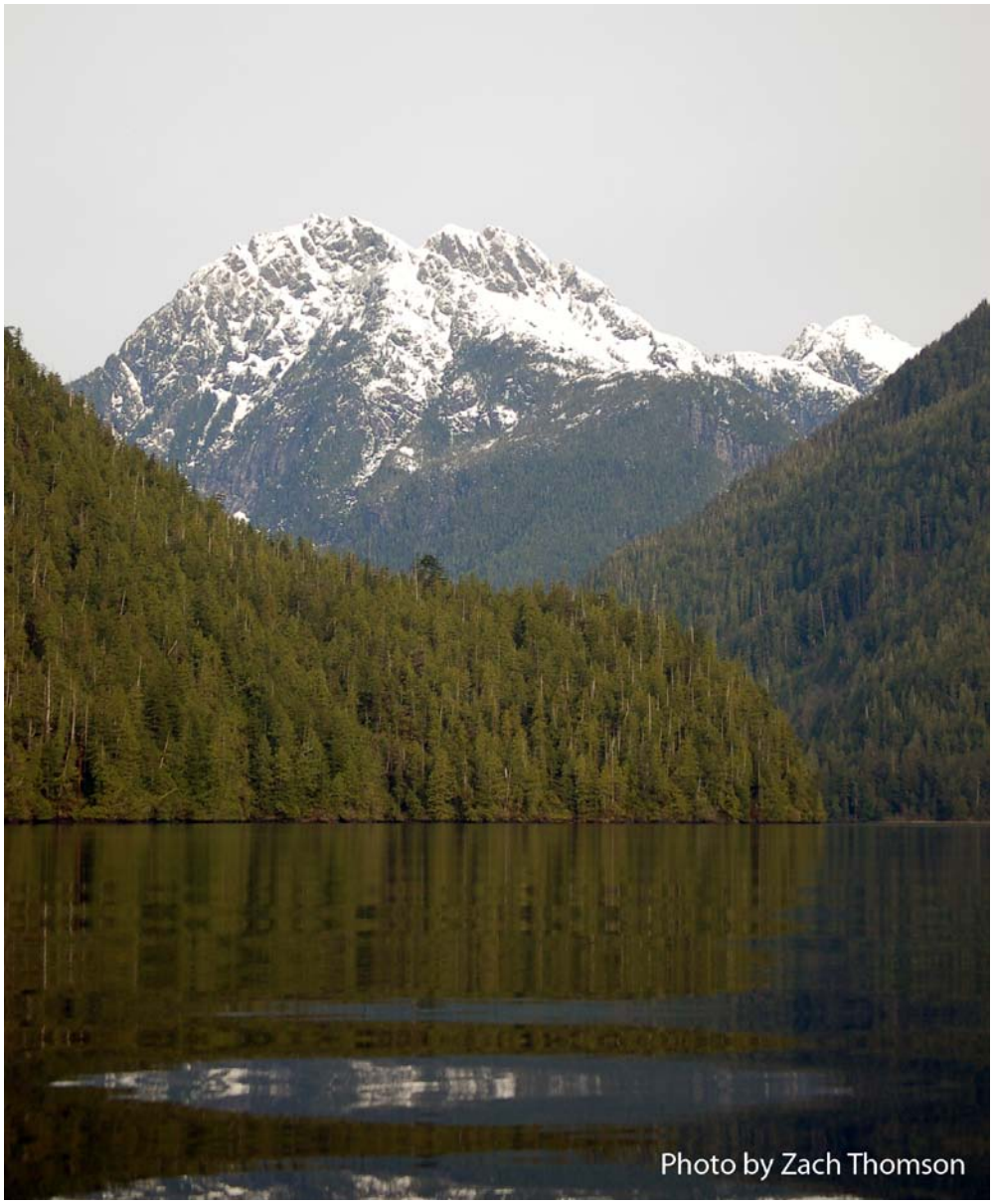
Synechococcus diversity and ecology in open ocean and coastal environments: a comparison



Zachary Thomson

School of Oceanography, Box 357940
University of Washington
Seattle, WA 98195-7940
Phone: (775) 240-6164

Non-Technical Summary

Bacteria are universally diverse and abundant organisms and this is true in the oceans as well. Marine cyanobacteria, or photosynthetic bacteria, are responsible for over 20% of the world's atmospheric oxygen that we breathe. Their abundance in the oceans makes them a very unique organism to study. By taking water samples from various regions around the world, we can determine why certain types of bacteria live in certain parts of the oceans and why they can't survive in others. These organisms are also very genetically diverse. Their sub-microscopic size makes new technological advances in molecular biology crucial to the study of these organisms. By analyzing the distribution and genetic sequences of these bacteria, we can determine the differences between organisms that look exactly alike. This project focuses on analyzing two specific types of bacteria (clades I and IV) within the genus *Synechococcus*. Both of these clades were found to be highly abundant in nutrient rich waters and were absent in nutrient poor waters. The ratio between these organisms also shifted from an IV dominated community in the open ocean to a much more even representation in the coastal samples which may be linked to iron limitation in the open ocean. However, the abundance of these organisms throughout both research sites demonstrates their importance to global carbon cycling and the ecosystem in general.

Acknowledgments

I would like to thank the University of Washington's School of Oceanography for funding the research and making everything possible. There are also a few individuals that I would like for not only helping shape my project but giving me the chance to gain invaluable experience as an undergraduate. I would like to thank Gabrielle Rocap for providing guidance, critique, and a great work environment. With her help, I feel that I have received the best experience possible as an undergraduate and am ready for new endeavors. I would also like to thank all the other graduate students and workers in the CEG lab for helping me whenever I needed it. Thank you to all the professors and Colleen Durkin for putting in all the time and effort to make this possible, even when things looked painfully dark. Lastly I would like to thank my Dad for all of his support.

