained from Niskin bottles of the CTD, deployed at various depths ranging from 0 – 200 m. Water samples (1-8 l) were filtered through 0.2 µm Sterivex filters (Millipore, Australia) for DNA extraction and through CellTraps (MEM-TEQ Ventures, UK) for flow sorting and single cell genomics. Samples collected for analytical flow cytometry were fixed with freshly prepared paraformaldehyde (10% w/v) for a minimum of 1 hour prior to freezing with liquid nitrogen. All samples were stored at -80°C.

5.4.3 *Planned work*

Processing and analyses of these samples could not be done due to time constraints. Future work will involve DNA extraction from the water samples filtered on Sterivex filters and characterisation of the microbial community including picocyanobacteria (*petB*), prokaryotes (16S rDNA) and eukaryotes (18S rDNA) and determine potential interactions. Additional community characterisation will include determining the abundance of different microbial groups using flow cytometry. The integration of

community composition with gene expression analyses including transcriptomics and proteomics will provide valuable insights into community structure and functional characterisation.

Supplementary information is provided in CD.

5.5 References

Abdalla R, Zaghloul F, Hussein N. (1995). A statistical modelling of phytoplankton eutrophication in the Eastern Harbour, Alexandria, Egypt. Bulletin of the National Institute of Oceanography and Fisheries (Egypt) 21: 125-146.

Abell G, Bowman J. (2005). Ecological and biogeographic relationships of class Flavobacteria in the Southern Ocean. FEMS Microbiol Ecol 51: 265-277.

Ajani P, Hallegraeff G, Pritchard T. (2001). Historic overview of algal blooms in marine and estuarine waters of New South Wales, Australia. Proc Linn Soc NSW 123:1-22.

Ajani P, Ingleton T, Pritchard T, Armand L. (2011). Microalgal blooms in the coastal waters of New South Wales, Australia. Proc Linn Soc NSW 133: 15-31.

Amaral-Zettler L, McCliment E, Ducklow H, Huse S. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. PLos One 4: e6372.

Bertrand E, McCrow J, Moustafa A, Zheng H, McQuaid J, Delmont T, *et al*. (2015). Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. Proc Natl Acad Sci USA 112: 9938-9943.

Caron D, Countway P, Jones A, Kim D, Schnetzer A. (2012). Marine protistan diversity. Annu Rev Mar Sci 4:467-493.

Cresswell G. (2000). Currents of the continental shelf and upper slope of Tasmania. Pap Proc R Soc Tasmania 133: 21-30.

Dufresne A, Ostrowski M, Scanlan D, Garczarek L, Mazard S, Palenik B, *et al*. (2008). Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. Genome Biology 9:R90.

Edgar R. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26: 2460-2461.

Guillou L, Viprey M, Chambouvet A, Welsh R, Kirkham A, Massana R, *et al*. (2008). Widespread occurrence and genetic diversity of marine parasitoids belonging to *Syndiniales* (*Alveolata*). Environ Microbiol 10: 3349-3365.

Hedge L, Johnston E, Birch G, Booth D, Creese B, Doblin M, *et al*. (2013). Living harbour. Dynamic Science: A Systematic Review of the Science of Sydney Harbour The Sydney Institute of Marine Science, Sydney, New South Wales, Australia (ISBN: 978-0-646-91693-0, Publisher: The Sydney Institute of Marine Science, Sydney, New South Wales, Australia).

Hill K, Rintoul S, Coleman R, Ridgway K. (2008). Wind forced low frequency variability of the East Australian Current. Geophys Res Lett 35: L08602.

Hoe Chang F, Sutherland J, McVeagh M, Gall M. (2014). Molecular phylogeny, pigment composition, toxicology and life history of *Pseudochattonella* cf. *verruculosa* (Class Dictyochophyceae) from Wellington Harbour, New Zealand. Harmful Algae 34: 42-55.

Hosoi-Tanabe S, Honda D, Fukaya S, Otake I, Inagaki Y, Sako Y. (2007). Proposal of *Pseudochattonella verruculosa* gen. nov., comb. nov. (Dictyochophyceae) for a formar raphidophycean alga *Chattonella verruculosa*, based on 18S rDNA phylogeny and ultrastructural characteristics. Phycol Res 55: 185-192.

Humily F, Farrant G, Marie D, Partensky F, Mazard S, Perennou M, *et al*. (2014). Development of a targeted metagenomic approach to study a genomic region involved in light harvesting in marine *Synechococcus*. FEMS Microbiol Ecol 88: 231-249.

IPCC, 2014: Climate Change 2014: Synthesis report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core writing team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Jardillier L, Zubkov M, Pearman J, Scanlan D. (2010). Significant CO_2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. ISME J 4:1180–1192.

Koh E, Cowie R, Simpson A, O'Toole R, Ryan K. (2012). The origin of cyanobacteria in Antarctic sea ice: marine or freshwater? Environ Microbiol Rep 4: 479-483.

Lane D. (1991). 16S/23S RRNA sequencing. In Stackebrandt E, Goodfellow M (ed.). Nucleic acid techniques in bacterial systematics. Wiley, New York, pp 125-175.

Lane D, Pace B, Olsen G, Stahl D, Sogin M, Pace N. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proc Natl Acad Sci USA 82: 6955-6959.

Laybourn-Parry J. (2002). Survival mechanisms in Antarctic lakes. Phil Trans R Soc Lond B 357: 863–869.

Li W. (1994). Primary productivity of prochlorophytes, cyanobacteria, and eucaryotic ultraphytoplankton: measurements from flow cytometric sorting. Limnol Oceanogr 39:169–175.

Mackey K, Buck K, Casey J, Cid A, Lomas M, Sohrin Y, *et al*. (2012). Phytoplankton responses to atmospheric metal deposition in the coastal and open-ocean Sargasso Sea. Front Microbiol 3: 359.

Magoc T, Salzberg S. (2011). FLASH: Fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27: 2957-63.

Massana R. (2011). Eukaryotic picoplankton in surface oceans. Annu Rev Microbiol 65: 91-110.

Mazard S, Ostrowski M, Partensky F, Scanlan D. (2012). Multi-locus sequence analysis, taxonomic resolution and biogeography of marine *Synechococcus*. Environ Microbiol 14: 372-386.

Morris R, Rappé M, Connon S, Vergin K, Siebol W, Carlson C, *et al*. (2002). SAR11 clade dominates ocean surface bacterioplankton communities. Nature 420: 806-810.

Murray S, Wiese M, Stuken A, Brett S, Kellmann R, Hallegraeff G, *et al*. (2011). *sxtA*-based quantitative molecular assay to identify Saxitoxin-producing harmful algal blooms in marine waters. Appl Environ Microbiol 77: 7050-7057.

Oksanen J, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hare R, *et al*. (2015). vegan: Community Ecology Package. R package version 2.3-0. CRAN.R-project.org/

Park T-G, Bell E, Pearce I, Rublee P, Bolch C, Hallegraef G. (2007). Detection of a novel ecotype of *Pfiesteria piscicida* (Dinophyceae) in an Antarctic saline lake by real-time PCR. Polar Biol 30: 843-848.

Partensky F, Blanchot J, Vaulot D. (1999). Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. Bull Inst Oceanogr 19: 457–475.

Poloczanska E, Babcock R, Butler A, Hobday A, Hoegh-Guldberg O, Kunz T, *et al*. (2007). Climate change and Australian marine life. Oceanogr Mar Biol 45: 407-478.

R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Rijstenbil J. (1987). Phytoplankton composition of stagnant and tidal ecosystems in relation to salinity, nutrients, light and turbulence. Neth J Sea Res 21: 113-123.

Scanlan D. (2012). Marine picocyanobacteria. In: Whitton B (ed.). Ecology of Cyanobacteria II: Their Diversity in Space and Time. Springer Science + Business Media B.V., pp 503-533.

Scanlan D, Ostrowski M, Mazard S, Dufresne A, Garczarek L, Hess W, *et al*. (2009). Ecological genomics of marine picocyanobacteria. Microbiol Mol Biol Rev 73: 249-299.

Schloss P, Westcott S, Ryabin T, Hall J, Hartmann M, Hollister E, *et al*. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75: 7537–7541.

Sherr E, Sherr B. (2002). Significance of predation by protists in aquatic microbial food webs. Anton Leeuw Int J G 81:293-308.

Sohm J, Ahlgren N, Thomson Z, Williams C, Moffett J, Saito M, *et al*. (2015). Co-occurring *Synechococcus* ecotypes occupy four major oceanic regimes defined by temperature, macronutrients and iron. ISME J 1-13.

Stuart R, Dupont C, Johnson A, Paulsen I, Palenik B. (2009). Coastal strains of marine *Synechococcus* species exhibit increased tolerance to copper shock and a distinctive transcriptional response relative to those of open-ocean strains. Appl Environ Microbiol 75(15): 5047-5057.

Wilkins D, Lauro F, Williams T, Demaere M, Brown M, Hoffman J, *et al*. (2013). Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics. Environ Microbiol 15: 1318-1333.

Wilmotte A, Demonceau C, Goffart A, Hecq J, Demoulin V, Crossley A. (2002). Molecular and pigment studies of the picophytoplankton in a region of the Southern Ocean (42–54°S, 141–144°E) in March 1998. Deep-Sea Res II 49: 3351-3363.

Worden A, Not F. (2008). Ecology and diversity of picoeukaryotes. In: Kirchman D (ed.). Microbial Ecology of the Oceans. Wiley, Hoboken, USA, p 594.

Yilmaz P, Parfrey L, Yarza P, Gerken J, Pruesse E, Quast C, *et al*. (2014). The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. Nucl Acids Res 42: D643-D648.

Zwirglmaier K, Jardillier L, Ostrowski M, Mazard S, Garczarek L, Vaulot D, *et al*. (2008). Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. Environ Microbiol 10: 147-161.

Chapter 6:

Impact of DNA damaging agents on genome-wide transcriptional profiles in two marine *Synechococcus* species

frontiers in
MICROBIOLOGY

ORIGINAL RESEARCH ARTICLE
published: 16 August 2013
doi: 10.3389/fmicb.2013.00232

Impact of DNA damaging agents on genome-wide transcriptional profiles in two marine *Synechococcus* species

Sasha G. Tetu[1], Daniel A. Johnson[2], Deepa Varkey[1], Katherine Phillippy[2†], Rhona K. Stuart[3], Chris L. Dupont[3,4], Karl A. Hassan[1], Brian Palenik[3] and Ian T. Paulsen[1]*

[1] *Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, NSW, Australia*
[2] *Microbial and Environmental Genomics Department, J. Craig Venter Institute, Rockville, MD, USA*
[3] *Marine Biology Research Division, Scripps Institution of Oceanography, University of California, San Diego, CA, USA*
[4] *Microbial and Environmental Genomics Department, J. Craig Venter Institute, San Diego, CA, USA*

Edited by:
Martin G. Klotz, University of North Carolina at Charlotte, USA

Reviewed by:
Martin Hagemann, University Rostock, Germany
Donald A. Bryant, The Pennsylvania State University, USA

***Correspondence:**
Ian T. Paulsen, Department of Chemistry and Biomolecular Sciences, Eastern Ave, Macquarie University, Sydney, 2109 NSW, Australia
e-mail: ian.paulsen@mq.edu.au

†Present address:
Katherine Phillippy, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, Bethesda, USA

Marine microorganisms, particularly those residing in coastal areas, may come in contact with any number of chemicals of environmental or xenobiotic origin. The sensitivity and response of marine cyanobacteria to such chemicals is, at present, poorly understood. We have looked at the transcriptional response of well characterized *Synechococcus* open ocean (WH8102) and coastal (CC9311) isolates to two DNA damaging agents, mitomycin C and ethidium bromide, using whole-genome expression microarrays. The coastal strain showed differential regulation of a larger proportion of its genome following "shock" treatment with each agent. Many of the orthologous genes in these strains, including those encoding sensor kinases, showed different transcriptional responses, with the CC9311 genes more likely to show significant changes in both treatments. While the overall response of each strain was considerably different, there were distinct transcriptional responses common to both strains observed for each DNA damaging agent, linked to the mode of action of each chemical. In both CC9311 and WH8102 there was evidence of SOS response induction under mitomycin C treatment, with genes *recA*, *lexA* and *umuC* significantly upregulated in this experiment but not under ethidium bromide treatment. Conversely, ethidium bromide treatment tended to result in upregulation of the DNA-directed RNA polymerase genes, not observed following mitomycin C treatment. Interestingly, a large number of genes residing on putative genomic island regions of each genome also showed significant upregulation under one or both chemical treatments.

Keywords: cyanobacteria, *Synechococcus*, transcriptome, microarray, toxic stress, ethidium bromide, mitomycin C, DNA damage

INTRODUCTION

Ocean environments comprise a vast component of the Earth's biosphere and play a key role in global biogeochemical processes. Unicellular marine cyanobacteria, including *Synechococcus* and *Prochlorococcus* species, are estimated to constitute 20–40% of total marine chlorophyll biomass and carbon fixation, and hence significantly impact the carbon cycle and global climate processes (Partensky et al., 1999). Considering their ecological importance, it is important to understand what stresses these organisms are susceptible to and how they respond to these challenges.

