NICHE CHARACTER IN A TEMPORALLY VARYING ENVIRONMENT

by

Shane Patrick Nowack

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Mathematics

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2014

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Isaac Klapper, and my co-advisor, Dr. David M. Ward. I would also like to thank my readers Dr. Jack Dockery, and Dr. Mark Pernarowski for correcting all of the errors and their helpful suggestions. Additionally I would like to thank Ray Spitieri and Jack Dockery for help with the modeling work, Fred Cohan for his comments and suggestions in review, and Millie Olsen, Eric Becraft, Chris Klatt, Al Parker, and Ben Jackson for the assistance they provided in the field. I would also like to thank Millie Olsen, George Schaible, and Eric Becraft for their assistance in maintaining laboratory cultures and obtaining sequence data. Finally, I acknowledge funding provided for this project by NSF-DMS 1022836, Montana Space Grant Consortium, and the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Genomic Science Program 395 (GSP). This contribution originates from the GSP Foundational Scientific Focus Area (FSFA) at the Pacific Northwest National Laboratory (PNNL) under contract 112443.

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES – CONTINUED

LIST OF FIGURES – CONTINUED

LIST OF FIGURES – CONTINUED

ABSTRACT

One of the major goals in the field of ecology is to understand the connection between an organism and its environment. In this thesis both theoretical and empirical approaches were used to investigate the effects of environmental variation on niche structure.

A mathematical model was developed to make predictions about the consequences of temporal frequency regime on optimal behavior. Three different time scales of environmental variation were studied: faster than the growth rate, slower than growth rate, and similar to growth rate. The model results predicted that (i) optimal behavior appears to be independent of fast environmental variation, (ii) niche width is largely determined by slow environmental variation, and (iii) biological clocks may have evolved from environmental variations that occur with a frequency that is comparable to the growth rate of the organism.

Representatives of the predominant organisms inhabiting the microbial mats found in the effluent channels of Mushroom Spring, Yellowstone National Park, were cultivated, and the growth rates of the isolated strains were measured with respect to light, temperature, and availability of dissolved inorganic carbon. The growth rate measurements suggested that closely-related *Synechococcus* species with distinct ecological adaptations exist within the Mushroom Spring community, and may explain the genetic diversity found *in situ*. The results also suggested that the fundamental light niche is interconnected with other environmental parameters, such as temperature and dissolved inorganic carbon availability. To compare the results of the mathematical and microbiological approaches, environmental light data that were collected in the vicinity of Mushroom Spring were incorporated into the mathematical model. The optimal fundamental light niche that was predicted by the model and the measured light niche of one of the cultivated strains exhibited qualitative similarities.

Collectively, this interdisciplinary approach has led to the identification of several environmental characteristics that are hypothesized to be important in determining niche structure.

CHAPTER 1

INTRODUCTION

Motivation

Theodosius Dobzhansky (1956) defined an adaptive trait to be "an aspect of the developmental pattern which facilitates the survival and/or reproduction of its carrier in a certain succession of environments." The term adaptation is somewhat ambiguously defined in the field of biology, but in this thesis, the terms adaptation and adaptive trait (Dobzhansky, 1956) are considered to be one and the same.

Understanding the ecological adaptations of microbial species has proven to have application in many areas of research. For instance, phenotypic traits that define certain ecological adaptations can often be traced back to gene presence/absence in the genome (Barros and Offenbacher, 2009; Roukos et al., 2010). If the functionality of the observed phenotypic trait has an identified utility, such as a byproduct that inhibits the growth of other harmful microorganisms (antibiotic), bioremediation properties (Swenson et al., 2000), or say, cost and time efficient biofuel production (Radakovits et al., 2010; Pittman et al., 2011), then artificial selection of the desired phenotype (Swenson et al., 2000) or genetic engineering (Radakovits et al., 2010) to select for the desired genotype can be used to mass produce the organism. As another example, the effects of environmental variation on niche structure can be

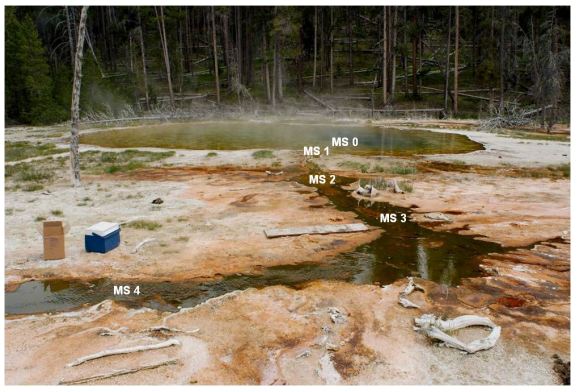
investigated by comparing the historical environmental data from a given ecosystem to the observed ecological adaptations of its inhabitants (Nowack et al., submitted; Chapter 2). This information can then be used to make predictions about the stability of the ecosystem with respect to future environmental changes, such as anthropogenically-induced changes (Rinnan et al., 2007; Buckley et al., 2010; Chevin et al., 2010). These two examples not only offer practical applications for the study of ecological adaptations of microbial species, but they also address two research objectives of considerable interest in the fields of ecology and biology: utilizing the functionality of the organisms among us and understanding the connection between an organism and its environment (Schwenk et al., 2009; Angiletta and Sears, 2011).

A Model Ecosystem

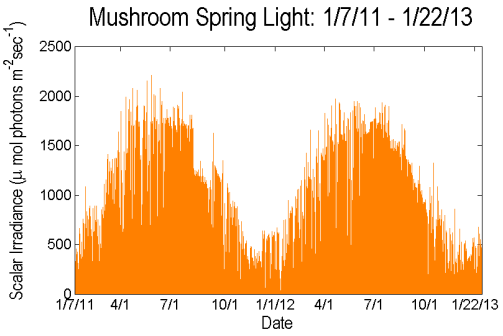
Much of the work presented here was motivated by the following goal: to theoretical predict the niche character of an organism that is inhabiting a temporally varying environment, and compare the theoretical predictions to what is found in nature. To obtain the empirical data required to make such comparisons, a model ecosystem, where the environmental conditions and ecological adaptations of the inhabitants could be measured, was sought. Microbial ecologists have been studying the microbial populations found in the hot springs of Yellowstone National Park (YNP) for decades, in particular, the microbial mat communities found in the effluent channels of Mushroom Spring and Octopus Spring (Ward et al., 1998; Ward

et al., 2006). These natural ecosystems are attractive model systems to study the effects of environmental variation on niche structure for two primary reasons. First, they are relatively simple, very few organisms can withstand the hot temperatures. And second, temporal environmental variations occur over many time scales, particularly with respect to temperature and light; geothermal fluid surges deriving from the hot spring source, day/night cycles, and seasonal changes all contribute to the observed environmental variability (Figure 1.1). These characteristics, combined with the vast amounts of available information regarding the distributions of the mat inhabitants (Ferris and Ward, 1997; Ramsing et al., 2000; Ward et al., 2006, Becraft et al., 2011; submitted), provided an ideal setting in which to study the effects of temporal environmental variations on niche character.