The unicellular marine cyanobacteria *Synechococcus* is a dominant phototroph that accounts for a significant portion of primary production in the world's oceans. They typically range from 0.8-1.5 µm in size and have a cosmopolitan distribution that is not yet completely understood. *Synechococcus* clades I and IV have been generalized as coastal dwelling organisms. While the amount of phylogenetic data on these organisms is increasing dramatically, the ecology of *Synechococcus* is more complex and unresolved. Using qPCR and clone libraries, open ocean and coastal dwelling cells of clades I and IV were analyzed. Time constraints restricted any clone data but qPCR found *Synechococcus* concentrations over 10^5 cells mL^{-1} in the sub-arctic North Pacific. The Barkley Sound estuary contained quantities ranging from $3.09x10^3$ to $3.25x10^4$ cells mL^{-1}. Clade I and IV's abundance in the high nutrient, low chlorophyll waters of the North Pacific was attributed to cell size which yields a very high surface area to volume ratio. These cells' capacity to compete with larger diatoms in the high nutrient coastal waters as well was attributed to clade I and IV's ability to regulate the uptake of phosphate and metals, a unique adaptation to these clades. The ratio of these clades did vary based on geographic region, despite high concentrations of both clades being present. The sub-arctic North Pacific yielded average clade I:IV ratios of .29:1 and Barkley Sound ratios were closer to 1:1. While no relationship between measured physical or chemical factors were linked to *Synechococcus* distribution, iron may play a role in the ratio of these two clades but further investigation is needed.

Introduction

Bacteria are universally diverse and abundant organisms. Their role in any food web can be substantial. The unicellular marine cyanobacteria *Synechococcus* is a dominant phototroph that accounts for a significant portion of primary production and carbon cycling in the world's oceans (Whitman et al. 1998). The incredibly small size of these organisms, typically 0.8-1.5 µm in diameter, makes them one of the smallest oxygen evolving organisms on the planet. Due to their size and generalized morphology, molecular techniques have become the most efficient way of studying *Synechococcus*. The clades, or branches of a phylogenetic tree, of *Synechococcus* are mainly derived by the differences in the internal transcribed spacer (ITS1) region, which is located between the 16S-23S ribosomal deoxyribonucleic acid (rDNA). This region is highly conserved within *Synechococcus* as a genus, yet has enough variation to allow specific analysis of individual clades. New molecular data using sequenced DNA segments is defining differences in the physiologies of these organisms in an attempt to determine their ecology and distribution. Quantitative polymerase chain reaction (qPCR) is a high-throughput, sensitive quantitative detection method that can target specified DNA segments. This technique allows for the analysis of cyanobacterial clade dynamics on both temporal and spatial scales (Ahlgren 2005). Clone libraries can provide valuable insight into the genetic variation of these organisms; however this technique is more restricted in a spatial sense. The variation in the ITS sequences between clades and within clades could help determine the physiological differences that are still unknown.

While *Synechococcus* is found nearly everywhere in the world's oceans, the water properties that factor into the distribution of these different clades is not yet thoroughly understood. Recent statistical research has shown that chemical factors account for 22.4% of the variance in *Synechococcus* distribution, compared to only 7.4% from physical parameters; however 63.4% of the variance is still unexplained (Zwirglmaier et al 2008). Clades also differ in the nitrogen sources that are utilized. While most of the clades can utilize urea, ammonia, nitrate and nitrite, the rate of assimilation can differ

(Moore et al. 2002). The relationship between nitrogen types and concentrations on the distribution of *Synechococcus* has not been completely determined (Ahlgren et al 2006).

Clades I and IV are some of the most abundant *Synechococcus* clades overall. They are generalized as having coastal distributions, yet can still be abundant in temperate open ocean waters. Recent genomic research has determined that both clades have adaptations that control nutrient uptake and regulation, specifically with phosphate and metals, which may give them an advantage in coastal environments or environments with higher nutrients (Zwirglmaier et al 2008). Clade I has been found to possess the ability to change the ratio of its photoharvesting pigments, also known as chromatic adaptation (Ting et al. 2002, Ahlgren et al. 2006). Whether these two clades compete for the same niche or if there are mutualistic benefits between the two is still being determined. Most studies have shown that clade IV typically outnumbers clade I in abundance throughout most of the year (Tai et al. 2009). However, recent research in the Puget Sound has shown a completely reversed ratio of these two clades, and so far this is the only research to date that has presented such a unique ratio (Rocap unpublished). Even when the ratio is 'normal', the factors contributing to this abundance is unknown (Zwirglmaier et al. 2008).

This study seeks to utilize both molecular techniques to analyze *Synechococcus* clades I and IV from a broad spatial scale down to individual sequences of clones to determine ecological and physiological differences between coastal and open ocean organisms. While the phylogenetic trees have been determined for *Synechococcus* clades, the genetic differences within clades from different regions have not been investigated. Clades I and IV have been found in both open ocean and coastal regions yet these waters can vary drastically in both physical and chemical properties. Thus, this paper seeks to compare how genetically similar cells within the same clade are when they are found in different environmental regimes. The genetic differences within clades abundant in both regions will be obtained using DNA from environmental samples. The direct comparison of the 16S-23S ITS1 regions of rDNA

from clones will provide insight into both the physiology of these organisms, as well as snapshot of the genetic diversity that can be found within these 'genetically distinct populations'.

Materials and Methods

Environmental Collection

The environmental samples were collected on two different research cruises. The first collection sites were in the Northeastern Pacific on the R/V Thomas G. Thompson from August 31st-September 17th 2008 (Fig. 1). The coastal samples were taken from the R/V Barkley Star in the Barkley Sound from March 24th-25th 2010 (Fig 2). Depth profiles containing eight samples spaced over the euphotic zone were taken in the sub-arctic North Pacific while only surface samples were obtained from the estuarine waters of Barkley Sound. Samples of 100 ml of seawater were taken in triplicate and filtered onto .2 µm polycarbonate filters. Nutrient and CTD data were also collected and analyzed in the Marine Chemistry Department at the University of Washington.