In recent years work has begun to unravel the gene level responses of *Synechococcus* and *Prochlorococcus* to a range of environmental stressors. Nitrogen limitation is thought to be common in the marine environment and these microbes have regulatory systems such as NtcA for regulating their responses to nitrogen (Lindell and Post, 2001). Phosphate limitation stress is also an issue in some marine environments and a subset of cyanobacteria encode the PhoB/R two component regulatory systems for responding to this (Tetu et al., 2009). Another challenge is high irradiance, especially UV wavelengths (Llabres and Agusti, 2006), and adaptation to high light conditions in marine *Synechococcus* involves expression of genes whose products are involved with dissipation of excess light, scavenging to eliminate reactive oxygen species and changes to photosynthetic machinery (Scanlan et al., 2009; Mella-Flores et al., 2012). Marine cyanobacteria have been shown to be more sensitive to copper stress than eukaryotic phytoplankton and their response to this stress has been characterized at the transcriptional level (Stuart et al., 2009). Iron, which is necessary for photosynthetic apparatus, is also considered to be important in the ecology of marine cyanobacteria, and transcriptional responses to changing iron availability have been characterized in *Prochlorococcus* strains (Thompson et al., 2011).

In contrast, relatively few studies to date have focused on how marine cyanobacteria respond to toxic chemicals, despite the finding of multidrug-like efflux systems in marine cyanobacterial genomes (Palenik et al., 2003). Numerous chemicals, including both xenobiotics and compounds produced by other

marine microbes potentially impact these important primary producers, particularly in coastal regions. Toxicity tests on marine *Synechococcus* have indicated that these species are particularly sensitive to exposure to the herbicide atrazine, possibly due to their small cell size and relatively high surface area to volume ratio (Weiner et al., 2004).

Genome sequences are now available for a number of marine cyanobacteria, including *Synechococcus* sp. WH8102, isolated from tropical Atlantic Ocean waters (Palenik et al., 2003), *Synechococcus* sp. CC9311 isolated from coastal waters (Palenik et al., 2006) as well as a number of others from habitats including the Sargasso Sea, Red Sea and the Mediterranean (Palenik et al., 2006; Dufresne et al., 2008). The availability of whole genome sequences allows comparative genomic analyses and techniques such as microarray analysis to be employed to elucidate mechanisms by which cyanobacteria cope with different challenges, such as exposure to toxic compounds.

To gain insight into how different strains of *Synechococcus* respond to toxic chemical exposure, microarray experiments were carried out on an open ocean and a coastal strain using ethidium bromide (EB) and mitomycin C (MC). For each chemical a "shock" treatment (2 h exposure) was applied, as examining short time courses reduces secondary effects due to long term toxicity. These two compounds were chosen as they are both DNA damaging agents which have different modes of action that have been well characterized previously. EB binds to DNA by intercalation between base pairs in the DNA helix, resulting in inhibition of DNA-directed RNA synthesis (Richardson, 1973). MC is a potent DNA cross-linker known to be capable of inducing the SOS response system in many bacterial species (Janion, 2008) and has been used to induce cyanobacterial prophages in environmental samples [for example (Sode et al., 1994)]. The SOS response, which occurs after single stranded DNA accumulates in a cell, is a global response to DNA damage that results in induction of DNA repair and mutagenesis pathways. In well studied bacteria, such as *Escherichia coli*, it is known to result in activation of RecA, which in turn inactivates the LexA repressor, decreasing the pool of this protein which during normal growth negatively regulates a set of genes including *umuD* and *umuC* involved in mutagenic repair (Janion, 2008). While some cyanobacteria such as the well studied freshwater species *Synechocystis* sp. PCC 6803 have lost the classic *E. coli* type LexA mediated SOS response (Domain et al., 2004), recent analyses of marine cyanobacteria *Prochlorococcus marinus* PCC9511 and *Synechococcus* sp. WH7803 have demonstrated a typical, coordinated SOS response to DNA damage, involving *lexA* and *recA* (Kolowrat et al., 2010; Blot et al., 2011).

Whilst neither EB nor MC is currently considered as a significant environmental pollutant and to our knowledge there is at present no information regarding levels of either chemical in the environment, both have been synthesized for more than 50 years, inevitably resulting in some degree of environmental exposure. Mitomycin C has a long history of use as an antitumor drug in the treatment of a wide range of cancers and is marketed in most countries worldwide (Bradner, 2001). In addition to a long history of routine laboratory use as a DNA stain, EB has been widely used in veterinary medicine to treat cattle infected with trypanosomes, a use which continues in parts of the world despite recent concerns regarding its mutagenicity (Roy Chowdhury et al., 2010). EB has also been widely used as a marker for multidrug efflux [for example (Mitchell et al., 1999; Patel et al., 2010; Hassan et al., 2011)] which makes it a particularly good test compound to gain comparative insights into the role of putative efflux pumps encoded within *Synechococcus* genomes in detoxification of the cell.

We hypothesized that a coastal *Synechococcus* strain would have a more robust response to chemical toxicant stress since xenobiotic compounds [likely sources include stormwater, industrial and agricultural runoff (Nogales et al., 2011)] and other toxic chemicals [for example antibiotics produced by marine bacteria (Long et al., 2003)] are likely to be in higher concentrations in coastal environments.

METHODS

Strains and growth conditions

Axenic *Synechococcus* sp. WH8102 and *Synechococcus* sp. CC9311 cultures (referred to as WH8102 and CC9311) were maintained in either SN (Waterbury and Willey, 1988) medium made with seawater from the Scripps Pier (Scripps Institution of Oceanography, La Jolla, CA), or in an artificial seawater medium (SOW) prepared as described previously (Su et al., 2006) with 9.0 mM $NaNO_3$ standard. For growth assays, batch cultures were grown in glass flasks, gently stirred, or in glass tubes, at 25°C under 30 μmol photons m^{-2} s^{-1} continuous white light. Chemical treatments were performed using each compound at a concentration that was inhibitory to long-term growth, 2 μg/mL for EB and 0.5 μg/mL for MC (**Figure A1**). To evaluate cell health at the time which chemical treated cells were harvested, flow cytometry (FACSCalibur, Becton Dickinson) was used to examine cell size and pigmentation. To investigate how these strains respond to these chemical treatments in terms of photosynthetic performance, a saturation pulse (200 ms) was applied to cells using a PHYTO-PAM fluorescence measuring system equipped with an emitter-detector-cuvette assembly unit ED-101US/D (Walz) to determine the quantum yield of photosystem II at multiple time points after the addition of each compound.

RNA isolation

Synechococcus strains were grown in 1.5-liter batch cultures to the early exponential phase (approximately 1.5×10^8 cells/ml), and half the culture was centrifuged at room temperature at 10,400 × g, and immediately resuspended in Trizol reagent (Invitrogen), then frozen at −80°C for RNA extraction in parallel with treated cells. EB (final concentration 2 μg/mL) or MC (final concentration 0.5 μg/mL) were added to the remaining culture. After a two hour incubation, these cells were harvested and suspended in Trizol reagent as outlined above. Total RNA was extracted from the cell pellet using Trizol reagent (Invitrogen) according to the manufacturer's instructions. The RNA was resuspended in 100 μl of DEPC-treated water. Using the RNeasy Mini kit (Qiagen) and following the manufacturer's protocol, the RNA was twice digested with DNase I (Qiagen) then eluted from the columns. The purity and yield of RNA was determined spectrophotometrically by measuring optical density at wavelengths of 260 and 280 nm. Samples were stored at −80°C.

DNA microarray transcriptional profiling

Full genome microarrays were synthesized for both WH8102 and CC9311 as described previously (Tetu et al., 2009). For WH8102 this consisted of a mixed population of PCR amplicons (2142 genes) and 70-mer oligonucleotides (389 genes), while for CC9311 arrays were constructed with 70-mer oligonucleotides for each of 2,892 genes. Each gene was represented six times on an array. Negative controls were 50% DMSO–50% deionized water, and positive controls included Arabidopsis PCR amplicons and 70-mer oligonucleotides. Each microarray experiment included a minimum of two biological replicates and a minimum of three technical replicates for each biological sample, and at least one "dye-swap" experiment per biological replicate.

An indirect labeling method was used to label cDNA as previously described (Peterson et al., 2004), where cDNA is synthesized in the presence of a nucleoside triphosphate analog containing a reactive aminoallyl group to which a fluorescent dye molecule, either cyanine 3 or cyanine 5 (Cy3/Cy5) is coupled. Approximately 4 μg of total RNA was used for indirect labeling, leading to the production of approximately 4 μg of cDNA with approximately 200 pmol of dye molecule incorporated per microgram of cDNA synthesized. Prior to hybridization, labeled cDNA was scanned spectrophotometrically to ensure optimal dye incorporation per sample for adequate signal intensity.

All hybridizations were performed as previously described (Peterson et al., 2004). Processing of the TIFF images from hybridized arrays was performed using TIGR-Spotfinder (www.tigr.org/software), and the datasets normalized by applying the LOWESS algorithm, using block mode and a smooth parameter of 0.33, available in the TIGR-MIDAS package (www.tigr.org/software). Statistical analysis was performed on the mean of $\log_2$-transformed signal ratios of the replicate spots using the Statistical Analysis of Microarrays (SAM) algorithms (Tusher et al., 2001) with a false discovery rate of less than 1%. Subsequent analyses considered genes significantly up- or downregulated based on a cutoff of 0.4-$\log_2$-fold/ −0.4-$\log_2$-fold change. Descriptions of the microarray experiments, quantitation data, and array design have been deposited into the gene expression omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) and have been assigned accession number GSE39818. Genome map circular figures were generated to show the location of the up- and downregulated genes in each *Synechococcus* genome using CGview software (Stothard and Wishart, 2005).

To identify sets of genes which displayed similar patterns of expression when exposed to different test conditions, *k*-means clustering analysis (Soukas et al., 2000) was performed using the TIGR Multi-Experiment Viewer TMEV software (Saeed et al., 2003). Gene trees were subsequently generated by average linkage hierarchical clustering with Euclidean distance as the distance metric.

Cloning and expressing sync2766+7

CC9311 gene sync_2766 (putative RND multidrug efflux transporter) was PCR amplified independently and together with adjacent sync_2767 (putative MFP subunit) using forward primers sync_2766F (5′-CAC CAT GGA ATG CCA GTC AAA ATT CTC-3′) and sync_2767F (5′-CAC CAT GAT TTT GCG GCT TCA ATC G -3′) respectively with reverse primer sync_2766R (5′-GTT TTC AGA AGT ATC TGG CAA AGA GTG-3′). PCR products were generated from CC9311 genomic DNA using Phusion polymerase with HF buffer using the following cycling parameters: 30 cycles (98°C, 10 s; 58°C, 30 s; 72°C, 2 min 20 s) after an initial denaturation at 98°C for 30 s. Products were cloned into Invitrogen Gateway pENTR/SD/D-TOPO vector and shuttled into pET-DEST42 following manufacturers' instructions [and as described previously (Ding et al., 2005)]. Plasmids containing cloned genes were sequenced prior to analysis to confirm the absence of deleterious mutations.

Immunoblots were conducted on whole cell lysates of *E. coli* BL21 cells harboring the pET-DEST42 vector with sync_2766 alone and together with sync_2767 or a control pET-DEST42 vector containing *Pseudomonas aeruginosa* PA01 gene PA0220, a putative amino acid transporter (referred to as the negative control). Cells were grown to mid-exponential phase and gene expression induced with 0.1 mM isopropyl-β-d-thiogalactopyranoside (IPTG) for 1 h at 37°C. Samples were run on a 10% SDS-PAGE gel and transferred to a PVDF membrane. Heterologously expressed V5-tagged proteins were detected using anti-V5 monoclonal antibody conjugated with horseradish peroxidase (HRP) (1/5000 dilution). The membranes were developed with chromogenic substrate 4-Chloro-1-Naphthol (4CN).

Broth-dilution MIC analyses were used to evaluate the resistance potential of the putative RND transporter in *E. coli* BL21(DE3). These analyses were conducted in cells cultured in Mueller-Hinton broth, using standard methodology as described elsewhere (Wiegand et al., 2008). The concentration ranges tested for MC and EB were 0–5 and 0–200 μg/mL respectively.

RESULTS AND DISCUSSION

GLOBAL TRANSCRIPTIONAL RESPONSE TO ETHIDIUM BROMIDE AND MITOMYCIN C EXPOSURE

A total of four sets of global expression microarray experiments were conducted examining the response of coastal strain CC9311 and open-ocean strain WH8102 to exposure to DNA damaging agents EB and MC. In both strains the mRNA levels of a large number of genes was strongly altered by each shock treatment, compared to the control (**Table 1**). *Synechococcus* CC9311 showed a particularly high degree of gene transcription up- and downregulation in response to the tested compounds, with more than 200 genes strongly (>2 fold) up- and downregulated by each treatment. (Lists of all significantly up- and downregulated genes for each strain and stress treatment are presented in **Table S1**). At the time point that gene expression was examined (2 h post addition) chemical treatments had not impacted average cell size or pigmentation as assessed by flow cytometry (data not shown). Measurements of photosynthetic quantum yield for each strain after chemical addition indicated that major reduction in photosynthetic activity occurred after the 2 h time point at which RNA was harvested (**Figure A2**).