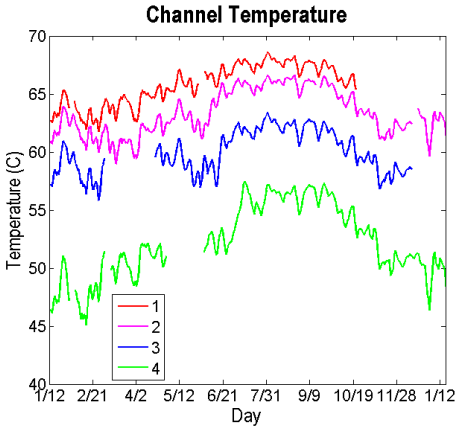
A third characteristic of these hot springs ecosystems worth noting is the existence of environmental gradients that create distinct spatial niches. As the distance from the source pool increases, the water temperature in the effluent channel decreases (Figure 1.1C), and, as depth from the surface of the mat increases, light intensity is attenuated (Figure 1.1D). While these are spatial characteristics and do not relate directly to the theoretical work discussed here, they are extremely important aspects of this ecosystem, and this is made apparent in the presentation of the empirical work in Chapters 4 and 5.



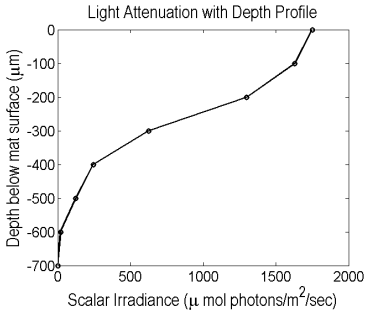
(A)



(B)



(C)



(D)

Figure 1.1. Environmental characteristics of Mushroom Spring, YNP. A) Mushroom Spring, YNP with temperature sites marked along the thermal gradient. B) Light data from Mushroom Spring over two-year period from 7 January 2011 to 22 January 2013. C) Effluent channel temperature data from 12 January 2011 to 12 January 2012 at Mushroom Spring, YNP. D) Light attenuation profile at a 60°C site in Mushroom Spring (modified from Becraft et al., submitted).

Project Summary

My thesis project can be separated into four distinct parts: (i) constructing and analyzing a mathematical model that investigates the effects that temporal environmental variations have on the inhabiting organisms, (ii) collecting *in situ* environmental data from the hot spring ecosystems to incorporate into the model as input, (iii) cultivating isolates that represent the predominant organisms found in the microbial mats of Mushroom Spring, and (iv) characterizing the ecological adaptations of the cultivated isolates in the laboratory. Combining these four parts has provided a collaborative, in-depth approach to investigating niche structure in a temporally and spatially varying environment.

In the remaining sections of this chapter, first, a summary of the literature related to the theoretical work is presented. Second, the original design of each part of the project is provided. And third, the implementation of the design, and some of the difficulties that had to be overcome to accomplish the objectives, are briefly discussed.

Literature Related to Theoretical Work

It has been stated that any theoretical study focused on understanding the connection between an organism and its environment must be based on the niche (Chase and Leibold, 2003). Hutchinson (1957) defined the fundamental niche of an

organism as an n-dimensional hypervolume such that every point in the hypervolume corresponds to an environmental state that would permit the organism to exist, in isolation, indefinitely. At the same time, Hutchinson also defined the realized niche. Hutchinson's fundamental niche represents the niche potential of an organism, while his realized niche reflects the effects of competition, predation and other biotic factors. Both the fundamental niche and the realized niche will be discussed in this thesis.

To supply a mathematical definition of niche it is helpful to define fitness, but like niche, there has been some debate on a universal definition of fitness (Krimbas, 2004). Here, any use of the word fitness is based on the following conceptual definition: (i) zero fitness in an environmental state means the organism cannot survive indefinitely at that condition, (ii) positive fitness in an environmental state means the organism can survive indefinitely at that condition, and (iii) more fit in environmental state A compared to environmental state B, means the organism grows at a faster rate, on average, in A than in B. Thus, the fundamental niche can then be thought of as the set of environmental conditions where it experiences positive fitness in an ideal setting (e.g., the laboratory). The magnitude of this set of environmental conditions will be referred to as the niche width (breadth). Mathematically, the niche will be assumed to be a one-dimensional space that abstractly encapsulates the environment as a whole (unless otherwise noted); hence, the niche

width will be the length of the interval along the environmental axis where fitness is positive.

Richard Levins, a pioneer of mathematically modeling organismal responses to temporally varying environmental conditions, posed this question in one of his classical works (1968): "If a large body size is optimal in a cold environment and a small body size is optimal in a warm environment, what is the optimal body size in a fluctuating environment?" To investigate this question Levins developed a patch model, in which the environment that the organism was inhabiting was either in one of two states (patches). He used the terms fine-grained and coarse-grained to describe the frequency with which the organism changed patches. A fine grained environment is one in which the organism changes patches often over its lifetime (high frequency), and a coarse-grained environment is one in which the organism changes patches rarely over its lifetime (low frequency). Levins then determined the optimal phenotype (niche) in a patchy environment by considering all phenotypes that were of a specified form (optimal fitness at a given point in phenotype space and decreasing fitness in a symmetric fashion as the distance from the optimal phenotype increased (Gaussian-like with fixed variance)), and that satisfied a total fitness constraint (this constraint, or fitness tradeoff, later became known as a specialist-generalist tradeoff (Gilchrist, 1995)). When the two environmental states considered were sufficiently different, in a fine-grained environment the optimal phenotype was determined to be a specialist that chose the most common environment.