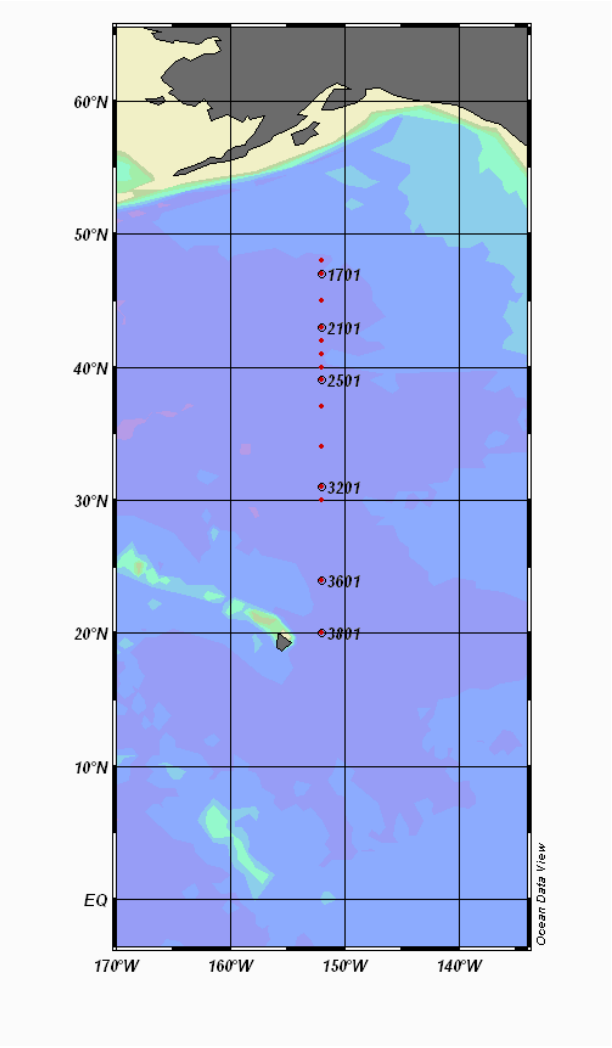


Figure 1. Station map of North Pacific transect. Labeled stations contain full euphotic depth profiles.

DNA Extraction

A quick and 'dirty' DNA extraction was used for these experiments, and lysed the cells in solution. Environmental filters were broken down using 650 µL of TloE and beadbeat for 2 minutes. Then, 500 µL of the resulting mixture was extracted into clean 1.5 mL tubes and heat shocked for 15 minutes at 95^{0}C. The DNA solution was then stored at -80^{0}C until analysis.

Quantitative PCR

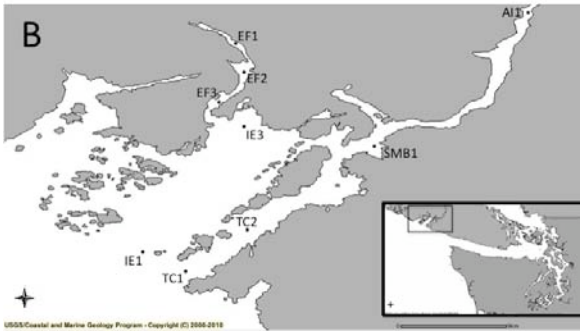


Figure 2. Station map of Barkley Sound

Environmental DNA was used in qPCR using primers and protocols designed by Nathan Ahlgren and Gabrielle Rocap (2005). Clade I was amplified using primers Syn1F2 (CTTTGTCTAGTTCACAACCCATTA) and Syn1R2 (AAATCACAATCCATACGAGTTCAT) with a final Mg^{2+} concentration of 3.5 mM. Clade IV was amplified using primers Syn4F1 (GCAAGAGCCGAGACTCTTAGAT) and Syn4R1b (CTCTTAAACGCTTACTGCGGT) with a final Mg^{2+} concentration of 1.5 mM. Both clades were run with an annealing temperature of 61^0C. Experiments were analyzed in iCycler and visualized in Ocean Data View. All stations and depths were analyzed in the North Pacific. Samples in Barkley Sound were restricted to the surface.

DNA Isolation and Cloning

The ITS1 region was amplified using 16S-1247f (CGTACTACAATGCTAC) and 23S-241r (TTCGCTCGCCRCTACT) primers which were designed by the Rocap Lab to target all clades of *Synechococcus*. Station 21 from the North Pacific and station IE1 from Barkley Sound were chosen to be cloned, due to the highest probability of obtaining sequences from both regions of the same clade (Fig. 1 and 2). Reactions were performed with 2 µl each of Taq buffer (10x), 8mM dNTPs, and primers (forward and reverse), along with .2 µl of Taq, 1.4 µl of Mg^{2+} and 9.4 µl of sterile water for a total of 19 µl of reagents. A final concentration of 2 mM of Mg^{2+} was used in each reaction. The cycling program consisted of 30 cycles of a 95^0C denaturing stage for one minute, 54^0C annealing stage for one minute and 72^0C extension stage for two minutes, with an extra ten minute extension at the end of the program. The resulting product was then visualized using gel electrophoresis, and compared to a ladder and positive control. The positive control was obtained from stock cultures that had been put through

the same DNA extraction protocol. The bands were positioned at the target spot on the ladder around 1400-1500 base pairs (bp). The products were then pooled in a gel purification reaction, ligation reaction and then transformed in competent E. coli cells using a TOPO cloning kit and protocol. Colonies were then picked at random and sequenced by the University of Washington. The sequences were then manually analyzed using Sequencher.