TRANSCRIPTIONAL RESPONSES VARIED FOR GENES IN DIFFERENT FUNCTIONAL CATEGORIES

To explore each organism's response to the applied toxic chemical stresses, genes showing significant changes in transcription

levels were grouped according to clusters of orthologous groups (COG) functional categories (**Table 2**). As expected, genes in the "DNA replication, recombination and repair" grouping tended to be transcriptionally upregulated in both strains under both stress conditions. In CC9311 genes in the "posttranslational modification, protein turnover, chaperone" category also tended to be transcriptionally upregulated, whilst in WH8102 these genes were disproportionately downregulated, particularly under conditions of EB stress. For example, both strains have genes encoding molecular chaperone DnaK2 heat shock protein homologs (SYNW2508 and sync_2923). In CC9311 this gene was significantly upregulated under both stress conditions, while the homolog in WH8102 was significantly downregulated in the EB treatment (no significant change in MC treatment). Genes encoding proteases FtsH2 and FtsH3 (sync_0355 and sync_0825/SYNW0305 and SYNW1587) DnaJ protein (sync_0023 and SYNW0024) and co-chaperone GrpE (sync_0022 and SYNW0023) similarly showed transcriptional upregulation in CC9311 and downregulation or no significant change in WH8102.

Interestingly, in CC9311 transcription of "secondary metabolite biosynthesis, transport and catabolism" related genes was affected differently by the two toxins, showing a tendency toward upregulation in EB and downregulation in MC treatment (**Table 2**). A high proportion of genes in the "energy production and metabolism" grouping were transcriptionally downregulated in both strains, although the trend was especially pronounced in WH8102. The "nucleotide transport and metabolism" category genes were also observed to be disproportionately downregulated in both treatments in both strains (**Table 2**).

Table 1 | Numbers of genes in *Synechococcus* sp. WH8102 and CC9311 whose transcription was strongly (more than twofold) upregulated or downregulated under EB and MC shock treatments as detected by microarray analysis.

Strain	EB		MC	
	Upregulated	**Downregulated**	**Upregulated**	**Downregulated**
WH8102	97 (317)	77 (382)	93 (236)	80 (301)
CC9311	255 (786)	228 (781)	211 (715)	226 (745)

Total numbers of significantly affected genes are given in brackets.

Table 2 | Proportion of *Synechococcus* sp. CC9311 and WH8102 genes assigned to each COG category that were significantly transcriptionally upregulated and downregulated by EB or MC treatment.

	COG function	**CC9311**				**WH8102**			
		EB		**MC**		**EB**		**MC**	
		% ↑	**% ↓**	**% ↑**	**% ↓**	**% ↑**	**% ↓**	**% ↑**	**% ↓**
Q	Secondary metabolites biosynthesis, transport and catabolism	36	21	14	36	17	0	13	17
P	Inorganic ion transport and metabolism	23	25	32	26	16	27	10	18
I	Lipid metabolism	34	34	18	28	7	13	9	13
H	Coenzyme metabolism	23	28	20	28	3	16	5	7
F	Nucleotide transport and metabolism	18	42	18	25	0	15	0	15
E	Amino acid transport and metabolism	21	32	23	25	9	25	3	14
G	Carbohydrate transport and metabolism	13	30	19	31	7	20	9	14
O	Posttranslational modification, protein turnover, chaperones	37	18	32	17	10	32	10	9
C	Energy production and conversion	18	32	20	34	10	34	6	20
U	Intracellular trafficking and secretion	25	19	13	6	12	24	12	0
N	Cell motility and secretion	8	15	31	31	13	0	0	13
M	Cell wall/membrane/envelope biogenesis	27	20	27	18	10	14	9	7
T	Signal transduction mechanisms	25	40	22	36	17	14	10	17
D	Cell division and chromosome partitioning	16	47	16	47	14	5	0	0
L	DNA replication, recombination and repair	36	20	34	13	15	7	15	1
K	Transcription	33	25	21	25	15	13	6	13
J	Translation, ribosomal structure and biogenesis	24	21	16	37	9	17	5	19
V	Defense mechanisms	41	34	31	25	10	10	13	3
R	General function prediction only	27	21	31	15	10	10	8	6
S	Function unknown	26	32	26	23	9	12	4	8
	Not in COG	28	27	25	27	18	11	13	14

DNA DAMAGE AGENTS ACTIVATED THE ACCESSORY GENOMES OF BOTH CC9311 AND WH8102

In both strains many genes within regions previously flagged as putative genomic islands (Palenik et al., 2003, 2006) showed significant transcriptional upregulation in either one or both stress treatments (**Figure 1**). Fourteen genome regions in WH8102 have been identified with anomalous G + C content, ten of which were adjacent to putative phage integrases, potentially representing phage associated genomic islands (Palenik et al., 2003). While these regions contain less than 10% of the ORFs in this genome, these putative islands contained a surprisingly large proportion of the highly upregulated genes. In the case of EB 37% of genes upregulated by more than 2 fold (36 of 97 genes) were located in such regions, while for MC stress the proportion was greater still with 47% of such genes located in putative genomic islands (44 of 93 genes upregulated by more than 2 fold). In WH8102 three of the non-phage associated island regions are of particular interest. The region which spans SYNW0424 to SYNW0460 contains a large number of genes which are highly upregulated under MC stress (**Figure 1B**). Many of the affected genes are predicted to function in carbohydrate modification of the cell envelope (COG:M), including numerous genes annotated as putative glycosyltransferases. The gene region SYNW0951 to SYNW0961 also contains a majority of genes strongly upregulated by MC stress [one exception being SYNW0953 encoding a giant 1.2MDa motility-related protein (Strom et al., 2012)] and to a lesser extent EB stress (most show significant upregulation, but falling below the 2 fold cutoff), however, there is almost no functional information for this set of genes. A third island region spanning SYNW2477 to SYNW2491 contains a cluster of genes strongly upregulated by EB (**Figure 1B**). This cluster spans the two previously mentioned putative ABC transporter operons predicted to be involved in transport of zinc and cyanate.

In CC9311 previous genome analyses flagged 19 regions of atypical trinucleotide composition as potential genomic islands (Palenik et al., 2006). Again these regions contained a higher than average proportion of the highly upregulated genes in both stress treatments. While these genomic islands contain roughly 4% of the encoded putative ORFs, more than 10% of genes upregulated by >2 fold in each of the stress treatments were within these genomic regions (28 of 255 in the EB treatment and 26 of 211 in MC treatment). Again, certain island regions were observed to contain clusters of highly upregulated genes. Most notably the large genomic island region spanning sync_0167 to sync_0193, includes a large number of genes showing high levels of upregulation in one or both stress treatments (**Figure 1A**). Similar to the large island in WH8102, this island contains a large proportion of genes associated with modification of the cell envelope (COG:M), including genes involved with capsular polysaccharide biosynthesis. Another putative recently horizontally acquired region showing a concentration of highly upregulated genes is the phage associated gene cluster sync_2414 to sync_2448, which encodes mainly genes of unknown function (**Figure 1A**). Putative RND multidrug efflux transporter encoding genes sync_2766 and sync_2767, which were upregulated by more than 2 fold in both stress treatment, are also located in a putative genomic island region.

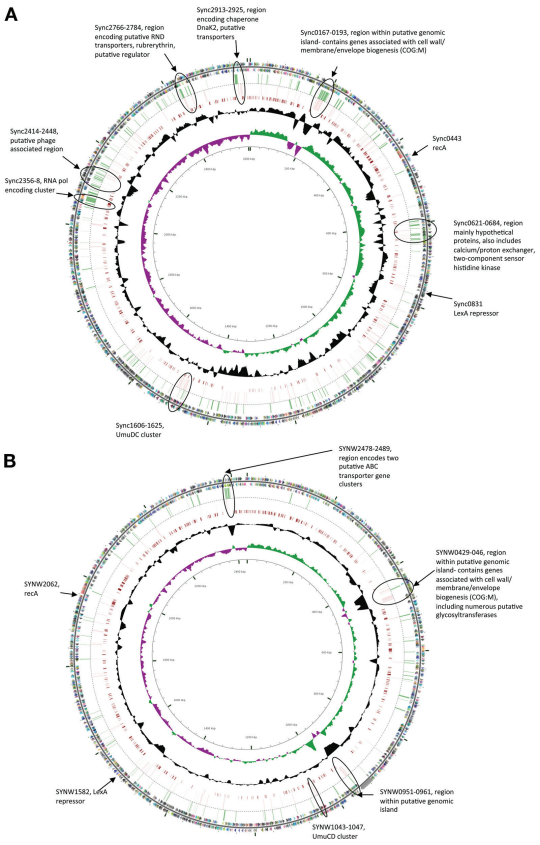


FIGURE 1 | Circular figures showing chromosomes of *Synechococcus* sp. CC9311 (A) and WH8102 (B) indicating location of genes strongly upregulated (>2 fold) by EB and MC treatments. Outermost two circles indicate positions of CDSs in plus (circle 1) and minus (circle 2) strands colored by functional category: translation, ribosomal structure, and biogenesis (pink); transcription (orange); DNA replication, recombination and repair (fuchsia); cell division and chromosome partitioning (yellow); posttranslational modification, protein turnover, chaperones (brown); cell wall/membrane/envelope biogenesis (olive); cell motility and secretion (dark green); inorganic ion transport and metabolism (green); signal transduction mechanisms (blue/green); energy production and conversion (aqua); carbohydrate transport and metabolism (blue); amino acid transport and metabolism (teal); nucleotide transport and metabolism (turquoise); coenzyme metabolism (dark blue); lipid metabolism (lavender); secondary-metabolite biosynthesis, transport, and catabolism (purple); general function prediction only (light gray); function unknown (dark gray); and no COG (gray black). Genes whose expression was upregulated by more than 2 fold in microarray experiments when exposed to EB are indicated by green bars in the third circle, while those similarly upregulated by MC are indicted by pink bars in the fourth circle. Moving toward the center, the next circle map pairwise blastn alignments (expected threshold = 1e-20) between CC9311 and WH8102. Circle six shows G + C content (deviation from average), and the seventh circle illustrates G + C skew in green (+) and purple (−). The scale (in Kbp) is indicated on the innermost circle. Genes of particular interest are circled and described.

Interestingly, other stress treatments in cyanobacterial strains have also resulted in a disproportionally high impact on transcription of genes within genomic island regions. In *Prochlorococcus* MED4 and MIT9313 a large fraction of the

differentially expressed genes in response to iron starvation were observed to reside on genomic islands and/or hypervariable regions (Thompson et al., 2011). Copper shock experiments in *Synechococcus* CC9311 also resulted in upregulation of a large number of genes in genomic island regions (Stuart et al., 2009), and recent examination of particular genes in these island regions has shown genomic island genes can confer adaptive advantage to stress conditions (Stuart et al., 2013).

COMPARISON OF DNA DAMAGE STRESS RESPONSES FOR CC9311 AND WH8102

Many of the orthologous genes shared by these two *Synechococcus* genomes showed different transcriptional responses in the two test strains following each shock treatment. In CC9311 many of the genes observed to be strongly downregulated (>2 fold) in EB stress have orthologs in WH8102 (155/288), however, only 10 of these orthologs were strongly downregulated in WH8102 (**Table A1**). Similarly for the 112 strongly upregulated CC9311 genes which have orthologs, only eight of these also showed greater than 2 fold upregulation in WH8102. The results of the MC shock treatment were similarly divergent between the two strains (**Table A1**). While a smaller number of WH8102 genes with orthologs in CC9311 showed high (>2 fold) up and downregulation, again only a small proportion of the CC9311 orthologs responded in the same way (**Table A1**). In all experiments a small number of orthologous genes responded in the opposite fashion in the two tested strains, for example showing strong upregulation in one strain and strong downregulation in the other, or vice versa (**Table A1**). In previous work looking at the response of these two *Synechococcus* strains to copper shock there was also only a modest overlap in significantly regulated orthologous genes (Stuart et al., 2009).

Previous genome analyses of both strains have shown a distinct difference in the two-component regulatory systems of these strains; WH8102 encodes relatively few histidine kinase sensors, only five compared to CC9311, which has 11 (Palenik et al., 2006). Analysis of the transcriptional response of these sensor kinase genes revealed significant upregulation of six of these genes in CC9311 in at least one of the toxicant exposure conditions, while none of the WH8102 genes were significantly upregulated in either condition (**Table 3**). Sensor kinase sync_0675 in CC9311 showed particularly high levels of upregulation on exposure to both MC and EB, to our knowledge there has been no previous work into the conditions regulating this gene. These observations suggest that the coastal strain CC9311 has different regulatory circuits to WH8102, which may allow it to react to conditions by switching on or off transcription of a greater proportion of "core" *Synechococcus* genes, potentially contributing to the higher tolerance/resilience of CC9311 to the tested toxic compounds.