On the other hand, in a coarse grained environment the optimal phenotype was a generalist that could survive in both environments, and had fitness within the niche space reflecting the amount of time spent in each patch.

Levins' work was the first major study to optimally determine the fundamental niche in a temporally varying environment. Other related, major theoretical works that investigated the effects of temporally varying environments on niche character include those of Slatkin and Lande (1976), Lynch and Gabriel (1987), and Gilchrist (1995). In each of these studies a fixed form for the niche function was assumed and the governing parameters of the niche functions were optimally determined. Details of the results of these studies are presented in the conclusion of Chapter 2, but the main point here is that all four of these works assumed fixed-form niche functions to determine optimal behavior in temporally varying environments. The main objective of Chapter 2 is to determine the consequences of temporal frequency regime on optimal behavior, which is done by performing an optimization procedure on the niche, and without making any predetermined assumptions about niche form.

The optimization procedure that is discussed in Chapter 2 investigates the effects of temporal variation alone on optimal behavior. But, as has been observed many times (references below), often there does not appear to be a single optimal species in any given environment, but rather, several coexisting species operating as a community (or co-occurring – and I do not differentiate between the two terms here). In Chapter 3, the competitive ability of the theoretically-determined optimal

species is investigated in order to study the competitive interaction between the optimal species and another sub-optimal species. Some of the major works related to species coexistence and competitive exclusion are now briefly reviewed.

The theory of competitive exclusion (Gause, 1934) states that only one species can exist per niche, but past observations of many closely-related phytoplankton species seemingly coexisting on a single limiting nutrient appeared to contradict this theory. This so-called "paradox of the plankton" (Hutchinson, 1961) has led to many questions regarding speciation and niche structure. Two well-known questions originating from Hutchinson's work are: what are the underlying mechanisms responsible for enabling species with identical or nearly identical niches to coexist? And, along an environmental gradient, where there appears to be an infinite number of niches, why are there not an infinite number of species, i.e., what are the mechanisms responsible for the empty niche space and the species discretization that is observed in nature?

Several studies have proposed that temporal variation in environmental conditions may increase niche dimension, offering an explanation for how one limiting nutrient can support similar, coexisting species in a single spatial niche (Lenas and Pavlou, 1995; Smith and Waltman, 1995; Chesson and Huntly, 1997; Wolkowicz and Zhao, 1998; Chesson, 2000; Litchman and Klausmeier, 2001). The periodically varying chemostat (Lenas and Pavlou, 1995; Smith and Waltman, 1995; Wolkowicz

and Zhao, 1998; Litchman and Klausmeier, 2001) has often been used as a modeling platform to study coexistence in a temporally varying environment. In these studies the incoming nutrient concentration and/or the dilution rate (wash-out) is varied, and the parameter space that supports coexistence is determined. Results have suggested that a necessary condition for coexistence is invasibility (Smith and Waltman, 1995; Litchman and Klausmeier, 2001; Siepielski and McPeek, 2010) – both species must be able to invade the other at different times of the environmental cycle in order for the two species to coexist. The coexistence mechanism itself is often proposed as a trade-off, such as a gleaner-opportunist tradeoff (Smith and Waltman, 1995; Litchman and Klausmeier, 2001) in which the gleaner species is one that thrives under low nutrient concentrations (or is able to survive at low densities), whereas an opportunist is a species that thrives when nutrient is at a high concentration (or when the population is at a high density). Similarly, other chemostat studies have demonstrated that two species that have complementary growth rates at different temperatures can coexist (Descamps-Julien and Gonzalez, 2005). Other theoretical studies investigating temporal-based coexistence (and not done in a chemostat), most notably the work of Peter Chesson (Chesson and Huntly, 1997; Chesson, 2000; Chesson, 2008), have also proposed certain trade-offs that may be possible coexistence mechanisms. For example, Chesson coined the term relative nonlinearity to describe a temporally-induced coexistence mechanism; he defined

this as different species having different nonlinear responses to competition in a fluctuating environment.

The preceding competition studies investigated temporal variation alone, but there is also literature that introduces a spatial dimension to study speciation along a gradient. For example, May and MacArthur (1972) utilized the ideas of "species packing" and "limiting similarity" to quantify how species must differ in resource consumption in order to coexist in a temporally varying environment. May and MacArthur (1972), and many others (Doeboeli and Diekmann, 2003; Pigolotti et al., 2007; Leimar et al., 2008; Yamauchi and Miki, 2009 – these four studies do not have a temporal component), applied Lotka-Volterra competition models with various, fixed-form competition kernels, to study speciation along environmental gradients. This continues to be an active of area of research, and recent results from these studies (Doeboeli and Diekmann, 2003; Pigolotti et al., 2007; Leimar et al., 2008; Yamauchi and Miki, 2009) have provided hypotheses for the observed phenotypic clustering and empty niche space that is often found in nature. It has been noted on more than one occasion that the form of the competition kernel, i.e., the choice of a fixed-form function that quantifies the competitive ability of a species, plays an important role in determining the theoretical results (Pigolotti et al., 2007; Yamauchi and Miki, 2009).

While the characteristics of Mushroom Spring define a simple, natural ecosystem that is conducive to studying both the effects of temporal variations on niche

structure and speciation along an environmental gradient, the focus of the theoretical work presented here is on the former. One reason for addressing the existence of environmental gradients and highlighting some of this literature is for breadth, in terms of acknowledging that other factors likely play an important role in determining community structure, possibly in tandem with temporal variations. Another reason is the repeated reference to these environmental gradients in the empirical work that is included in Chapters 4 and 5. A thorough literature review related to the empirical work is provided in the introduction section of Chapter 4.