Results

Water Dynamics

The transect in the North Pacific covered nearly 30^0 latitude (Fig. 1). Thus water dynamics changed drastically from cold eutrophic waters, to tropical oligotrophic waters (Fig. 3). Surface water temperatures ranged from 11^0C to over 26^0C. Between 49^0N and 40^0N, the water appears to be vertically stratified and the mixed layer deepens to the south (Fig 3c.). Nutrient concentrations were also highly variable with NO_2 ranging from 7 uM L^{-1} to 0.08 uM L^{-1} in the surface waters. Barkley Sound's coastal waters were high in nutrients, with NO_2 peaking at 26.19 uM L^{-1}. Nutrient concentrations were typically highest at the mouth of Barkley Sound and in Alberni Inlet while low concentrations were found in Effingham Inlet (Appendix A). When comparing the two surface samples chosen for cloning, the North Pacific sample was about 5^0C warmer but only had 27% of the nitrate compared to Barkley Sound.

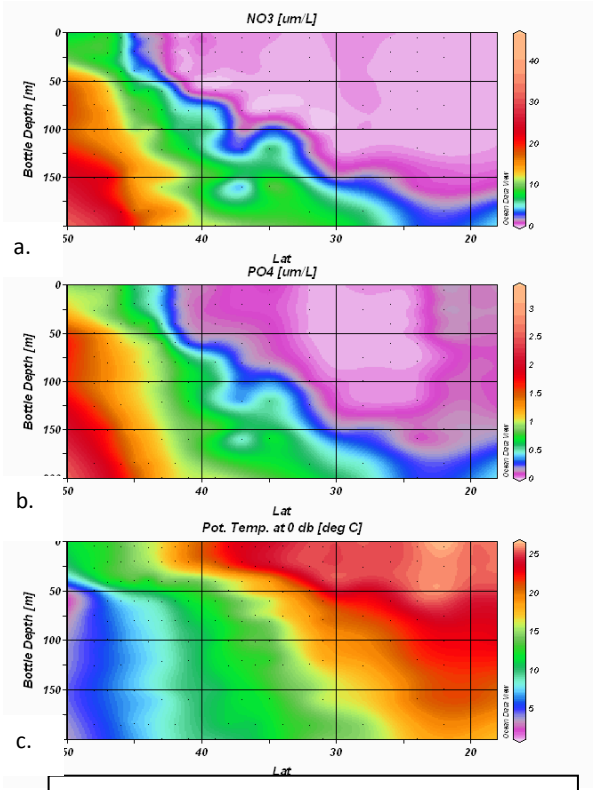


Fig. 3 Nutrient and temperature profiles from the North Pacific transect.

Cloning

Unfortunately, the creation of clone libraries was unsuccessful. Environmental *Synechococcus* DNA was successfully amplified in a PCR reaction and was the correct length, 1400-1500 base pairs, according to the gel electrophoresis. Interestingly, the amplified DNA in Barkley Sound consisted of multiple bands around the target length while the DNA from the North Pacific was only one length (Fig. 4a and 4b). The purified gels were quantified and the North Pacific sample had 24.66 ng μl^{-1} and the Barkley Sound sample contained 57.51 ng μl^{-1}. However, after ligation and plating of the DNA into competent E. Coli cells most of the colonies were unsuccessful in taking up the environmental DNA from both sites. The colonies that looked to contain successful insertions were PCR'd and tested on a gel. However, the gel came back with the primers having amplified the entire plasmid instead of just the insert. It was determined that no colonies had successfully taken up the environmental DNA from either station.

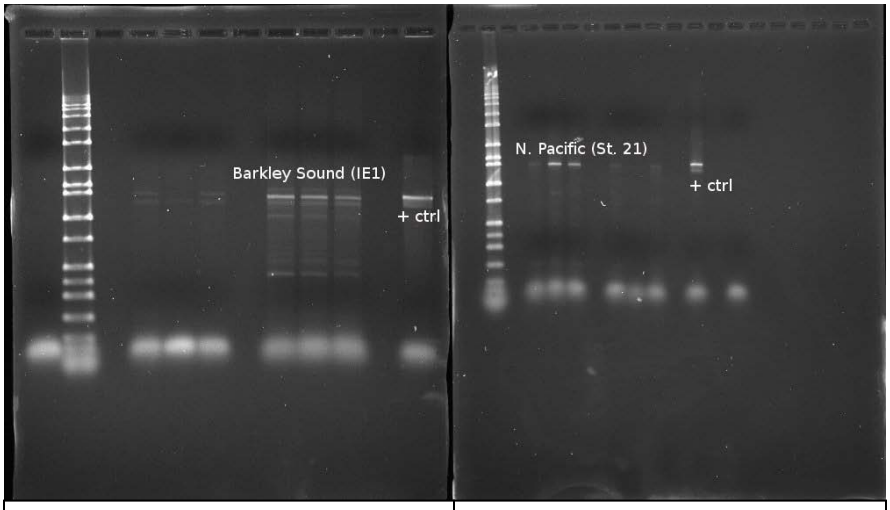


Fig 4a. Gel visualization of PCR product from Barkley Sound station

Fig 4b. Gel visualization of PCR product from North Pacific station

Quantitative PCR

In the North Pacific, Clades I and IV were in the highest abundance north of 37^0N. Clade I had a maximum abundance of $5.55x10^4$ cells mL^{-1} but clade IV reached $2.30x10^5$ cells mL^{-1}, which is over four times greater abundance (Fig. 5). The highest abundances of both clades were found in the surface waters. Clades I and IV were also vertically partitioned in the water column in the northern stations, with clade I comprising a larger percentage of the population deeper in the water column (Fig. 6).

Neither clade was found at detectable levels south of 34^0N. Their abundance drastically decreased after 37^0N as the waters became more nutrient deplete.

Synechococcus was found to be present in the Barkley Sound region at all stations. The maximum abundance of both clades was seen in the Alberni Inlet with clades I and IV peaking at $2.46x10^4$ cells mL^{-1} and $3.25x10^4$ cell mL^{-1} respectively (Fig. 7a and 7b). Overall, the abundances of these organisms throughout the Barkley Sound was fairly consistent. The ratio between clades I and IV ranged from .72:1 to 1.54:1, but generally hovered around a 1:1 ratio.