TRANSCRIPTIONAL RESPONSES ARE LINKED TO THE "MODE OF ACTION" OF CHEMICAL TOXICANTS

In both strains a set of genes showing a significant transcriptional response to both DNA damaging agents was observed. In CC9311 there were 88 genes strongly (more than 2 fold) upregulated and 107 strongly downregulated by both tested compounds. While many of these genes are presently annotated as hypothetical proteins, of those with functional assignations, there were a number encoding regulatory and transport related proteins. These include the two-component response regulator (sync_0115), transcriptional regulator MarR family (sync_2782), two type II alternative RNA polymerase sigma factors (sync_0098, sync_1018) and the sensor histidine kinase (sync_0675) previously discussed. In addition, the genes encoding the protease FtsH2 (sync_0355) and components of a putative RND multidrug efflux transporter (sync_2766 and sync_2767) were transcriptionally upregulated. The strong transcriptional increase observed for the regulatory protein-encoding genes in CC9311 may account for the large proportion of this genome which is affected by these stresses, as these regulators may affect transcription of numerous other genes. WH8102 encodes a homolog only for sync_0115, gene SYNW0126, which itself was not significantly upregulated. In WH8102 there were 22 genes strongly upregulated and 23 strongly downregulated by both tested compounds, almost all of which have not been ascribed a function. Those few with functional predictions include an encoded ribosomal protein

Table 3 | Transcriptional response of sensor kinases to toxic shock treatments in *Synechococcus* sp. CC9311 and WH8102.

CC9311 sensor kinase	MC fold change (log_2)	EB fold change (log_2)	WH8102 homolog	MC fold change (log_2)	EB fold change (log_2)
sync_0675	1.32 (±0.26)	1.4 (±0.37)			
sync_1079	0.36 (±0.22)	0.7 (±0.43)			
sync_0573	−1.32 (±0.25)	−0.95 (±0.40)			
sync_2219	0.46 (±0.19)	NS	SYNW0551	NS	−0.59 (±0.16)
sync_1006	−0.7 (±0.18)	−0.9 (±0.65)	SYNW0753	NS	−0.35 (±0.18)
sync_0706	−0.25 (±0.09)	NS			
sync_1133	NS	0.81 (±0.43)			
sync_0668	0.61 (±0.16)	NS			
sync_0263	−0.5 (±0.23)	−0.92 (±0.44)			
sync_0286	NS	1.37 (±0.40)	SYNW0246	NS	NS
sync_1233	NS	NS	SYNW0807	NS	NS
			SYNW0948*	−0.15 (±0.06)	NS

**SYNW0948 is involved in phosphate sensing and has no CC9311 homolog.*

SYNW2091, ABC transporter component SYNW2479 and CpeT homolog SYNW2003.

While there were overlaps in the transcriptional response between the two shock treatments, there were also clear differences in the response to each chemical, many of which can be accounted for by differences in the mode of action of these compounds. MC is a potent DNA crosslinker and is known to be capable of inducing the SOS response system in many bacterial species (Janion, 2008), while EB binds to DNA by intercalation between base pairs in the DNA helix, resulting in inhibition of DNA-directed RNA synthesis (Richardson, 1973). In both strains there was evidence of SOS response induction under MC treatment. In both CC9311 and WH8102 genes encoding RecA (sync0443/ SYNW2062), the LexA repressor (sync0831/ SYNW1582) and UmuC (sync1607/SYNW1043) were significantly upregulated under MC but not under EB treatment. Another gene, sync1474/SYNW1405, encoding a conserved hypothetical protein present in both examined strains was also highly upregulated under MC treatment but not in the EB experiment. This gene, whose function is presently unknown, appears to be specific to cyanobacteria and has homologs in most sequenced *Synechococcus* genomes. In CC9311 *umuD* encoding sync1606 was also significantly upregulated by MC stress (and to a lesser degree in the EB treatment), while the homolog in WH8102 (SYNW1044) was not. Interestingly, in CC9311 there is a cluster of genes downstream of *umuDC* that were also significantly upregulated in the MC treatment (sync1608-sync1613, sync1615, sync1620-sync1623) (**Figure 1A**). However, as most of these encode proteins of unknown function, it is again difficult to speculate on the exact role of this cluster, which may encode a clade specific SOS response.

It has previously been suggested that cyanobacteria lack an SOS response similar to that of *E. coli* (Domain et al., 2004; Patterson-Fortin et al., 2006) based largely on work in the model freshwater species *Synechocystis* sp. PCC6803, where LexA appears to instead be involved in carbon assimilation. In this strain the *lexA* gene is not DNA damage inducible and its amino acid sequence shows loss of residues required for activity of archetypal LexA proteins, which function to repress expression of DNA repair genes in other prokaryotes (Patterson-Fortin et al., 2006). However, recent experimental and bioinformatic analyses have indicated that it is likely that many of the small, marine cyanobacterial species do indeed have an inducible pathway involved in DNA repair, similar to the *E. coli* SOS response and involving LexA. Li and colleagues examined the available 33 sequenced cyanobacterial genomes and found that, while a small number, including *Synechocystis* sp. PCC6803, did not harbor a recognizable *lexA* gene, the majority do encode a LexA homolog with conserved DNA-binding domains (Li et al., 2010). In their analysis of UV stress response in *Prochlorococcus marinus* PCC9511, Kolowrat and colleagues found that classic, *E. coli*-like DNA repair pathways appeared to operate, noting conservation of active residues in LexA and the presence of a putative LexA binding site upstream of *recA*, *umuC*, *umuD* and *lexA*, and hypothesizing this strain possesses an active SOS response mechanism involving the *lexA* gene (Kolowrat et al., 2010). Similarly, *Synechococcus* sp. WH7803 was recently shown to respond to oxidative stress via up-regulation of a putative LexA regulon, which includes *lexA*, *recA* and *umuDC* (Blot et al., 2011).

In their work Li and colleagues computationally predict *lexA* regulons for CC9311 and WH8102 based on the location of putative LexA boxes within the operator region of a set of genes (Li et al., 2010). Interestingly, only a subset of these genes were found to be significantly upregulated under MC stress in this work, 15 out of 27 for CC9311 and 10 out of 46 for WH8102 (**Table 4**) indicating that further refinement of the LexA boxes using experimental information may be possible.

EB, which acts as an inhibitor of DNA-directed RNA synthesis, led to significant transcriptional upregulation of genes encoding DNA-directed RNA polymerases in both organisms. In CC9311 EB treatment resulted in strong upregulation of the DNA-directed RNA polymerase gene cluster sync_2356-2358 and a number of surrounding genes (**Figure 1A**), none of which were significantly upregulated by MC. In WH8102 two out of three genes in the homologous DNA-directed RNA polymerase gene cluster were significantly upregulated in the EB experiment (SYNW0613 and 0614 but not 0615), but to a lower degree, showing $\log_2$ fold changes 0.76 and 1.0, respectively. Part of this effect in both species is that EB may be generally inhibiting transcription and the cell is responding by additional DNA-directed RNA polymerase production. In WH8102 EB treatment also resulted in upregulation of two adjacent sets of genes encoding putative ABC transporter encoding operons SYNW2479-2481 and SYNW2485-2487, annotated as putative zinc and cyanate transporters respectively. All but one of these genes (SYNW2479) were not significantly upregulated in the MC treatment. Only two of these genes have homologs in CC9311 (sync_1497/SYNW2479 and sync1498/SYNW2480) and in CC9311 these genes were not significantly affected by EB.

DOES sync2766 ENCODE A RND MULTIDRUG EFFLUX TRANSPORTER?

Among the set of CC9311 genes highly upregulated in both DNA damage treatments were sync_2766 and sync_2767 which are annotated as putative RND multidrug efflux transporter and membrane fusion protein (MFP) subunit respectively. Gene sync_2766 contains a conserved domain AcrB (COG0841), indicative of a cation/multidrug efflux pump, while sync_2767 contains the conserved domain RND_mfp (TIGR01730). These genes reside in a region of the genome which was previously flagged as a putative genomic island and protein BLAST searches with the translated amino acid sequences reveal sporadic distribution in marine cyanobacteria (data not shown). We performed a series of expression studies using sync_2766 cloned independently and together with sync_2767 into expression vector pET-DEST42 in *E. coli* BL21(DE3) cells to look for evidence of efflux pump activity. Expression was observed by immunoblotting only when both sync_2766 and sync_2767 were present in the construct. Minimum inhibitory concentration (MIC) assays were carried out with both EB and MC, however, growth of cells expressing sync_2766-7 was not significantly different to the negative control. For EB MICs of 15 μg/mL were routinely observed for both control and experimental constructs, while MICs of 0.09 μg/mL were recorded for MC. In gram-negative

Table 4 | Transcriptional response of *Synechococcus* sp. WH8102 and CC9311 genes previously flagged as having predicted LexA boxes (Li et al., 2010).

WH8102 predicted LuxA box genes	MC fold change (log_2)	EB fold change (log_2)	CC9311 predicted LuxA box genes	MC fold change (log_2)	EB fold change (log_2)
SYNW0037	NS	NS	sync_0035	0.40	NS
SYNW0347	0.32	0.28	sync_0036	2.75	NS
SYNW0450	NS	NS	sync_0129	NS	NS
SYNW0684	1.16	0.76	sync_0443	1.48	0.58
SYNW0711	−0.23	NS	sync_0453	NS	NS
SYNW0712	NS	NS	sync_0748	NS	−0.36
SYNW0713	NS	−0.14	sync_0749	0.34	NS
SYNW0714	NS	NS	sync_0750	0.39	NS
SYNW0715	NS	NS	sync_0751	0.33	NS
SYNW0716	NS	NS	sync_0841	0.88	NS
SYNW0717	NS	NS	sync_0967	0.67	NS
SYNW0718	NS	NS	sync_1344	0.96	0.66
SYNW0719	NS	NS	sync_1373	0.38	1.08
SYNW0720	NS	NS	sync_1473	1.39	−0.84
SYNW0873	NS	NS	sync_1474	2.16	NS
SYNW0874	−0.37	NS	sync_1558	−1.79	−2.26
SYNW0958	0.82	NS	sync_1596	NS	0.93
SYNW0959	NS	NS	sync_1606	4.10	1.15
SYNW1043	1.39	0.26	sync_1607	2.25	0.79
SYNW1044	NS	NS	sync_1623	3.68	1.02
SYNW1045	1.01	NS	sync_2244	−0.60	NS
SYNW1138	1.98	0.32	sync_2245	−0.52	NS
SYNW1140	0.33	0.21	sync_2246	NS	−0.88
SYNW1148	NS	NS	sync_2247	−0.77	−0.69
SYNW1149	NS	NS	sync_2248	−0.60	−1.24
SYNW1203	NS	NS	sync_2443	NS	NS
SYNW1204	NS	NS	sync_2444	−1.34	−2.49
SYNW1205	NS	NS			
SYNW1319	−0.53	NS			
SYNW1466	NS	0.28			
SYNW1467	2.98	1.97			
SYNW1661	NS	0.46			
SYNW2062	3.08	NS			
SYNW2106	NS	NS			
SYNW2107	NS	NS			
SYNW2108	NS	−0.61			
SYNW2109	NS	−0.56			
SYNW2110	NS	NS			
SYNW2111	NS	NS			
SYNW2342	−0.49	0.20			
SYNW2343	NS	NS			
SYNW2344	NS	−0.39			
SYNW2345	NS	−0.75			
SYNW2346	NS	−0.37			
SYNW2347	NS	−0.83			
SYNW2410	0.82	1.58			

Dark grey shading indicates instances where gene transcript levels were not significantly different between the control and test conditions. Light grey shading indicates gene transcript levels observed to be significantly reduced in the test condition compared to the control values.

cells, high-level drug resistance mediated by RND efflux systems, typically relies on the formation of tripartite complexes that include an inner-membrane pump (such as sync_2766), a MFP (such as sync_2767) and an outer-membrane spanning channel, which may be distally encoded in the genome (Nikaido and Takatsuka, 2009). High-level resistance may not have been observed from Sync_2766/7 in *E. coli*, since this organism lacks a cognate outer-membrane protein for this transport system.

Alternatively, it is possible that the high intrinsic resistance of *E. coli* to these compounds (relative to *Synechococcus* sp.), may have swamped any efflux activity attributable to sync_2766.

Synechococcus Spp. DISPLAY UNIQUE RESPONSES TO DIFFERENT STRESS CONDITIONS

Comparing the results of this study to previous microarray experiments on the response of marine *Synechococcus* to nutrient limitation [phosphate starvation (Tetu et al., 2009), (Ostrowski et al., 2010)] heavy metal stress [copper toxicity (Stuart et al., 2009)] and nickel starvation (Dupont et al., 2012), it appears that the transcriptional response to different "types" of stress tend to be quite specific and can differ substantially between strains. In each study the set of genes showing the highest transcriptional increase differed. In many cases this is not surprising as the function of these genes was clearly linked to the nature of the induced stress, for example increasing transcription of alkaline phosphatases under phosphate starvation conditions and transcriptional activation of SOS response genes after MC treatment.