Design

Model Development

The hot spring communities and the literature discussed above played an influential role in model design. It was decided that a temporally varying chemostat model would be an appropriate, concrete modeling platform that could be molded in such a way that would allow investigation into the questions of interest. The chemostat is often used to study microbial behavior in aquatic ecosystems and it was ideal for the reasons that follow. First, the relative simplicity of chemostat operation is tractable from a mathematical-analysis perspective: a self-explained nutrient vessel, an action-vessel where organisms utilize the available nutrient source and produce more cells and byproducts, and a waste vessel, where cells and byproducts from the action vessel are disposed. Second, the standard temporally varying chemostat

model (Smith and Waltman, 1995; Wolkowicz and Zhao, 1998) is extendable in such a way that allows the model to accept environmental data as input. Then, with the inclusion of a fitness response function in the growth rate and consumption terms, that is environmentally-dependent, the chemostat now possesses a niche component. That is, the key difference in the chemostat model discussed here compared to the one found throughout the literature is the introduction of a function that represents the niche. Third, the organisms of interest with respect to the motivating hot spring communities reside in the effluent channels of the source pool, in which the overflowing water provides nutrients to mat inhabitants found on the surface, and also carries away byproducts and cells, much like a standard chemostat. It should also be noted that a chemostat model has a notable drawback, with respect to modeling the niches of the organisms inhabiting Mushroom Spring. Namely, the inhabitants of Mushroom Spring are found both on the surface and in the subsurface of the microbial mats. While the chemostat may provide a conceptually agreeable modeling platform for the surface populations, the same cannot be said about the subsurface populations. Mathematically modeling the populations along thermal and vertical gradients would require a spatial component to be included in the model. Models such as these are discussed in Klapper and Dockery (2010).

In Chapter 2 a one-species chemostat model is analyzed to find the optimal fitness response of an organism with respect to a single varying environmental parameter (e.g., temperature or light), and the fitness response function was used as

a proxy for the optimal fundamental niche. The model allows the user to enter any environmental history data of their choice. Thus far optimization results for two different environmental inputs have been obtained: a sinusoidally varying environment and environmental light data collected from Mushroom Spring, YNP.

In Chapter 3 the number of species in the temporally varying chemostat is increased to investigate the competitive ability of the optimally-determined species. Two-species competition in the chemostat between the optimal species and an arbitrary sub-optimal species is studied to test the validity of the choice of the objective function that was made. Specifically I ask, can the optimal species always outcompete any other species in a two species system?

Data Collection

Environmental data from Mushroom and Octopus Spring (MS and OS, respectively), YNP were collected in the form of water temperature and light intensity. Water temperature data (time point every 20 minutes) at eight different sites along the flow path of the main effluent channel of both hot springs, MS and OS sites 1-4 (Figure 1.1), were collected from 12 January 2011 to 12 February 2012. Light intensity data in the vicinity of the source pools of the two hot springs were collected from 7 January 2011 to 22 January 2013 (time point every 30 minutes). The data were collected with the intent of supplying the model with *in situ* environmental input.

Additionally, twelve times over the two-year period from 7 January 2011 to 22 January 2013, water samples from MS and OS sites 0-5 and mat samples from MS and OS sites 1-4 were collected (Figure 1.1). The mat and water samples were collected for two main purposes: (i) to compare water and mat species distributions and (ii) to identify possible shifts in the distribution of ecological species over the seasons. Thus far only a subset of these samples has been analyzed, and a manuscript is being drafted that compares mat and water populations over a three-month period (Jackson et al., in preparation).

Cultivation and Phenotypic Analyses

To cultivate isolates representative of the predominant inhabitants found in Mushroom Spring, YNP, first, identification of the predominant organisms is required. A universally accepted definition of microbial species is lacking, and this poses a challenge to identifying the predominant organisms. Molecular cut-offs such as 70% DNA-DNA reassociation (Wayne et al., 1987) and 97% identity at the 16S rRNA locus (Stackebrandt and Goebel, 1994), as well as ecological species (ecotypes) definitions (Cohan, 2002; Ward and Cohan, 2005; Cohan and Perry, 2007) have all been proposed, but none are universally accepted. Here, the term species is defined as described by Cohan and Perry (2007), as populations whose members are ecologically interchangeable.

David Ward from the Land Resources and Environmental Sciences Department

at Montana State University and his students and collaborators have been investigating microbial species and speciation for many years (Ferris and Ward, 1997; Ward, 1998; Ramsing et al., 2000; Ferris et al., 2003; Ward and Cohan, 2005; Ward et al., 2006; Allewalt et al., 2006; Becraft et al., 2011; Melendrez et al., 2011; Becraft et al., submitted; Melendrez et al., in preparation). Particular focus has been placed on identifying the fundamental species-like units with respect to the primary producers of the microbial mats in Mushroom Spring – oxygenic phototrophs from the genus *Synechococcus* (Brock, 1978; Ward et al., 2012), a type of thermophilic cyanobacterium. In both single- and multi-locus approaches a theory-based evolution algorithm, Ecotype Simulation (Koeppel et al., 2008), was used to predict ecological species populations from DNA sequence variation found in mat samples collected from Mushroom Spring (Becraft et al., 2011; Melendrez et al., 2011; Becraft et al., submitted; Melendrez et al., in preparation). In Chapter 4 and Chapter 5 sequence variation in the *psaA* gene (encoding a protein required for photosynthesis) found in environmental samples was used to demarcate ecological species. The hypothesis that variation at the *psaA* locus can discern the ecologically distinct *Synechococcus* populations in Mushroom Spring is tested in Chapter 5.

Ward and colleagues were awarded funding from a Department of Energy Joint Genome Institute Community Sequencing Program to obtain 18 genome sequences of predicted *Synechococcus* species from Mushroom Spring. The purpose of the funding was to investigate the ecological adaptations within and among predicted species

populations, which would further test the hypothesis that the genetic diversity found *in situ* (Ferris and Ward, 1997; Ward et al., 2006; Becraft et al., 2011, submitted) is due to the existence of distinct ecological species. The proposed plan, to accomplish this objective, was to cultivate multiple representatives within predicted species populations, and then apply a combined phenotypic/comparative genomic analyses on the cultivated strains.

My role in this project was to, first, obtain the 18 unicyanobacterial isolates through cultivation methods, and then, second, to measure the growth rates of the isolates with respect to different environmental variables (e.g., temperature and light). Accomplishing these tasks would result in obtaining the information needed to compare the theoretical predictions of the mathematical model to the ecological adaptations of the inhabitants. Moreover, obtaining the targeted isolates would allow several other hypotheses to be tested, three of which are now stated. First, closely-related *Synechococcus* species with distinct ecological adaptations exist in the Mushroom Spring microbial mat, and may explain the distribution patterns of genetic diversity found *in situ*. Second, the environmental parameters that comprise the fundamental niche are interconnected. And third, variation in the *psaA* gene is a species demarcation method that is able to discern the ecologically distinct *Synechococcus* populations. These hypotheses have been tested and the results are presented in Chapters 4 and 5.