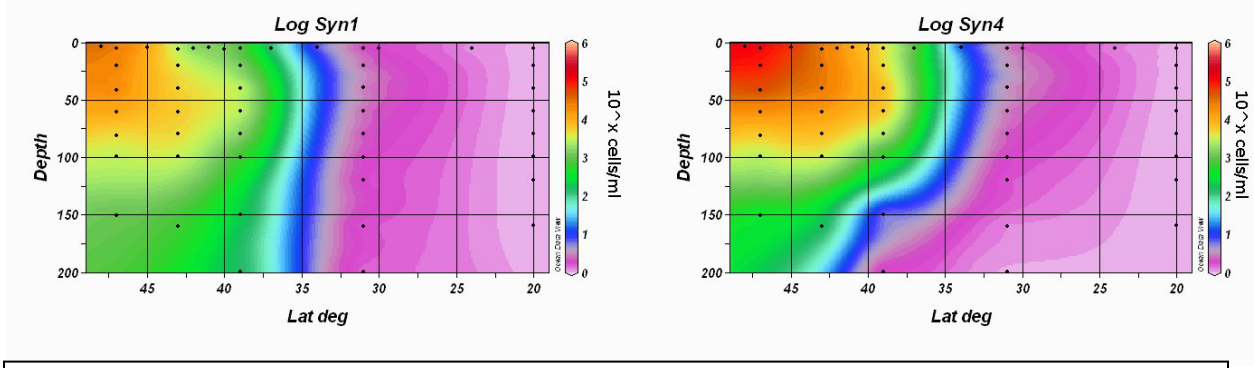


Fig 5. North Pacific distribution of clades I and IV. Color mapping of cells/ml on a logarithmic scale. Black dots represent actual sample sites.

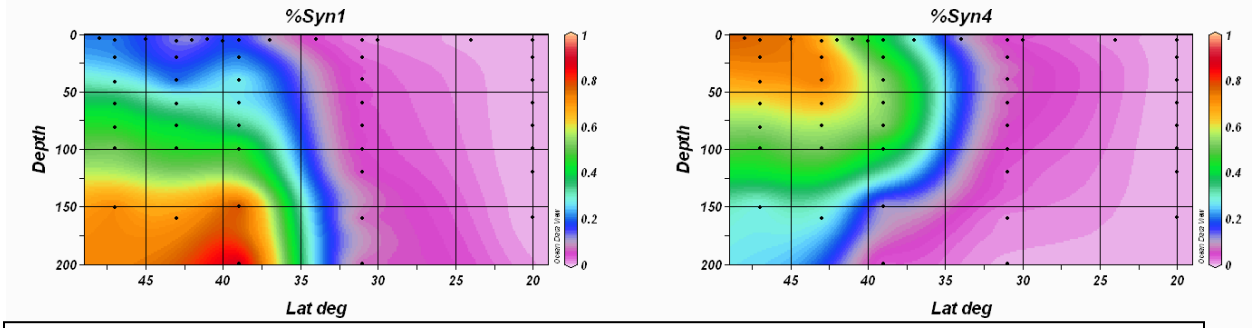


Fig 6. North Pacific distribution of clades I and IV. Color mapping of *Synechococcus* clades represented as a percentage of total cells/ml.

Clone Libraries

Due to time restrictions, the problems with cloning were not able to be formally addressed. A number of factors could have resulted in the failure of suitable colonies for clones. However, with the