In the analysis of global transcriptional responses of *Synechococcus* strains CC9311 and WH8102 to copper toxicity (Stuart et al., 2009) it was shown that the coastal strain CC9311 responded with a more strain specific oxidative or copper acclimation response than the open-ocean strain WH8102, where a more generic stress response was observed. It was also noted that many of the copper toxicity responsive genes in CC9311 may have been acquired by horizontal gene transfer, however, these genes reside in different island regions to those observed to be responsive to EB and MC. Interestingly, many of the Cu-stress induced WH8102 genes flagged as generic stress response genes, including chaperone proteins DnaJ/DnaK, GroEL/GroES, HtpG and endopeptidase Clp, were observed in this work to be strongly downregulated under conditions of EB toxicity.

One common response observed for both strains following stress with high-copper concentrations and both DNA damaging agents applied in this work was the tendency for photosystem genes to be downregulated. Stuart et al. (2009) reported that of the 37 total photosystem genes, 14 were downregulated in CC9311 while 10 genes were downregulated in WH8102 under high-copper shock conditions, and none of these photosystem genes were upregulated in either strain under these conditions. In this study CC9311 was observed to downregulate 26 photosystem genes under conditions of EB toxicity and 28 when exposed to MC. The open ocean strain WH8102 downregulated 12 photosystem genes when exposed to EB and 24 when exposed to MC. In both cases, however, there were a small number of significantly upregulated photosystem genes (in CC9311 five genes following EB exposure and two genes following MC exposure, in WH8102 three genes following EB exposure). Measurements of photosystem II quantum yield, generated by pulse amplitude modulated fluorometry using the saturation pulse method, indicated that these transcriptional changes occurred before physiologically detectable changes in photosynthesis (**Figure A2**). Genes involved with nitrate transport also showed a tendency to downregulation in both DNA damage stress and copper shock treatments, with putative nitrate transporters SYNW2463-4 and sync_2899 strongly downregulated in both studies.

The large number of datasets concerning transcriptional responses of WH8102 to a range of stress conditions [copper toxicity (Stuart et al., 2009), phosphate starvation (Tetu et al., 2009) and nickel starvation (Dupont et al., 2012)] enabled us to conduct cluster analyses which revealed a cluster of stress related genes upregulated in most treatments (**Figure 2**). This set of upregulated genes encodes functions such as Clp endopeptidase, heat shock proteins, chaperones and cell division proteins. However, as noted above, these were typically downregulated in the EB treatment. This may be due to the possibility that EB directly inhibits transcription (interestingly, late phosphate stress similarly led to downregulation of these genes).

CONCLUSIONS

Synechococcus sp. WH8102 and CC9311 are useful model marine cyanobacteria and can be used to look at how related organisms which live in different environments (open and coastal ocean, respectively) respond to changes such as exposure to toxicants, nutrient deprivation or other stresses. Genome sequences are available for both of these strains, which have provided insight into the genetic basis for adaptations to their different marine environments. Coastal aquatic environments tend to be more dynamic than open ocean environments and, fitting with this, coastal strain CC9311 was previously observed to have genetic mechanisms which may impart a greater capacity to sense and respond to changes in its environment than open ocean strain WH8102 (Palenik et al., 2006). In this work the two tested strains showed a number of differences in their transcriptional response to each of the DNA damage agents tested. Coastal strain

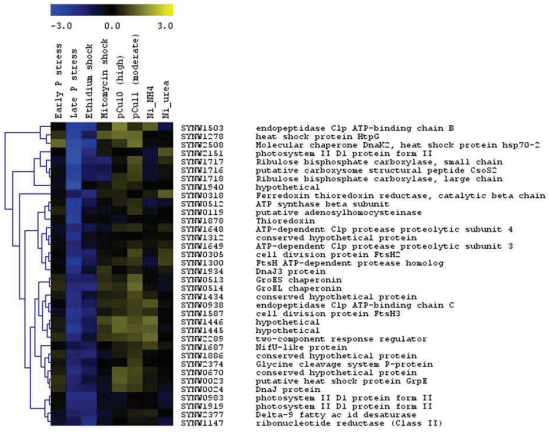


FIGURE 2 | Hierarchical trees showing clusters of *Synechococcus* sp. WH8102 genes whose expression was affected similarly by growth under a large number of stress conditions. The expression levels for this strain under eight different stress conditions: early and late phosphate stress (Tetu et al., 2009), EB and MC shock (this study), moderate and high copper shock (pCu11; pCU10) (Stuart et al., 2009) and Nickel (Ni) deprivation in cultures growing on $NH4^+$ and urea (Dupont et al., 2012) were compared. For each set of experiments data from two independent biological replicates was used. The color bar indicates the $\log_2$ ratio.

CC9311 was observed to alter transcription of a much greater proportion of the genome in response to both toxic compounds tested. Examination of the orthologous genes shared between the strains showed that genes in CC9311 were more likely to illicit a transcriptional response than their WH8102 counterparts. The coastal strain may be able to elicit a more effective global transcriptional response to these compounds through upregulation of two-component response regulators, which could act to coordinate the transcriptional changes observed in this genome.

The data presented here extends our understanding of how model *Synechococcus* strains respond to different stress conditions at the global transcriptional level and helps validate SOS regulatory elements in these strains. Our results indicate that strains residing in different niches, e.g., coastal compared to open ocean environments, may elicit quite different global responses to different encountered stresses. Given the importance of these primary producers to global carbon cycling and marine ecosystem health it is important to continue to extend our understanding of *Synechococcus* stress responses, especially regarding potential anthropogenic pressures.

ACKNOWLEDGMENTS

This work was supported by grants from the US Department of Energy's Genomes to Life program (www.doegenomestolife.org) under project Carbon Sequestration in *Synechococcus* sp.: from Molecular Machines to Hierarchical Modeling, US Department of Energy grant DOE DE-FG03-O1ER63148 to Brian Palenik and Ian T. Paulsen; Australian Research Council Discovery Grant DP110102718: Lifestyle choices: genomic analysis of niche adaptations in marine *Synechococcus* to Ian T. Paulsen; NSF: MCB0744334 to Brian Palenik. We thank Drs Sophie Mazard and Martin Ostrowski for their assistance with operation of the flow cytometer and PAM fluorometer.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Evolutionary_and_Genomic_Microbiology/10.3389/fmicb.2013.00232/abstract

Table S1 | Complete list of *Synechococcus* sp. WH8102 and CC9311 genes whose expression was significantly affected by test conditions, according to SAM analysis. Genes are ordered according to $\log_2$ fold change (shown in numerator column) with those showing the greatest change in expression under test conditions listed first.

REFERENCES

Blot, N., Mella-Flores, D., Six, C., Le Corguille, G., Boutte, C., Peyrat, A., et al. (2011). Light history influences the response of the marine cyanobacterium *Synechococcus* sp. WH7803 to oxidative stress. *Plant Physiol.* 156, 1934–1954. doi: 10.1104/pp.111.174714

Bradner, W. T. (2001). Mitomycin C: a clinical update. *Cancer Treat. Rev.* 27, 35–50. doi: 10.1053/ctrv.2000.0202

Ding, X. Z., Bhattacharjee, A., Nikolich, M. P., Paulsen, I. T., Myers, G., Seshadri, R., et al. (2005). Cloning, expression, and purification of *Brucella suis* outer membrane proteins. *Protein Expr. Purif.* 40, 134–141. doi: 10.1016/j.pep.2004.12.017

Domain, F., Houot, L., Chauvat, F., and Cassier-Chauvat, C. (2004). Function and regulation of the cyanobacterial genes lexA, recA and ruvB: LexA is critical to the survival of cells facing inorganic carbon starvation. *Mol. Microbiol.* 53, 65–80. doi: 10.1111/j.1365-2958.2004.04100.x

Dufresne, A., Ostrowski, M., Scanlan, D. J., Garczarek, L., Mazard, S., Palenik, B. P., et al. (2008). Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol.* 9:R90. doi: 10.1186/gb-2008-9-5-r90

Dupont, C. L., Johnson, D. A., Phillippy, K., Paulsen, I. T., Brahamsha, B., and Palenik, B. (2012). Genetic identification of a high-affinity Ni transporter and the transcriptional response to Ni deprivation in Synechococcus sp. strain WH8102. *Appl. Environ. Microbiol.* 78, 7822–7832. doi: 10.1128/AEM.01739-12

Hassan, K. A., Brzoska, A. J., Wilson, N. L., Eijkelkamp, B. A., Brown, M. H., and Paulsen, I. T. (2011). Roles of DHA2 family transporters in drug resistance and iron homeostasis in *Acinetobacter* spp. *J. Mol. Microbiol. Biotechnol.* 20, 116–124. doi: 10.1159/000325367

Janion, C. (2008). Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. *Int. J. Biol. Sci.* 4, 338–344. doi: 10.7150/ijbs.4.338

Kolowrat, C., Partensky, F., Mella-Flores, D., Le Corguille, G., Boutte, C., Blot, N., et al. (2010). Ultraviolet stress delays chromosome replication in light/dark synchronized cells of the marine cyanobacterium *Prochlorococcus marinus* PCC9511. *BMC Microbiol.* 10:204. doi: 10.1186/1471-2180-10-204

Li, S., Xu, M. L., and Su, Z. C. (2010). Computational analysis of LexA regulons in Cyanobacteria. *BMC Genomics* 11:527. doi: 10.1186/1471-2164-11-527

Lindell, D., and Post, A. F. (2001). Ecological aspects of ntcA gene expression and its use as an indicator of the nitrogen status of marine *Synechococcus* spp. *Appl. Environ. Microbiol.* 67, 3340–3349. doi: 10.1128/AEM.67.8.3340-3349.2001

Llabres, M., and Agusti, S. (2006). Picophytoplankton cell death induced by UV radiation: evidence for oceanic Atlantic communities. *Limnol. Oceanogr.* 51, 21–29. doi: 10.4319/lo.2006.51.1.0021

Long, R. A., Qureshi, A., Faulkner, D. J., and Azam, F. (2003). 2-n-Pentyl-4-quinolinol produced by a marine *Alteromonas* sp. and its potential ecological and biogeochemical roles. *Appl. Environ. Microbiol.* 69, 568–576. doi: 10.1128/AEM.69.1.568-576.2003

Mella-Flores, D., Six, C., Ratin, M., Partensky, F., Boutte, C., Le Corguille, G., et al. (2012). *Prochlorococcus* and *Synechococcus* have evolved different adaptive mechanisms to cope with light and UV stress. *Front. Microbiol.* 3:285. doi: 10.3389/fmicb.2012.00285

Mitchell, B. A., Paulsen, I. T., Brown, M. H., and Skurray, R. A. (1999). Bioenergetics of the staphylococcal multidrug export protein QacA. Identification of distinct binding sites for monovalent and divalent cations. *J. Biol. Chem.* 274, 3541–3548. doi: 10.1074/jbc.274.6.3541

Nikaido, H., and Takatsuka, Y. (2009). Mechanisms of RND multidrug efflux pumps. *Biochim. Biophys. Acta* 1794, 769–781. doi: 10.1016/j.bbapap.2008.10.004

Nogales, B., Lanfranconi, M. P., Pina-Villalonga, J. M., and Bosch, R. (2011). Anthropogenic perturbations in marine microbial communities. *FEMS Microbiol. Rev.* 35, 275–298. doi: 10.1111/j.1574-6976.2010.00248.x

Ostrowski, M., Mazard, S., Tetu, S. G., Phillippy, K., Johnson, A., Palenik, B., et al. (2010). PtrA is required for coordinate regulation of gene expression during phosphate stress in a marine *Synechococcus*. *ISME J.* 4, 908–921. doi: 10.1038/ismej.2010.24

Palenik, B., Brahamsha, B., Larimer, F. W., Land, M., Hauser, L., Chain, P., et al. (2003). The genome of a motile marine *Synechococcus*. *Nature* 424, 1037–1042. doi: 10.1038/nature01943

Palenik, B., Ren, Q., Dupont, C. L., Myers, G. S., Heidelberg, J. F., Badger, J. H., et al. (2006). Genome sequence of *Synechococcus* CC9311: insights into adaptation to a coastal environment. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13555–13559. doi: 10.1073/pnas.0602963103

Partensky, F., Hess, W. R., and Vaulot, D. (1999). *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol. Mol. Biol. Rev.* 63, 106–127.

Patel, D., Kosmidis, C., Seo, S. M., and Kaatz, G. W. (2010). Ethidium bromide MIC screening for enhanced efflux pump gene expression or efflux activity in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 54, 5070–5073. doi: 10.1128/AAC.01058-10

Patterson-Fortin, L. M., Colvin, K. R., and Owttrim, G. W. (2006). A LexA-related protein regulates redox-sensitive expression of the cyanobacterial RNA helicase, crhR. *Nucleic Acids Res.* 34, 3446–3454. doi: 10.1093/nar/gkl426

Peterson, S. N., Sung, C. K., Cline, R., Desai, B. V., Snesrud, E. C., Luo, P., et al. (2004). Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. *Mol. Microbiol.* 51, 1051–1070. doi: 10.1046/j.1365-2958.2003.03907.x

Richardson, J. P. (1973). Mechanism of ethidium bromide inhibition of RNA polymerase. *J. Mol. Biol.* 78, 703–714. doi: 10.1016/0022-2836(73)90290-8

Roy Chowdhury, A., Bakshi, R., Wang, J., Yildirir, G., Liu, B., Pappas-Brown, V., et al. (2010). The killing of African trypanosomes by ethidium bromide. *PLoS Pathog.* 6:e1001226. doi: 10.1371/journal.ppat.1001226

Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., et al. (2003). TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 34, 374–378.