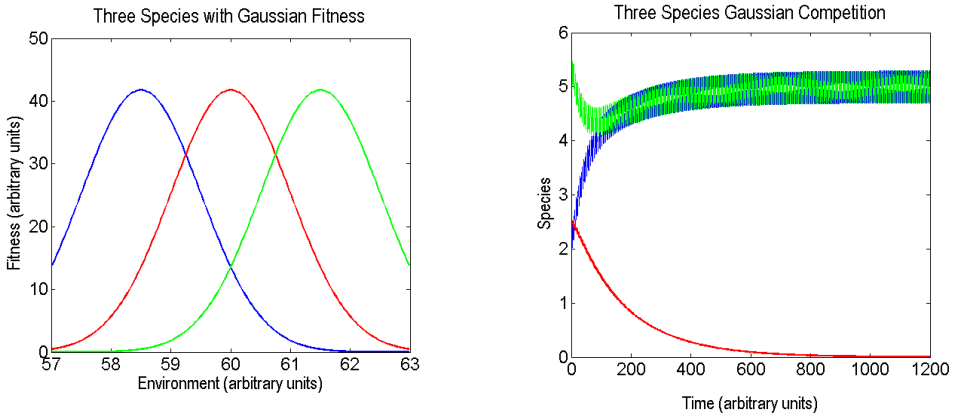
Another graduate student in Ward's lab, Millie Olsen, is performing a comparative genomic analyses of the same strains that I had phenotypically characterized. The results from the phenotypic analyses have provided information regarding the ecological adaptations of several of the strains, while the comparative genomics analyses have been used to search for gene presence/absence differences that may explain the observed adaptations, or alternatively, lead to new hypotheses that can be experimentally tested.

Implementation of Design

Model Implementation

The n-species chemostat model is a system of $n + 1$ nonlinear ordinary differential equations (ODE). Numerical simulations, with $n = 3$ and with an assumed sinusoidally varying environmental history in place, were initially computed. All three species were assumed to have Gaussian fitness (Figure 1.2). The objective was to find two sub-optimal species that could work together to outcompete the optimal species, and then continue to stably coexist on a single limiting nutrient. Sub-optimal in this instance indicates that the mean of the Gaussian is not at the average of the fluctuating environment, and optimal indicates that the mean of the Gaussian is at the average of the environment, with variance of all three Gaussians identical. If this could be accomplished, it would not only provide evidence to support the hypothesis that temporal environmental fluctuations may lead to spatial

coexistence (Smith and Waltman, 1995; Chesson and Huntly, 1997; Chesson, 2000; Litchman and Klausmeier, 2001), but more specifically, possibly lead to new hypotheses about communities competing against other communities as a mechanism that leads to the niche structure that is found in nature. Using MATLAB's (Math-



(A) (B)

Figure 1.2. Three-species competition with species possessing Gaussian fitness responses. A) Fitness responses of three competing Gaussian species in the chemostat. B) Results of *in silico* competition experiment showing two suboptimal Gaussian species outcompeting the optimal Gaussian species. The optimal Gaussian is one that has mean at the average environmental condition.

works; Natick, MA) built-in ODE solvers to compute the solutions of the chemostat equations led to identifying some problems that the built-in solvers had with this particular system of equations (see Chapter 2, equation 2.2). That is, the different ODE solvers (ode45, ode15s, ode23t, ode23tb, ode113) produced different results.

In response I began writing my own codes first, and checking them against MATLAB's built-in functions before believing any of the output that was produced. I determined this process to be necessary as MATLAB's built-in functions tend to be more computational inexpensive than the codes that I write, and the codes that I write are easier to troubleshoot and check for accuracy. Being made aware of the limitations of proprietary software and its built-in functions was an important and useful lesson, and fortunately it was one I learned early on in my graduate work.

After resolving the numerical inconsistencies, I found that it was relatively easy to find two sub-optimal species that could outcompete the optimal species. The results (Figure 1.2) were then shared with a collaborator who is an expert in microbial speciation (Fred Cohan, Wesleyan University). It was suggested that Gaussian fitness functions may not accurately reflect what is found in nature, leading us to take the the novel approach that became a big part of my thesis work. That is, without making any initial assumptions about the form of the niche function, what is the relationship between the optimal niche function and the environmental history? Optimal is defined here to mean that a species that possesses the optimal niche function is able to outcompete any other species (with any other niche function) in a two-species chemostat.

To calculate the optimal species a constrained optimization procedure was applied. Initially, the arithmetic mean of the nutrient concentration in the chemostat was selected as the the objective function to be minimized. The optimization

procedure was performed over three frequency regimes to investigate the effects of temporal variation alone on optimal behavior. Two time-scale asymptotic approximations (Holmes, 1995) to the ODEs were found for the limiting frequency cases (where the environment is changing either very quickly or very slowly with respect to the effective growth rate of the species), and allowed for a calculus of variations approach to be applied to determine the optimum. In the intermediate frequency cases (again, with respect to the effective growth rate) a numerical optimization approach was required.

Numerical Approach. Numerically defining the optimization problem, by defining the fitness response function as a piecewise constant function in the environmental variable, was quite challenging and the details of this setup are included in Appendix A, Section A.4. Solving the numerically-defined problem in a computationally efficient manner was also challenging, and resulted in a thorough investigation of constrained optimization methods and how they handled the intricacies of my problem. Broyden's method (Broyden, 1965), penalty methods, conjugate gradient methods, and augmented Lagrangian methods (Nocedal and Wright, 2006) were all implemented to solve the optimization problem, but these methods took days to compute the optimal fitness response for a single frequency, with rather insufficient resolution (12 to 24 piecewise constant subintervals). Finally, a sequential quadratic programming (SQP) method (Powell, 1978) was implemented that solved

the numerical optimization problem, of a given frequency, in an hour or less, with 200 piecewise-constant intervals. SQP methods are a generalization of Newton's method that solve the Kuhn-Karush-Tucker equations for constrained optimization problems (with both inequality and equality constraints). The key to efficient computation with SQP methods lies in the choice of the updating scheme of the Hessian of the Lagrangian, with which the Broyden-Fletcher-Goldfarb-Shanno method (Nocedal and Wright, 2006) is commonly incorporated. All numerical methods that were implemented produced consistent output, with the computational efficiency being the biggest variable.