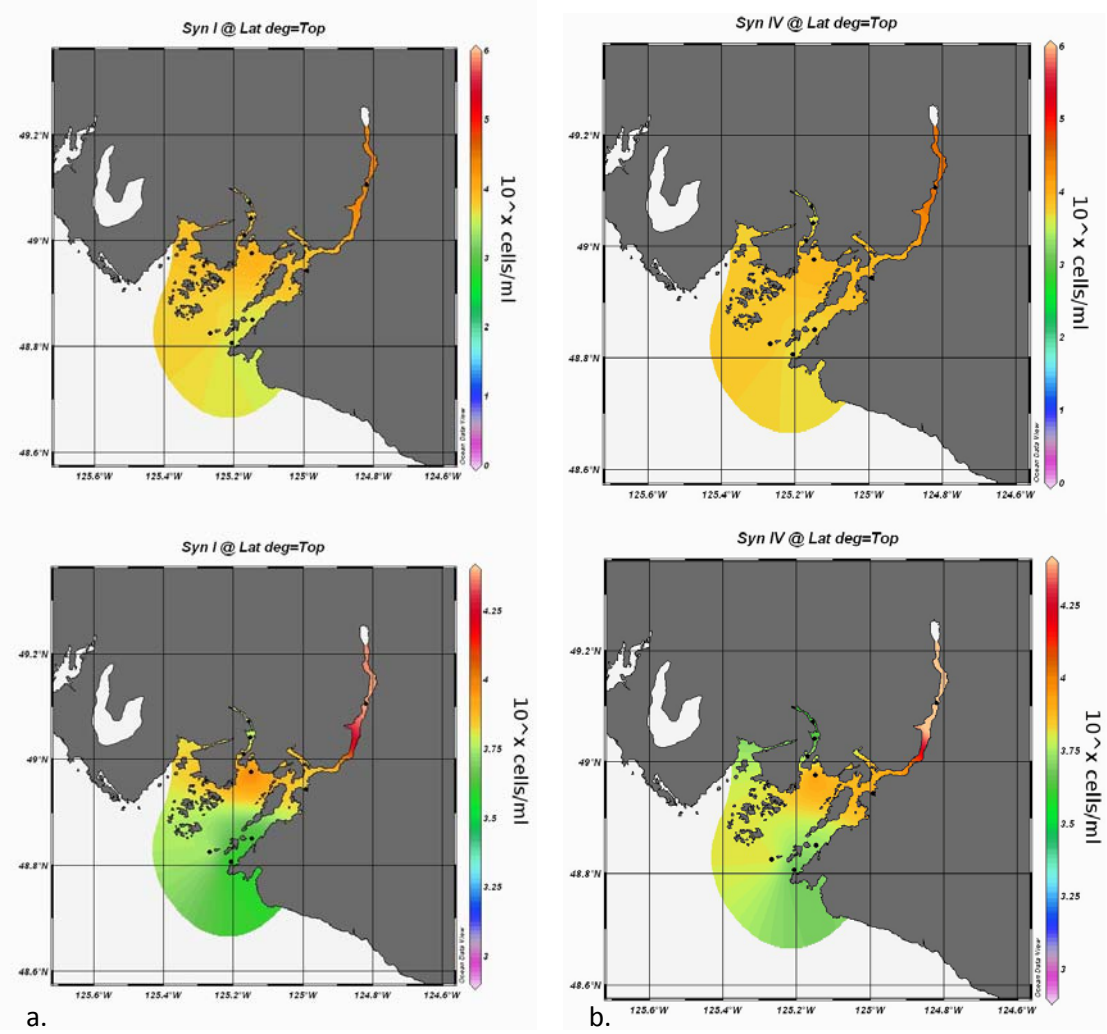


Fig 7a and 7b. Abundance of *Synechococcus* clades from surface samples in Barkley Sound. Top graph's scale matches North Pacific samples. Bottom graph's scale for more detailed differences.

positive amplification of the environmental DNA and successful qPCRs, the issues had to result from a step after the PCR and gel extraction. Also, with the abundance of DNA in the NANO-drip results post-gel purification, there should have been DNA to sit in the open plasmids. It is presumed that a problem occurred either in the ligation, plating or post-plate PCR of colonies.

Discussion

North Pacific Distribution

The transect in the sub-arctic North Pacfic is known as a high nutrient, low chlorophyll (HNLC) region. These areas of the ocean are iron limited and produce low biomass even with high

concentrations of macronutrients (Martin et al. 1988). Both clades I and IV were found to be very abundant above 37^0N, sometimes exceeding 10^5 cells mL^{-1}. Their small size gives them a much higher surface area to volume ratio (SA:V). Organisms with a high SA:V ratio have more surface area to take up nutrients, yet physiologically need less nutrients to survive due to their small volume (Martin et al. 1988). This is crucial in regions of nutrient limitation. While macronutrients are plentiful, iron is still limiting and having a high SA:V ratio allows clades I and IV to thrive and outcompete larger phytoplankton. The absences of these organisms south of 34^0N could be based on temperature, nutrients or competition. The rapid decrease in abundance south of 37^0N follows the appearance of warmer waters and very reduced macronutrient concentrations. Even with a high SA:V ratio, clades I and IV physiologically need higher nutrients than those present in nutrient poor, or oligotrophic, waters. Oligotrophic settings are also typically dominated by other cyanobacteria, specifically *Prochlorococcus,* which most likely are outcompeting *Synechococcus* clades I and IV (Johnson et al. 2006).

Clades I and IV are also vertically partitioned in the water column. While the highest abundance of each clade is near the surface, clade IV is dominant, comprising over 60% of the total *Synechococcus* population in the upper 50 meters. This is also consistent with most previous research showing that when both I and IV occupy the same region, I is found in lower abundance (Tai et al. 2009). The ratio between I and IV in the North Pacific was typically well under 1:1, especially in the surface waters. Clade I, on the other hand, seems to be able to survive deeper in the euphotic zone, with nearly 1000 cells ml^{-1} at 160 meters depth and a ratio to clade IV of 3.53:1. Whether this is an adaptation by clade I to survive in colder or darker waters is not determinable by this data set. However, clade I's ability to change the ratio of its photoharvesting pigments (chromatic adaptation), probably plays a role in its ability to survive deeper, darker waters than other clades (Ting et al. 2002). With concentrations of both I and IV increasing in the surface waters as the transect moves north, it is unknown from these samples the extent to which these organisms can survive. With both clades able to survive water temperatures at

6^0C, it seems that these organisms could have a very large distribution in the high latitude, high nutrient waters of the North Pacific and thus play a significant role in both community structure and carbon cycling.

Barkley Sound Distribution

The consistent and abundant distribution of *Synechococcus* throughout the Barkley Sound was very surprising. Early indications through microscopy of very high diatom concentrations led to the assumption that concentrations of these organisms might be scarce. The recent discovery of clade I and IV's ability to control the uptake of key nutrients like phosphate and metals is a unique adaptation that is beneficial to the survival of these organisms in coastal waters (Zwirglmaier et al 2008). Smaller phytoplankton are typically outcompeted in nutrient rich waters like the waters in Barkley Sound. Clades I and IV's ability to regulate nutrient uptake allows it to compete with much larger diatoms and still thrive in regions that typically see very few small phytoplankton.

The distribution of *Synechococcus* in Barkley Sound is also determined by the movement of water masses. Classic estuarine circulation has fresher surface waters flushing out of the system while deeper saline waters move into the system. Analyzing the water mass movements based on salinity, it appears that waters in the Alberni Inlet, the area of highest *Synechococcus* abundance, exit the sound through the Imperial Eagle Channel instead of Trevor Channel (Linder 2010 unpublished). Interestingly, the region where all these waters intersect in Imperial Eagle Channel (IE3) was the only other station besides the one in Alberni Inlet to reach over 10^4 cells mL^{-1}. Whether the movement of waters from Alberni Inlet through Imperial Eagle affects the distribution of *Synechococcus* in Effingham Inlet is unknown. Clade IV does decrease in abundance deeper into Effingham Inlet, but Clade I shows no clear pattern.

Ecology of Clade I and IV

This study has reinforced the distribution of clade I and IV in coastal waters and has shown that they are not only highly abundant in HNLC regions, but can also be partitioned vertically in the water column. However, it is still unknown what environmental factors control the distribution of these organisms. Using linear regression analysis to compare the abundance of each clade to physical and chemical parameters from both regions yielded very little information. There was almost no correlation between temperature, NO_3, PO_4, NO_2, and NH_4. The highest R^2 value was .157 and was associated with NH_4. As NH_4 increased, so did the abundance of both clades. This is consistent with other research that has found no significant correlation of bottom-up effects on the abundance of these organisms (Tai et al. 2009). Top-down factors, such as predation, are still relatively poorly understood and were not analyzed in this study.