Scanlan, D. J., Ostrowski, M., Mazard, S., Dufresne, A., Garczarek, L., Hess, W. R., et al. (2009). Ecological genomics of marine picocyanobacteria. *Microbiol. Mol. Biol. Rev.* 73, 249–299. doi: 10.1128/MMBR.00035-08

Sode, K., Oozeki, M., Asakawa, K., Burgess, J., and Matusunaga, T. (1994). Isolation of a marine cyanophage infecting the unicellular cyanobacterium *Synechococcus* species NKBG 042902. *J. Mar. Biotechnol.* 1, 189–192.

Soukas, A., Cohen, P., Socci, N. D., and Friedman, J. M. (2000). Leptin-specific patterns of gene expression in white adipose tissue. *Genes Dev.* 14, 963–980.

Stothard, P., and Wishart, D. S. (2005). Circular genome visualization and exploration using CGView. *Bioinformatics* 21, 537–539. doi: 10.1093/bioinformatics/bti054

Strom, S. L., Brahamsha, B., Fredrickson, K. A., Apple, J. K., and Rodriguez, A. G. (2012). A giant cell surface protein in *Synechococcus* WH8102 inhibits feeding by a dinoflagellate predator. *Environ. Microbiol.* 14, 807–816. doi: 10.1111/j.1462-2920.2011.02640.x

Stuart, R. K., Brahamsha, B., Busby, K., and Palenik, B. (2013). Genomic island genes in a coastal marine *Synechococcus* strain confer enhanced tolerance to copper and oxidative stress. *ISME J.* 75, 5047–5057.

Stuart, R. K., Dupont, C. L., Johnson, D. A., Paulsen, I. T., and Palenik, B. (2009). Coastal strains of marine *Synechococcus* species exhibit increased tolerance to copper shock and a distinctive transcriptional response relative to those of open-ocean strains. *Appl. Environ. Microbiol.* 75, 5047–5057. doi: 10.1128/AEM.00271-09

Su, Z., Mao, F., Dam, P., Wu, H., Olman, V., Paulsen, I. T., et al. (2006). Computational inference and experimental validation of the nitrogen assimilation regulatory network in cyanobacterium *Synechococcus* sp. WH (8102). *Nucleic Acids Res.* 34, 1050–1065. doi: 10.1093/nar/gkj496

Tetu, S. G., Brahamsha, B., Johnson, D. A., Tai, V., Phillippy, K., Palenik, B., et al. (2009). Microarray analysis of phosphate regulation in the marine cyanobacterium *Synechococcus* sp. WH8102. *ISME J.* 3, 835–849. doi: 10.1038/ismej.2009.31

Thompson, A. W., Huang, K., Saito, M. A., and Chisholm, S. W. (2011). Transcriptome response of high- and low-light-adapted *Prochlorococcus* strains to changing iron availability. *ISME J.* 5, 1580–1594. doi: 10.1038/ismej.2011.49

Tusher, V. G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U.S.A.* 98, 5116–5121. doi: 10.1073/pnas.091062498

Waterbury, J. B., and Willey, J. M. (1988). Isolation and growth of marine planktonic cyanobacteria. *Meth. Enzymol.* 167, 100–105. doi: 10.1016/0076-6879(88)67009-1

Weiner, J. A., Delorenzo, M. E., and Fulton, M. H. (2004). Relationship between uptake capacity and differential toxicity of the herbicide atrazine in selected microalgal species. *Aquat. Toxicol.* 68, 121–128. doi: 10.1016/j.aquatox.2004.03.004

Wiegand, I., Hilpert, K., and Hancock, R. E. (2008). Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* 3, 163–175. doi: 10.1038/nprot.2007.521

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 April 2013; accepted: 29 July 2013; published online: 16 August 2013.
*Citation: Tetu SG, Johnson DA, Varkey D, Phillippy K, Stuart RK, Dupont CL, Hassan KA, Palenik B and Paulsen IT (2013) Impact of DNA damaging agents on genome-wide transcriptional profiles in two marine Synechococcus species. Front. Microbiol. **4**:232. doi: 10.3389/fmicb.2013.00232*
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology.

APPENDIX

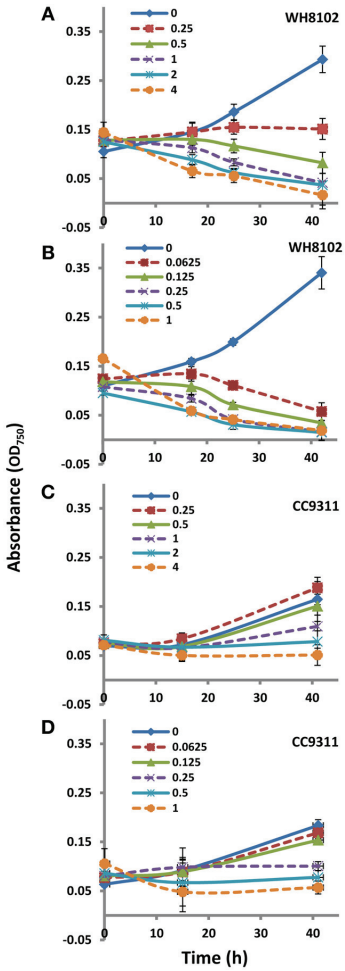


FIGURE A1 | Growth of WH8102 and CC9311 on treatment with EB and MC at concentrations ranging from 0–4 and 0–1 μg/ml^{-1} respectively. **(A)** Growth of WH8102 on treatment with EB; **(B)** Growth of WH8102 on treatment with MC; **(C)** Growth of CC9311 on treatment with EB; **(D)** Growth of CC9311 on treatment with MC. The axes represent absorbance at 750 nm against time in hours. The experiment was conducted in triplicate and the mean plotted. The error bars represent standard deviation.

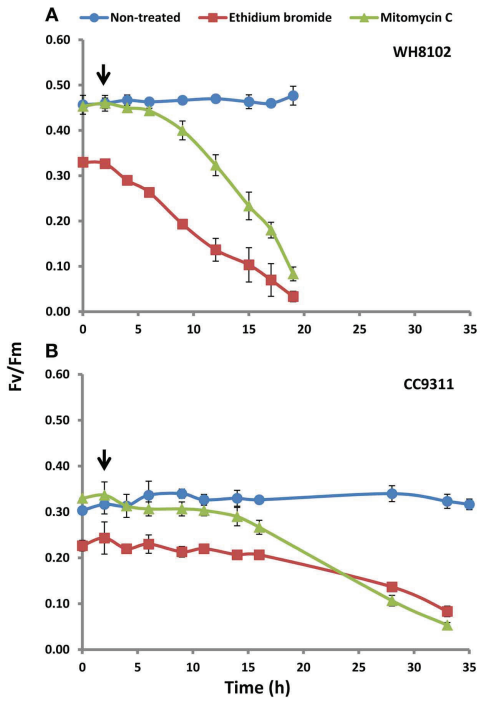


FIGURE A2 | Photosynthetic efficiency of photosystem II in (A) WH8102 and (B) CC9311 on treatment with EB and MC, and with no treatment as determined using PAM fluorometry. The axes represent Fv/Fm (ratio of variable fluorescence and maximum fluorescence) against treatment time in hours. The experiment was carried out in triplicate and the mean plotted. The error bars represent standard deviation. The arrow represents the point of 2 h exposure at which RNA was isolated for microarray analysis.

Table A1 | List of *Synechococcus* sp. CC9311 (A) and WH8102 (B) orthologous genes which showed strong (more than 2-fold) up- or downregulation in both strains in EB or MC treatments.

A

CC9311 gene with predicted function	Log_2 fold change	WH8102 ortholog	Log_2 fold change
	EB treatment		
sync_1804 possible beta-carotene ketolase	2.63	SYNW1368	1.39
sync_1816 predicted membrane protein	2.07	SYNW1452	1.29
sync_1654 conserved hypothetical protein	1.84	SYNW1555	1.67
sync_0955 possible Small, acid-soluble spore proteins, a	1.78	SYNW1906	1.78
sync_1805 conserved hypothetical protein	1.69	SYNW1369	1.06
sync_0622 possible cystathionine gamma-synthase	1.43	SYNW0674	1.69
sync_0950 AbrB family transcriptional regulator	1.22	SYNW1518	1.35
sync_2354 radical SAM enzyme, Cfr family	1.13	SYNW0617	1.12
sync_0355 cell division protein FtsH2	1.44	SYNW0305	−1.30
sync_1975 light-independent protochlorophyllide reductase	−3.03	SYNW1725	−1.03
sync_2092 hypothetical-related protein	−2.43	SYNW1893	−1.32
sync_1569 glutamine synthetase, type I	−1.76	SYNW1073	−1.70
sync_1928 cytochrome c oxidase subunit I	−1.70	SYNW1529	−1.04
sync_2155 Photosystem I reaction center subunit III	−1.56	SYNW1835	−2.15
sync_1973 light-independent protochlorophyllide reductase, N subunit	−1.55	SYNW1723	−1.15
sync_2802 conserved hypothetical	−1.50	SYNW2385	−1.72
sync_2484 Ferredoxin thioredoxin reductase, catalytic beta chain	−1.35	SYNW0318	−1.62
sync_0014 RNA-binding region RNP-1 (RNA recognition motif)	−1.08	SYNW0014	−1.09
sync_2113 NADH-ubiquinone/plastoquinone oxidoreductase, B subunit	−1.01	SYNW1873	−1.71
sync_1409 Photosystem I reaction center subunit psaK	−2.73	SYNW1290	1.24
sync_1581 GUN4-like family protein	−1.51	SYNW1067	1.27
	MC treatment		
sync_0831 LexA repressor	2.51	SYNW1582	1.96
sync_1607 Nucleotidyltransferase/DNA polymerase	2.25	SYNW1043	1.39
sync_1474 conserved hypothetical protein	2.16	SYNW1405	2.49
sync_0443 recA protein	1.48	SYNW2062	3.08
sync_1545 ABC-type Fe3+ transport system periplasmic component	1.05	SYNW1797	−1.66
sync_1409 Photosystem I reaction center subunit psaK	−3.18	SYNW1290	−2.30
sync_0133 photosystem I iron-sulfur center, subunit VII	−2.59	SYNW0144	−1.87
sync_2323 allophycocyanin, alpha subunit	−2.44	SYNW0485	−1.36
sync_0488 phycocyanin, alpha subunit	−2.24	SYNW2023	−1.18
sync_0216 Heme oxygenase	−2.23	SYNW0171	−1.89
sync_1569 glutamine synthetase, type I	−2.18	SYNW1073	−1.89
sync_2092 hypothetical-related protein	−2.06	SYNW1893	−2.64
sync_0463 photosystem I reaction center subunit II	−2.04	SYNW2044	−2.00
sync_2872 urea ABC transporter, periplasmic urea-binding protein	−1.83	SYNW2442	−1.23
sync_2324 allophycocyanin, beta subunit	−1.78	SYNW0484	−1.95
sync_0513 phycobilisome linker polypeptide	−1.77	SYNW1999	−1.51
sync_0489 phycocyanin, beta subunit	−1.76	SYNW2022	−2.06
sync_2907 Type II alt RNA polymerase sigma factor, sigma-70 family	−1.72	SYNW2496	−1.19
sync_0393 photosystem I core protein PsaA	−1.70	SYNW2124	−1.45
sync_0504 C-phycoerythrin class II alpha chain	−1.67	SYNW2009	−1.81
sync_0914 Predicted protein	−1.66	SYNW0691	−1.25
sync_0132 acyl carrier protein	−1.62	SYNW0143	−1.31
sync_0620 cysteine synthase A	−1.56	SYNW0673	−1.12
sync_1137 Photosystem II protein Y-related protein	−1.56	SYNW0898	−1.31
sync_2155 Photosystem I reaction center subunit III	−1.52	SYNW1835	−1.34
sync_0495 C-phycoerythrin class I beta chain	−1.39	SYNW2008	−2.33

(Continued)