A numerical approach was also required to compute the optimal fitness response when the environmental light data were incorporated into the model. Model results were then compared to the laboratory-determined light niche of one of the isolated *Synechococcus* strains obtained from the cultivation efforts.

Two Species Chemostat to Test Optimality. Once an optimization strategy was in place the next step was to investigate whether the objective function chosen, the arithmetic mean of the nutrient concentration, would result in a fitness response that was indeed optimal. That is, validation that the optimal species was able to outcompete any other species in a two-species chemostat was sought. A stability analysis from Floquet theory (Chicone, 2006; Rasmussen, 2007) was applied to the limiting frequency cases by linearizing about a known solution of the two-species

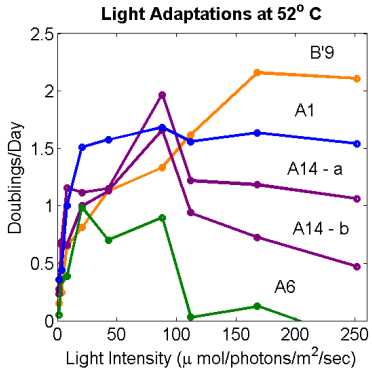
system (such as the asymptotic solutions of the one species system). The details of the stability analysis are included in Chapter 3. While the stability analysis did not produce the desired result of confirming that the optimal species did outcompete all other species (see below), it did lead to the observation that minimizing the geometric mean of the nutrient concentration, versus the arithmetic mean, may be a better choice for the objective function, at least in the low frequency case. Numerical results in Appendix A, Section A.2, support this statement. The idea of minimizing the geometric mean was borrowed from earlier work (Haldane and Jayakar, 1963; Gilchrist, 1995), and with the geometric mean as the objective function, it was discovered that in the low frequency case one eigenvalue of the monodromy matrix of the variational equation was always exactly zero, which has been both a source of interest and confusion. With one zero eigenvalue, stability can not be concluded using this analysis. This approach did lead to results regarding the invasibility properties of the optimal species, and also other, sub-optimal species. An alternative approach to investigate stability is currently being studied and is briefly discussed in Appendix B, Section B.2.

Implementation of Empirical Approaches

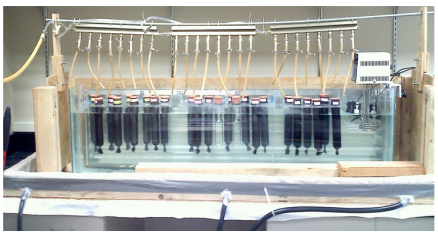
Cultivation. The basic cultivation approach is simple in theory: collect mat samples from the field, homogenize and dilute the mat samples to extinction, pick

the colonies that grow on the plates representative of the highest dilution, scale-up the colony in liquid medium to a large enough volume to extract DNA, and then obtain sequence data (see Appendix C, Figure C.1). If the sequence matches a predominant sequence found *in situ*, and does not contain any other sequences, then successful cultivation of an isolate representative of a predominant organism has been achieved. Theoretically this sounds simple, but it has been proposed that less than 1% (Giovannoni and Stingl, 2005) of the microbes on this planet have been successfully cultivated and are in pure culture. My efforts attempting to cultivate isolates representative of the predominant organisms found in Mushroom Spring reflect this difficulty. After five years of applying various cultivation protocols, several of the proposed 18 genomes either have been or are in the process of being sequenced. The details of my cultivation efforts are included in Chapter 5.

Phenotypic Analyses. After obtaining isolates representative of the predominant species the goal was to phenotypically characterize the isolates with respect to light and temperature. Eight light conditions (using neutral density covering to attenuate the light), at an initial temperature of 52°C, were established. The light conditions ranged from 1 μmol photons/m^2/sec to 250 μmol photons/m^2/sec. Preliminary results suggested that higher light intensity capabilities than the current facilities allowed for were required to determine the upper-light limit of the organisms (see Figure 1.3A). Significant effort was made to purchase an incubator with higher



(A)



(B)

Figure 1.3. A) Preliminary light adaptation results with low-light incubator and without bubbled CO_2. The old equipment was able to achieve intensities of 250 μ mol photons/m^2/sec and not able to bubble CO_2. B) New growth chamber with CO_2 bubbling feature and scalar light intensity capabilities that reach up to 3000 μ mol photons/m^2/sec.

light capabilities. But, because extreme temperatures were required, above 52°C, a device that could stably maintain higher light at these temperatures was not found to be available on the market. In addition, upon presenting these results (Figure 1.3A) at a meeting, the photosynthesis expert present (Donald Bryant, Penn State University) informed me that I was not truly measuring light adaptations, since saturating levels of CO_2 were not being provided. Therefore, part of a summer was spent at Penn State University in Bryant's lab learning how they made these types of measurements and how to construct an apparatus that was able to attain high light intensities and bubble CO_2 (see Figure 1.3B).

The chamber that I constructed can achieve scalar light intensities of up to 3000 μmol photons/m^2/sec (approximately equivalent to the sunlight reaching the mats

at peak intensity near summer solstice (S. Nowack, unpublished observations)), and can bubble CO_2 to 24-100 mL cultures at one time. Thus far I have conducted growth experiments on 10 of the isolates under various combinations of imposed environmental conditions (focusing on light, temperature, and the availability of dissolved inorganic carbon) in order to test the hypotheses that were stated in the design section and to obtain data to compare to model results. The details of these experiments are described in Chapters 4 and 5.

CHAPTER 2

CONSEQUENCES OF TEMPORAL FREQUENCY REGIME ON OPTIMAL BEHAVIOR

Contribution of Authors and Co-Authors

Manuscript in Chapter 2

Author: Shane Nowack

Contributions: Partially obtained funding (MSGC), designed and analyzed the model, collected the field data, conducted the experiments, and wrote the manuscript.

Co-Author: Isaac Klapper

Contributions: Partially provided funding, assisted with experimental design, discussed the results and edited the manuscript at all stages.