The ratio of these two clades has also been an area of interest. In the North Pacific, the ratio of clades I:IV averaged .29:1 in the surface waters (Fig 8a). This is similar to previous research that found clade IV to be dominant in high latitude, high nutrient parts of the open ocean (Zwirglmaeir et al 2008). The ratio in Barkley Sound estuarine system was near 1:1 throughout (Fig 8b). While research has shown clade IV outnumbering clade I in coastal waters throughout most the year, around the spring bloom between March and May it is not uncommon for clade I to equal or outnumber IV (Tai et al 2009). The collection time of the Barkley Sound samples align with previous research at Scripps Pier showing a bloom in clade I in March, allowing their numbers to match that of clade IV. Thus the ratio found in Barkley Sound was well within expected relative abundances of each clade. Interestingly in Puget Sound, the ratio of clade I:IV averaged 3:1 throughout the year (Rocap unpublished) (Fig 8c). Puget Sound is much closer geographically to Barkley Sound than the Scripps sampling site and is another estuarine system, yet the ratio of clades I:IV differs drastically from both Barkley Sound and Scripps Pier. The waters in the Puget Sound have a much longer residence time, extending up to 120 days, which could

affect the relative abundance of each clade (Babson et al. 2006). While no iron sampling was taken on either cruise, iron could be the determining factor in which clade is dominant. It is interesting that

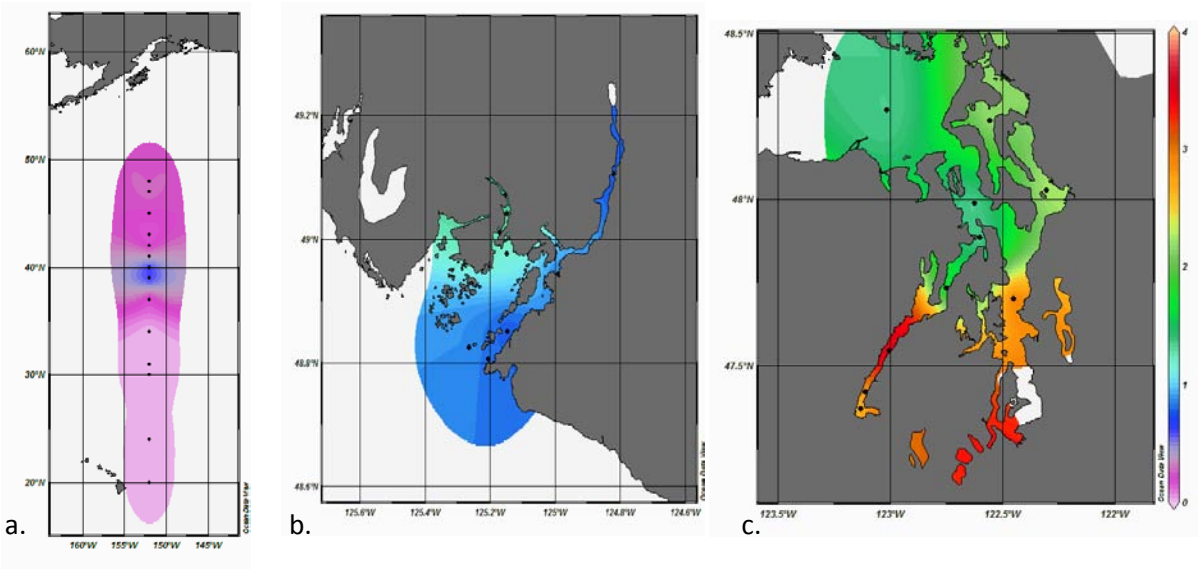


Fig. 8 Ratio of clades I:IV in different sampling regions. Warmer colors/higher numbers represent a higher abundance of clade I compared to clade IV. Scale is the same for each plot. A) North Pacific B) Barkley Sound C) Puget Sound

there is a gradient for the ratio of clades I:IV from the open ocean to coastal waters to estuarine waters that was found to be present. This may be attributed to the abundance or lack of iron in each region. Clade IV may have a higher affinity for iron and therefore can exclude clade I from HNLC regions. In Puget Sound where river runoff is abundant, thus having higher concentrations of iron, clade I is able to thrive and even outnumber clade IV. The ratio in Barkley Sound places this region as a transition zone from open ocean to long residence time estuarine system.

Conclusion

The analysis of the distribution of *Synechococcus* clades I and IV has shown that these organisms are highly abundant and can thrive in numerous environments. Their small size allows them to flourish in HNLC regions of the ocean and their unique adaptations to regulate the uptake of nutrients give them the ability to compete with larger phytoplankton and thrive in coastal waters. However, as with past research, no clear relationship between any physical or chemical factors were associated with the

distribution of these organisms. The ratio of I:IV was consistent with past research showing a dominance of clade IV in the open ocean while they were equal in abundance in Barkley Sound.

The wide variety of suitable habitats still drives at the question of the genetic diversity within these clades. Are the cells in the open ocean different strains than the cells thriving in the coastal environments? Or are their genetic sequences exactly the same and these cells are able to change gene expression to thrive in multiple environments? Successful cloning of these organisms from different regions should be able to answer these questions and determine how these organisms are so abundant. This wide distribution and high abundance make them undoubtedly vital to carbon cycling, oxygen production and the ecosystem in general. The continual study of these organism's genetic information as well as distribution will provide insights into their evolution, and unique adaptations that make them such abundant organisms.

References:

Ahlgren, N. A., and G. Rocap. 2006. Culture isolation and culture-independent clone libraries reveal new marine Synechococcus ecotypes with distinctive light and N physiologies. Applied and Environmental Microbiology 72: 7193-7204.