Table A1 | Continued

A			
CC9311 gene with predicted function	**Log_2 fold change**	**WH8102 ortholog**	**Log_2 fold change**
	EB treatment		
sync_0511 possible phycobilisome linker polypeptide	−1.38	SYNW2001	−1.44
sync_0562 Photosystem I reaction center subunit IV	−1.37	SYNW1960	−1.23
sync_0398 photosystem I reaction center subunit XI	−1.37	SYNW2118	−1.21
sync_0417 ribosomal protein L36	−1.36	SYNW2087	−1.44
sync_0399 photosystem I subunit VIII-related protein	−1.36	SYNW2117	−1.56
sync_0502 Phycoerythrin class II gamma chain, linker polypeptide (L-R 32.1)	−1.31	SYNW2010	−1.88
sync_2871 high-affinity branched-chain amino acid transport protein livH	−1.27	SYNW2441	−1.76
sync_2154 photosystem I reaction center subunit IX-related protein	−1.26	SYNW1836	−1.20
sync_0516 Possible phycobilisome linker polypeptide	−1.26	SYNW1989	−2.26
sync_2936 Sucrose phosphate synthase	−1.22	SYNW2520	−1.63
sync_1220 magnesium chelatase, H subunit	−1.18	SYNW0820	−1.16
sync_1973 light-independent protochlorophyllide reductase, N subunit	−1.17	SYNW1723	−1.46
sync_0293 ammonium transporter	−1.13	SYNW0253	−2.02
sync_1967 ribulose bisphosphate carboxylase, large subunit	−1.08	SYNW1718	−1.23
sync_0080 possible photosystem II protein PsbZ	−1.06	SYNW0081	−1.20
sync_0487 Protein of unknown function (DUF1001) superfamily	−1.71	SYNW2003	1.01
sync_2330 conserved hypothetical protein	−1.54	SYNW1951	2.06
B			
WH8102 gene with predicted function	**Log_2 fold change**	**CC9311 ortholog**	**Log_2 fold change**
	EB treatment		
SYNW1906 conserved hypothetical protein	1.78	sync_0955	1.78
SYNW0674 possible cystathionine gamma-synthase	1.69	sync_0622	1.43
SYNW1555 hypothetical	1.67	sync_1654	1.84
SYNW1368 possible beta-carotene ketolase	1.39	sync_1804	2.63
SYNW1518 conserved hypothetical protein	1.35	sync_0950	1.22
SYNW1452 predicted membrane protein (COG2259)	1.29	sync_1816	2.07
SYNW0617 conserved hypothetical protein	1.12	sync_2354	1.13
SYNW1369 conserved hypothetical protein	1.06	sync_1805	1.69
SYNW1067 conserved hypothetical protein	1.27	sync_1581	−1.51
SYNW1290 possible photosystem I reaction center subunit X	1.24	sync_1409	−2.73
SYNW1835 Photosystem I reaction center subunit III (PsaF)	−2.15	sync_2155	−1.56
SYNW2385 conserved hypothetical	−1.72	sync_2802	−1.50
SYNW1873 NADH dehydrogenase I chain 2 (or N)	−1.71	sync_2113	−1.01
SYNW1073 Glutamine synthetase, glutamate–ammonia ligase	−1.70	sync_1569	−1.76
SYNW0318 Ferredoxin thioredoxin reductase, beta chain	−1.62	sync_2484	−1.35
SYNW1893 hypothetical	−1.32	sync_2092	−2.43
SYNW1723 light-independent protochlorophyllide reductase	−1.15	sync_1973	−1.55
SYNW0014 RNA-binding region RNP-1	−1.09	sync_0014	−1.08
SYNW1529 cytochrome c oxidase subunit I	−1.04	sync_1928	−1.70
SYNW1725 Protochlorophyllide reductase iron-sulfur ATP-binding protein	−1.03	sync_1975	−3.03
SYNW0305 cell division protein FtsH2	−1.30	sync_0355	1.44
	MC treatment		
SYNW2062 RecA bacterial DNA recombination protein	3.08	sync_0443	1.48
SYNW1405 conserved hypothetical protein	2.49	sync_1474	2.16
SYNW1043 putative UmuC protein	1.39	sync_1607	2.25
SYNW1582 SOS function regulatory protein, LexA repressor	1.96	sync_0831	2.51
SYNW2003 CpeT homolog	1.01	sync_0487	−1.71

(Continued)

Table A1 | Continued

A			
CC9311 gene with predicted function	**Log_2 fold change**	**WH8102 ortholog**	**Log_2 fold change**
	EB treatment		
SYNW1951 conserved hypothetical	2.06	sync_2330	−1.54
SYNW2008 C-phycoerythrin class II beta chain	−2.33	sync_0495	−1.39
SYNW1989 Possible phycobilisome linker polypeptide	−2.26	sync_0516	−1.26
SYNW0253 Ammonium transporter family	−2.02	sync_0293	−1.13
SYNW2010 C-phycoerythrin class II gamma chain, linker	−1.88	sync_0502	−1.31
SYNW2441 putative urea ABC transporter	−1.76	sync_2871	−1.27
SYNW2520 putative sucrose phosphate synthase	−1.63	sync_2936	−1.22
SYNW2117 photosystem I subunit VIII (PsaI)	−1.56	sync_0399	−1.36
SYNW1723 light-independent protochlorophyllide reductase	−1.46	sync_1973	−1.17
SYNW2001 possible phycobilisome linker polypeptide	−1.44	sync_0511	−1.38
SYNW2087 50S ribosomal protein L36	−1.44	sync_0417	−1.36
SYNW1835 Photosystem I reaction center subunit III (PsaF)	−1.34	sync_2155	−1.52
SYNW0143 acyl carrier protein (ACP)	−1.31	sync_0132	−1.62
SYNW0898 possible photosystem II PsbY protein	−1.31	sync_1137	−1.56
SYNW1718 Ribulose bisphosphate carboxylase, large chain	−1.23	sync_1967	−1.08
SYNW1960 photosystem I subunit IV (PsaE)	−1.23	sync_0562	−1.37
SYNW2118 putative photosystem I reaction center subunit XI	−1.21	sync_0398	−1.37
SYNW0081 possible photosystem II protein PsbZ (ycf9)	−1.20	sync_0080	−1.06
SYNW1836 photosystem I reaction center subunit IX (PsaJ)	−1.20	sync_2154	−1.26
SYNW0820 Protoporphyrin IX Magnesium chelatase subunit	−1.16	sync_1220	−1.18
SYNW0673 O-acetylserine (thiol)-lyase A	−1.12	sync_0620	−1.56
SYNW1797 putative iron ABC transporter, substrate binding protein	−1.66	sync_1545	1.05

Genes which responded significantly, but in the opposite direction, are highlighted in gray.

Chapter 7:

Conclusions and future directions

7.1 Conclusions

Oceanic primary production is predicted to be impacted by climate change due to variations in environmental factors including temperature and nutrient availability (Henson et al, 2013). Changes in primary production can in turn have profound effects on biogeochemical cycles. Picocyanobacteria of the genus *Synechococcus* make a significant contribution to oceanic primary production and support marine life.

7.1.1 *Linking temperature preference and biogeography of Synechococcus isolates*

Synechococcus has a global occurrence in a range of temperature regimes with lineages exhibiting distinct temperature niche preference. The studies carried out in chapters 2 and 3 aimed to determine the growth physiology of representative isolates, from the most prominent environmental clades, in response to different temperature conditions.

Growth physiology determined at three environmentally temperatures relevant for the Australian marine ecosystem (18°C, 22°C and 26°C) showed significant differences among representative *Synechococcus* strains of clades I, II, III and IV (Figure 7.1). The growth patterns reflect the distinct environmental distribution of these high latitude (clade I and IV) and mid- to low-latitude (clade II and III) lineages. High latitude clades particularly clade IV displayed the lowest rate of growth at the highest tested temperature. The preference for lower temperatures displayed by clades I and IV correlates with their ecological niche preference of temperatures ranging from 10°C to 20°C. In isolates of clades II and III, the rates of growth correlated positively with temperature. At the highest tested temperature, the clade II isolate continued to increase the rate of growth suggesting the ability to tolerate further increases in temperature. At 18°C, growth was severely affected in isolates of clades II and III. Higher temperature preference observed in isolates of clades II and III support their predominance in warm, open ocean regions with highest relative abundance at temperatures ranging from 20°C – 28°C.

The capacity to sustain higher rates of growth at low temperatures would aid clades I and IV to outcompete clades II and III in high latitude niches. On the other hand, at higher temperatures clades II and III would dominate due to higher rates of growth. These growth patterns reflect the acclimation and predominance of clades I and IV in high latitude cold waters and clades II and III in warmer tropical waters. Such temperature associated growth patterns provide evidence for clade preference for specific temperature regimes and support the hypothesis of temperature playing a significant role in shaping the distribution of *Synechococcus* lineages in the oceanic environment.

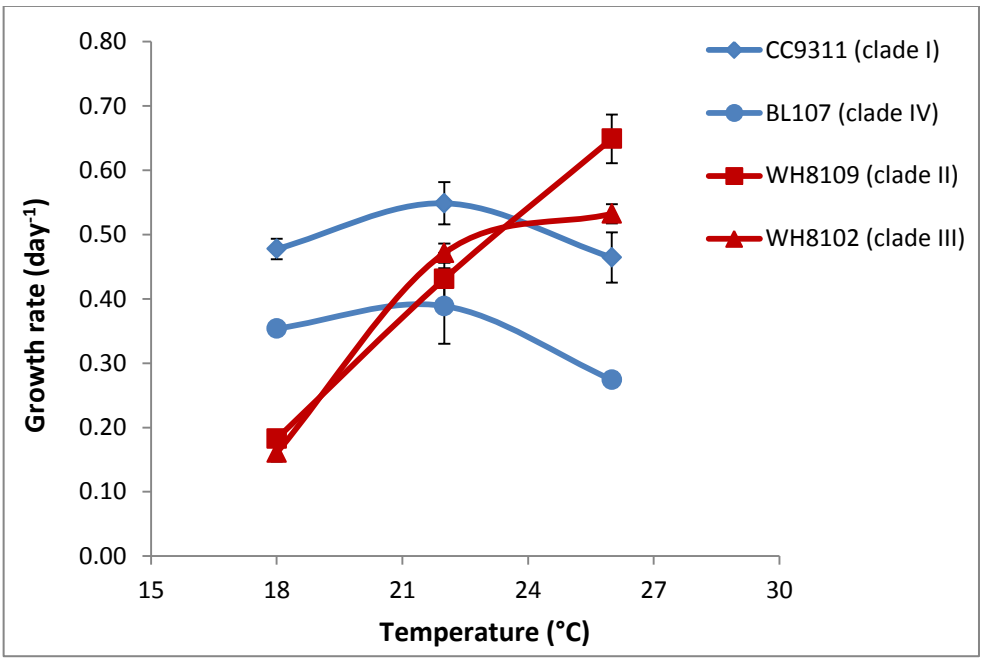


Figure 7.1: Growth rates of marine *Synechococcus* isolates of clades I, II, III and IV grown at 18°C, 22°C and 26°C.

Since individual *Synechococcus* lineages vary in their temperature-related growth physiology, changes in temperature regimes can alter their distribution patterns, which can in turn impact food webs and nutrient cycling. Monitoring shifts in community composition is an important indicator of long-term changes in ecosystem function.

7.1.2 *Temperature acclimation strategies in Synechococcus isolates*

In order to elucidate niche partitioning in marine *Synechococcus* lineages, it is important to uncover the cellular responses to environmental stimuli and the regulatory systems controlling them.

The work undertaken here provides insights to the overall cellular changes in response to temperature including those that are unique to individual isolates as well as common between them. Temperature acclimation was found to involve genes associated with various metabolic pathways, including the adjustment of photosynthetic machinery such as photosystems and phycobilisomes, electron transport components as well as membrane transport. Mechanisms to cope with photo-oxidative stress were also determined to be an important part of temperature acclimation. Though genes associated with most of these components are present in the genomes of most lineages, the different gene expression responses to temperature change observed in this study suggests varying regulatory and functional capacity.

Significant differences in the membrane lipid composition amongst representative isolates were observed. These inherent differences include a significantly lower level of lipid saturation in the high latitude isolate in comparison to lower latitude isolates. Temperature has an influence on membrane structure and acclimation requires adjustment of lipid composition with the involvement of fatty acid desaturases. At the low temperature niches of higher latitudes, lower lipid saturation would contribute to the maintenance of membrane fluidity and function. This adaptation would potentially provide clade IV isolates with a distinct advantage over the clade II and III isolates to thrive in the colder high latitude niches. In response to high temperature, the observed changes in the expression of fatty acid desaturases suggest a requirement for the adjustment of lipid saturation. The observed change was mostly a repression of associated genes, thus suggesting a need for increased level of saturation at high temperature. This study highlights the importance of the organisation and structure of lipid membranes in temperature acclimation and the differential ability of lineages to make the required adjustments. Such differential capacity to adjust lipid composition likely contributes to the observed niche partitioning of lineages.

By exploring the acclimation strategies and changes in resource allocation in different *Synechococcus* lineages, this work begins to enhance our understanding of the genomic basis and functional capacity underlying temperature niche preference. Additionally, whole cellular protein studies provide information on resource allocation to various biological processes and highlight the inherent differences among isolates as seen with the expression of membrane transporters in oligotrophic versus mesotrophic strains. Such information is useful to create metabolic and cellular models.

7.1.3 *Influence of temperature on the Synechococcus community in Australian coastal systems*

This is the first study to examine in detail the picocyanobacterial population structure in an Australian estuarine environment, the Sydney Harbour estuary. Marine *Synechococcus* sub-cluster 5.1 was the primary picocyanobacteria and the community exhibited significant seasonal partitioning in abundance and composition. Temperature was a significant contributor to the change in the *Synechococcus* lineages from the higher prevalence of clade II in late summer to clade I in late winter. The influence of temperature is also observed in the partitioning of lineages along the coast of Australia which is influenced by the East Australian Current. Clade II lineages dominate the communities and clade I and IV increase in abundance further south particularly in regions influenced by the upwelling of cold, nutrient-rich waters. In addition to temperature other co-varying factors would certainly have a role in shaping the community structure. The dramatic seasonal change in the abundance of the *Synechococcus* population in the Sydney Harbour estuary cannot be explained solely by temperature. Physical and chemical parameters such as metals, particularly in coastal areas as well as competition from other microbial groups and reduction of standing stocks through grazing and phage lysis, are all expected to play major roles. The influence of both these abiotic and biotic factors requires further investigation.