Co-Author: David M. Ward

Contributions: Partially provided funding, assisted with experimental design, discussed the results and edited the manuscript at all stages.

Manuscript Information Page

Shane Nowack, Isaac Klapper, and David M. Ward
Journal of Theoretical Biology
Status of Manuscript:
___ Prepared for submission to to a peer-reviewed journal
_x_ Officially submitted to a peer-reviewed journal
___ Accepted by a peer-reviewed journal
___ Published in a peer-reviewed journal

Abstract

It is widely accepted that temporal fluctuations in environmental conditions play a role in defining niche parameters. Here a general theory of effect of temporal history on optimal behavior is proposed. A constrained optimization procedure to predict an organism's fitness response to a variable environment is presented under the assumption that one environmental parameter is temporally fluctuating. Results suggest that fluctuations at a low or intermediate frequency, but not high frequency (with respect to the effective growth rate), may play an important role in defining fundamental niche width along the axis of the fluctuating environmental variable. On the other hand, fluctuations at intermediate, but not high or low frequencies, may naturally lead to a fitness advantage of biological clocking mechanisms. That is, among the model predictions is that biological clocks may be a behavior that emerges in response to temporal environmental fluctuations on intermediate time scales. To test some of the predictions, light data collected near Mushroom Spring in Yellowstone National Park are then incorporated into the model and the output is compared to the observed fundamental light niche of a particular Mushroom Spring inhabitant.

Introduction

Almost every ecosystem experiences temporal environmental variation at some

level – day/night cycles, seasonal changes, and climate changes occur over a wide range of time scales and influence the behavior of many organisms. Understanding the relationships between temporal fluctuations of a particular ecosystem and the niche characteristics of its inhabitants may provide valuable insight with respect to certain evolutionary processes such as adaptation, diversity, and extinction (Levins, 1968; Lynch and Lande, 1993; Gilchrist, 2000; Chevin et al., 2010), and more specifically, provide insight into mechanisms for response to temporal variation. Recent technological advances have made obtaining detailed records of the environment manageable; for example, the use of microsensor technology to precisely track *in situ* environmental variables such as light intensity, temperature, oxygen concentrations, and pH over impressively small spatiotemporal scales has become increasingly common over the last several years (Kühl, 2005; Becraft et al., 2011; Bernstein et al., 2013; Becraft et al., submitted). These kind of detailed data sets open new possibilities for predictive methods, such as mathematical models, to investigate the effects that environmental variation has on the niche character of an inhabiting organism.

In 1957 Hutchinson defined the fundamental niche of a species as an n-dimensional hypervolume where every point corresponds to an environmental state that would permit the species, in isolation, to exist indefinitely. At the same time, Hutchinson also defined the realized niche. Hutchinson's fundamental niche represents the

niche potential of an organism, while his realized niche reflects the effects of competition, predation and other biotic factors. Niches can be expected to depend on many factors, but here, as is common practice, we will simplify to one environmental dimension and view the Hutchinsonian niche in one dimension only, with other environmental factors considered fixed. Importantly, we view this single environmental dimension in the abstract. Rather than exploring consequences of a particular environmental condition (e.g., Q10 laws for growth as a function of temperature), the primary objective here is to study, in a pure sense, the magnitude and extent of variation that an organism's niche character inherits from its environment. Specifically we ask: how much ecological character can be predicted just from variability and without appealing to the details and peculiarities of a particular environmental effect?

Ecological niche models in variable environments are abundant in the literature. However, a seemingly standard practice in these models is the selection of a fixed form for the function that defines the niche properties. For example, Gilchrist (1995) modeled the temperature niche in a thermal sensitivity study of ectotherms by preselecting a Logan curve (he looked at Gaussian and rectangular curves as well) to model fitness with respect to temperature, and then optimally fit the niche-defining parameters to the curve (temperature optima and range for net growth in this case). Similarly, resource competition along an environmental gradient has often been used to model the realized niche and investigate competitive exclusion/coexistence

mechanisms in both constant and fluctuating environments (May and MacArthur, 1972; Chesson and Huntly, 1997; Litchman and Klausmeier, 2001; Gyllenberg and Meszéna, 2005; Pigolotti et al., 2007; Leimar et al., 2008; Yamauchi and Miki, 2009). In these realized-niche models the competition coefficients are derived from predetermined utilization functions, which again involves an assumption on the form of how the fitness of one species affects the fitness of another – usually in the form of a Gaussian. The model we propose does not introduce a predetermined form for the niche function; rather, the extent that temporal fluctuations drive niche structure is investigated through environmental information alone, i.e., the form of the niche is an emergent property. And indeed, the model does not predict simple forms such as Gaussians emerging. More importantly, allowing plasticity in fitness form may lead to insights that cannot be gained from study of fixed forms like Gaussians, including important aspects of the different ways that organisms respond to different kinds of temporal variability.

For purposes of choosing a concrete modeling platform, the impact of a temporally varying environment on fundamental niche character is investigated via a one-species model of a theoretical, spatially homogeneous chemostat. Chemostat systems provide rough proxies for many environmental ecosystems; at the same time, chemostat models have the advantage of mathematical tractability. We attempt here to study effects of temporal variations in as simple of a setting as possible, i.e., what

behaviors emerge from temporally varying environments even before the complexities of those environments are considered? Because of its simplicity, the chemostat is a good a place for such a study. A further advantage here for chemostats is the relevance to an application to surface populations of microorganisms that are found near the mat/water interface in effluent channels of hot springs. The organisms of interest are found within the top one millimeter of the mat surface and, by way of the overflowing water, have nutrients flowing in and nutrients and their byproducts flowing out, which in many ways mimics the three-vessel operation of a standard chemostat.

The chemostat environment in the model varies with respect to time according to a supplied environmental history function (e.g., temperature or light). The fitness response of the organism describes how the organism allocates its growing efforts with respect to the varying environment, is included as an unknown function in the growth rate and consumption terms of the model, and is determined by applying a constrained optimization procedure (a constraint is introduced to enforce a fitness trade-off). The optimal fitness response function will serve as a proxy for the fundamental niche. For cases when the environment is changing very slowly or very quickly the optimization is calculated by applying an asymptotic approach, while solutions for intermediate environmental fluctuations are computed numerically.