Ahlgren, N. A., G. Rocap, and S. W. Chisholm. 2005. Measurement of Prochlorococcus ecotypes using real-time polymerase chain reaction reveals different abundances of genotypes with similar light physiologies Environmental Microbiology 8: 441-454

Babson, A. L., A. Kawase, and P. Maccready. 2006. Seasonal and interannual variability in the circulation of Puget Sound, Washington: A box model study.

Martin, J. H. & Fitzwater, S. 1988. Iron deficiency limits phytoplankton growth in the north-east Pacific subarctic. *Nature* 331:341–3.

Moore, L. R., A. F. Post, G. Rocap, and S. W. Chisholm. 2002. Utilization of different nitrogen sources by the marine cyanobacteria Prochlorococcus and Synechococcus. Limnology and Oceanography 47: 989-996.

Rocap, G., D. L. Distel, J. B. Waterbury, and S. W. Chisholm. 2002. Resolution of Prochlorococcus and Synechococcus ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. Applied and Environmental Microbiology 68: 1180-1191.

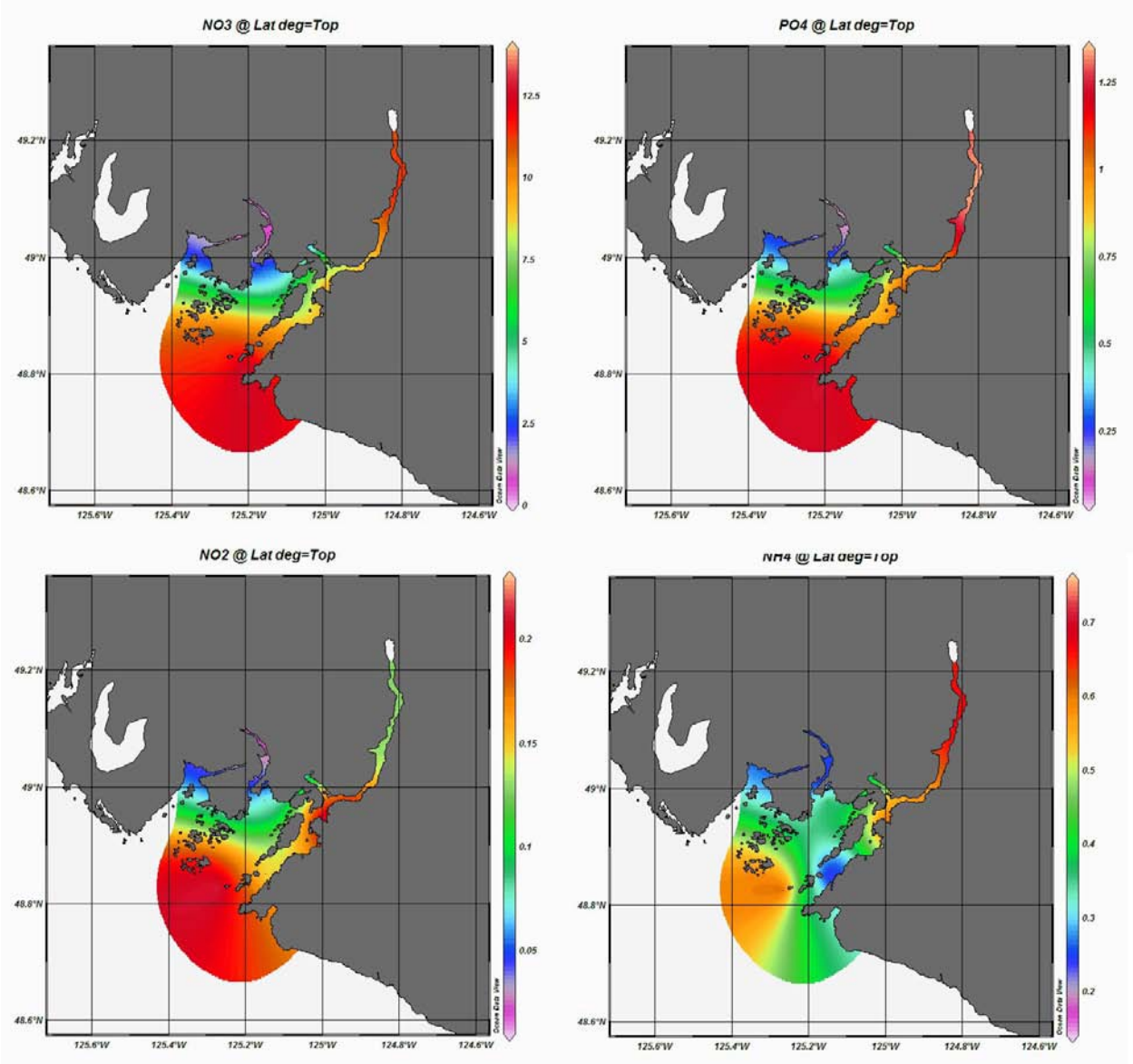
Tai, V., and B. Palenik. 2009. Temporal variation of Synechococcus clades at a coastal Pacific Ocean monitoring site. Isme Journal 3: 903-915.

Ting, C. S., **G.** Rocap, J. King and S. W. Chisholm. 2002. Cyanobacterial Photosynthesis in the Oceans: Origins and Significance of Divergent Light Harvesting Strategies. Trends in Microbiology. 10:134-142.

Whitman, W. B., D. C. Coleman, and W. J. Wiebe. 1998. Prokaryotes: The unseen majority. Proceedings of the National Academy of Sciences of the United States of America 95: 6578-6583.

Zwirglmaier, K. and others 2008. Global phylogeography of marine Synechococcus and Prochlorococcus reveals a distinct partitioning of lineages among oceanic biomes. Environmental Microbiology 10: 147-161.

Appendix A.



Surface nutrient profiles of Barkley Sound. All color mapping in mM/L.