7.2 Future directions

7.2.1 *Investigating the role of unique genes in temperature adaptation*

Changes in gene expression in response to temperature conditions suggest a potential involvement in temperature acclimation. Confirming their role in acclimation would require construction of gene knockout mutants. Target genes of particular interest would be those that display differential expression as well as unique occurrence in the genomes of isolates that are acclimated to a specific temperature regime. Though not differentially expressed, the gene (BL107_07284/CC9311_2793) encoding a fatty acid desaturase would be an interesting target due to the low temperature-associated adjustment of membrane saturation. This gene is unique to genomes of clades I and IV which are acclimated to colder temperature regimes. Another potential target, induced at high temperature, is a hypothetical gene (SYNW1579/Syncc8109_1366) which is conserved in the genomes of warm-adapted clades II and III.

Expression changes and gene deletion techniques will help determine the role of these genes in temperature adaptation and define the unique genes that enable lineages adapt to specific environmental niches. Thus this study provides a good basis from which to expand our functional understanding of lineages that inhabit specific niches.

7.2.2 *Exploring membrane saturation as an indicator of temperature niche preference*

Lipid composition influences the organisation, structure and fluidity of the membrane. This in turn can affect functions such as nutrient transport, photosynthetic and respiratory electron transport and state transitions of phycobilisomes between photosystems.

Adjustment in the composition of membrane lipids in response to the prevalent environmental conditions is a strategy that is used by other marine picocyanobacteria (Van Mooy et al, 2006). *Prochlorococcus* dominates oligotrophic subtropical gyres of the ocean wherein the concentration of phosphate is in the low nanomolar range. Among the acclimation strategies utilised by

Prochlorococcus to cope with low phosphorus availability is the biochemical adaptation strategy in which sulfur replaces phosphorus as a major component of lipids. This adaptation strategy supports their successful colonisation of nutrient-scarce regions (Van Mooy et al, 2006).

Since temperature affects membrane fluidity as well as cellular functions directly associated with the membrane, adjustments in lipid composition may be involved in countering these effects. The comparison of lipid composition of a high latitude lineage with that of mid- to low-latitude lineages, showed a distinct difference in the saturation levels. Analysis of a larger set of lineages from different temperature regimes will help determine if differences in membrane composition involving saturated and unsaturated fatty acids can be linked to temperature niches. Such analyses would be useful to test the hypothesis that lower saturation in the membrane lipid content contributes to the better adaptation of *Synechococcus* lineages such as clades I and IV to colder higher latitudes whilst those lineages in the mid- to lower-latitudes contain saturated fatty acids.

7.2.3 *Modeling changes in marine Synechococcus distribution along the Australian coasts*

Predictive models suggest large-scale changes in the global abundance and distributions of picocyanobacteria particularly in higher latitudes (Flombaum et al, 2013). In addition to shifts in the *Synechococus* population as a whole, increases in sea surface temperature may result in changes in the distribution of individual *Synechococcus* lineages and in their community composition. As reported in this study, changes in temperature have a significant effect on the growth physiology of *Synechococcus* isolates with distinct variations between lineages. Thus changes in sea surface temperature will likely impact on the distribution and community composition of these lineages which could in turn lead to major changes in biogeochemical cycling. Alteration in the abundance and distribution of *Synechococcus* lineages in addition to inherent differences in the requirement and utilisation of nutrients among them can in turn impact nutrient cycling. For example, representative isolates of clades II and III increase their biosynthesis of light-harvesting phycobilisomes in response to increase in temperature whilst those of clades I and IV decrease their phycobilisome content. Higher

phycobilisome synthesis would increase the nitrogen requirement and uptake. Due to the higher temperature preference and adaptability of clades II and III, a rise in temperature may result in these clades outcompeting their counterparts of clades I and IV, possibly influencing nitrogen demand.

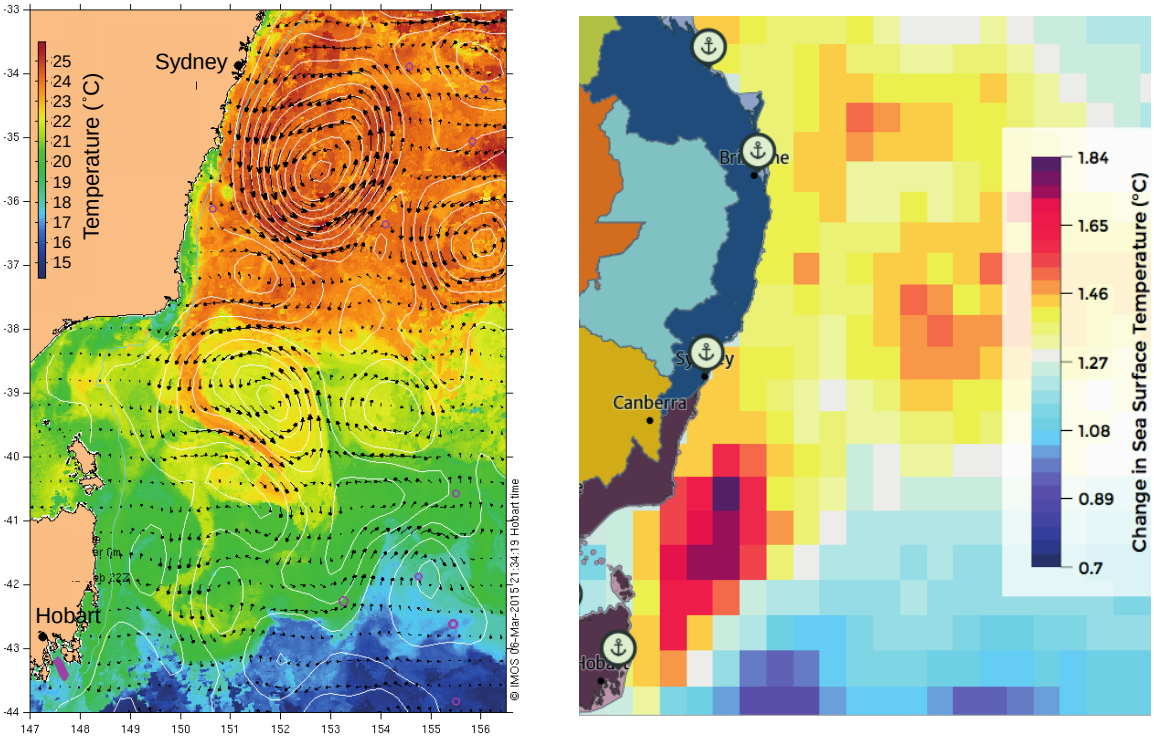


Figure 7.2: Sea surface temperature along the eastern coast of Australia, a) representative of the current conditions of temperature, with the EAC and associated eddy fields (OceanCurrent v1.0 at oceancurrent.imos.org.au), and b) predicted change in temperature in 2070 based on the emissions scenario RCP 4.5 (IPCC, 2013; marine climate projection at www.climatechangeinaustralia.gov.au).

Coastal ecosystems of Australia are strongly influenced by the EAC (Hedge et al, 2013). The intensification of the EAC and its intrusion further south has been linked to climate change associated rising temperatures (Schaeffer et al, 2013; Seymour et al, 2012). These changes could have a significant impact on ecosystems and the services provided as seen with the northward shift of the Gulf Stream and the warming trends in the Gulf of Maine (Pershing et al, 2015). Based on the IPCC emission scenario model RPC 4.5, sea surface temperatures along the eastern coast of Australia are predicted to increase by close to 2°C in some regions (Figure 7.2). The *Synechococcus* populations in

these coastal ecosystems could be significantly impacted. The current niche partitioning of lineages shows the predominance of clade II lineages in the main drift of the EAC and that of clade IV further south around Maria Island, Tasmania and an increased presence of clade I in coastal upwelling regions. Increases in sea surface temperature particularly in the south-eastern regions of Australia and the southward intrusion of the EAC might favor the dominance of clade II over the cold-adapted clades I and IV.

Exploring the distribution of picocyanobacteria in the global marine environment and their adaptability to environmental parameters is fundamental in understanding their evolutionary history and ecosystem function. A better understanding of temperature-related responses including modulation of growth physiology and metabolism as well as partitioning of *Synechococcus* lineages and community structure across different temperature regimes provides a strong foundation to monitor ecosystem health and predict the impacts of environmental change on their abundance, diversity and distribution.

7.3 References

Flombaum P, Gallegos J, Gordillo R, Rincón J, Zabala L, Jiao N, *et al*. (2013). Present and future global distributions of the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. Proc Natl Acad Sci USA 110: 9824-9829.

Hedge L, Johnston E, Birch G, Booth D, Creese B, Doblin M, *et al*. (2013). Living harbour. Dynamic Science: A Systematic Review of the Science of Sydney Harbour The Sydney Institute of Marine Science, Sydney, New South Wales, Australia (ISBN: 978-0-646-91693-0, Publisher: The Sydney Institute of Marine Science, Sydney, New South Wales, Australia).

IPCC, 2013: Climate Change 2013: The physical science basis. Working Group I contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker T, Qin D,

Plattner G, Tignor M, Allen S, Boschung J, *et al.* (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Pershing A, Alexander M, Hernandez C, Kerr L, Le Bris A, Mills K, *et al*. (2015). Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. Science 350: 809-812.

Schaeffer A, Roughan M, Morris B. (2013). Cross-shelf dynamics in a Western Boundary Current regime: Implications for upwelling. J Phys Oceanogr 43: 1042-1059.

Seymour J, Doblin M, Jeffries T, Brown M, Newton K, Ralph P, *et al*. (2012). Contrasting microbial assemblages in adjacent water masses associated with the East Australian Current. Environ Microbiol Rep 4: 548-555.

Van Mooy B, Rocap G, Fredricks H, Evans C, Devol A. (2006). Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments. Proc Natl Acad Sci USA 103: 8607-8612.

Appendix I:

Biosafety approval letter

Final Approval - EXEMPT

Biosafety Application Ref: 5201100898 – Final Approval

Dear Prof. Paulsen

Re: Functional genomic analyses of Escherichia coli, Psedomonas fluorescens complex, Acinetbacter baumannii, Burkholderia cenocepacia, Klebsiella pneumonia, Bacillus cereus and marine cyanobacteria

Thank you for your recent correspondence. Your responses have been reviewed by the Committee and Final Approval of the above application is granted, effective 30 November 2011.

Approval has been granted subject to your compliance with the Office of the Gene Technology Regulator's standard conditions for exempt work listed below:

1. The project must be conducted in accordance with the OGTR Guidance Notes for the Containment of Exempt Dealings (http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/exemptdealclass-2).
2. You must inform the Institutional Biosafety Committee if you complete or abandon the exempt dealings with GMOs.

The following personnel are authorised to conduct this research:

Prof. Ian Paulsen – Chief Investigator/Supervisor

Dr. Karl Hassan – Co-Investigator
Dr. Sasha Tetu – Co-Investigator
Dr. Sophie Mazard – Co-Investigator
Dr. Martin Ostrowski – Co-Investigator
Dr. Anahit Penesyan – Co-Investigator
Dr. Amanda Mackie – Co-Investigator
Dr. Amy Cain – Associate Investigator
Dr. Ole Andreas Okstad - Associate Investigator
Mr. Chee Kent Lim - Associate Investigator
Mr. Prasanth Subramani - Associate Investigator
Miss Bhumika Shah - Associate Investigator
Mr Daniel Farrugia - Associate Investigator
Miss Katy Virginia Breakwell - Associate Investigator
Miss Liping Li - Associate Investigator
Miss Deepa Varkey - Associate Investigator
Ali Khameneh - Associate Investigator
Mr Jan-Christoph Rieckmann – Research Assistant

Please note the following standard requirements of approval:

1. Approval will be for a period of 3 years subject to the provision of annual reports. If, at the end of this period the project has been completed, abandoned, discontinued or not commenced for any reason, you are required to submit a Final Report. If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. These reports are located at the following address:
http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/biosafety_research_ethics/forms

A Progress/Final Report for this project will be due on: 30 November 2012

2. Please remember to notify the Committee of any alteration to the project by completing a 'Request for Amendment' form and submitting it to Biosafety@mq.edu.au The 'Request for Amendment' form is located at the following address:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/biosafety_research_ethics/forms

3. If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide the Macquarie University's Research Grants Management Assistant with a copy of this email as soon as possible. Internal and External funding agencies will not be informed that you have final approval for your project and funds will not be released until the Research Grants Management Assistant has received a copy of this email.

If you need to provide a hard copy letter of Final Approval to an external organisation as evidence that you have Final Approval, please do not hesitate to contact the Committee Secretary at the address below.

Please retain a copy of this email as this is your formal notification of final Biosafety approval.

Yours Sincerely

Dr Subra Vemulpad

Chair, Macquarie University Institutional Biosafety Committee