To support the theoretical findings, empirical light data collected near the source pool of Mushroom Spring, Yellowstone National Park (YNP) were included in the

model as the environmental history function. Previous molecular studies have revealed that oxygenic phototrophic *Synechococcus* spp. are the primary producers of the microbial mats found in the effluent channels of Mushroom Spring (Brock, 1978; Ward et al., 2012), suggesting that light may be an important temporally varying, niche-defining characteristic in this ecosystem. The areas in and around this particular spring, as with any other geographical location, have a uniquely defined light history; weather patterns and shading from landscape features (such as trees and mountains) add complexity to the predictable daily, seasonal, and yearly variations. Therefore, light intensity data were collected directly from the Mushroom Spring location, included in the chemostat model, and then the model output was interpreted to make predictions about the inhabiting phototrophs regarding (i) the frequencies of light fluctuations that might be important in determining the fundamental light niche, (ii) the preferable intensity of light for optimal growth, and (iii) the upper light limit. The model output was then compared to a laboratory-determined light niche of a particular *Synechococcus* strain that was isolated near the surface of the Mushroom Spring mat.

Model Description

The n-species chemostat model stated below is a modification of the classic

chemostat model found throughout the literature (e.g., Smith and Waltman, 1995):

$$\frac{dS}{dt} = D(S^0 - S) - \sum_{j=1}^{N} \frac{1}{Y^{(j)}} \frac{r_s^{(j)} S}{K_s^{(j)}} x^{(j)} f^{(j)}(T(t)) \tag{2.1}$$

$$\frac{dx^{(j)}}{dt} = x^{(j)} \left(\frac{r_s^{(j)} S}{K_s^{(j)}} f^{(j)}(T(t)) - D \right).$$

Nutrient influx and dilution rate are equal and constant; nutrient, organisms, and

Table 2.1. Parameter definitions.

Quantity	Description	Units
$T(t)$	Environmental history function varying with time	varies
ω	Frequency of environmental fluctuation	1/time
$f^{(j)}(T(t))$	Fitness response function for species j	none
S^0	Incoming nutrient concentration	mass/volume
$S(t)$	Limiting nutrient concentration	mass/volume
$x^{(j)}(t)$	Biomass of species j	microbes/volume
D	Dilution rate	1/time
$K_s^{(j)}$	Half-saturation of species j	mass/volume
$r_s^{(j)}$	Maximal growth rate of species j	1/time
$Y^{(j)}$	Yield coefficient for species j	microbes/mass
t_L	Any time after chemostat steadies	time
R/ω	Duration of environmental cycle	time

byproducts flow out of the culture vessel at a constant rate and nutrient flows into the vessel at an equivalent rate. Note that standard Michaelis-Menton kinetics have been replaced by first-order kinetics; numerical results using either consumption function are qualitatively similar. Two important additional features are (i) that the chemostat is operating under a temporally varying environmental condition,

$T(t)$, and (ii) that a fitness response function, $f^{(j)}(T)$, is incorporated into the model in the growth rate and consumption terms of each species.

Single Species Model

Much of the discussion that follows is based on the one-species model

$$\begin{aligned} \frac{dS}{dt} &= D(S^0 - S) - \frac{r}{Y}\frac{S}{K_s}xf(T) \\ \frac{dx}{dt} &= x\left(\frac{rS}{K_s}f(T) - D\right). \end{aligned} \tag{2.2}$$

The solution of system (2.2) is

$$\begin{aligned} S(t) &= S^0 - \frac{K_s}{r}\frac{e^{\int_0^t (\frac{rS^0}{K_s}f(T(B))-D)dB}}{\int_0^t e^{\int_0^B (\frac{rS^0}{K_s}f(T(\gamma))-D)d\gamma} f(T(B))dB + C} \\ x(t) &= \frac{K_sY}{r}\frac{e^{\int_0^t (\frac{rS^0}{K_s}f(T(B))-D)dB}}{\int_0^t e^{\int_0^B (\frac{rS^0}{K_s}f(T(\gamma))-D)d\gamma} f(T(B))dB + C}, \end{aligned} \tag{2.3}$$

where C is an integration constant determined by initial conditions (solution details are provided in Appendix A, Section A.1). Notice that the quantity $\frac{rS^0}{K_s}f - D$ appears several times in (2.3), and this same quantity also appears as an "intrinsic rate of growth" factor in (2.2) (with S in place of S^0 in the species equation). The inverse time scale $\frac{rS^0}{K_s}\bar{f} - D$, where $\bar{f}$ is the average fitness response over one environmental cycle, will be referred to as the effective growth rate in the following sections as this quantity is a measure of the average rate of change of biomass in the chemostat over one cycle. Note, by observing the form of the solutions shown in

(2.3), that the nutrient and biomass concentrations at any time t have exponentially-decaying dependence on earlier times, that is, the chemostat system has memory built into it.

The Fitness Response Function

An important aspect of the theory described here is the optimization of the function $f(T)$, which measures the fitness of species x at each environmental condition T in the range, and which will be used to make predictions about the relationships between temporal environmental fluctuations of specified frequencies and niche character. Note that the environment itself is not a consumable entity, and the limiting nutrient S should be thought of as a chemical compound that is essential for biomass production (e.g., nitrate, phosphate, carbonate, etc.). In this regard, higher fitness at one environmental condition compared to another (a larger f at one T versus another) means the organism is producing more biomass at that particular condition, not from consuming the environment, but from consuming more of the limiting nutrient because the current state of the environment is one which is more favorable for growth.

We define optimal here to mean that a species that possesses an optimal fitness response function will survive and out-compete other species subject to the same constraints (trade-offs). Intuitively, the best competitor in a given environment should be an organism equipped with a fitness response function that results

in more biomass production, on average, than any other organism possessing any other fitness response function. Therefore, an objective function that maximizes the geometric mean (Haldane and Jayakar, 1963; Gilchrist, 1995) of the species concentration, where the mean is calculated over one environmental cycle, is formulated for this investigation. For mathematical convenience the objective function is formally stated as a minimization problem in substrate, justified by the relationship $\frac{x}{Y} + S = S^0$ (established in Appendix A.1). We have analyzed a two-species chemostat system and have shown in the low and high frequency regimes, that the species possessing the optimal fitness response function, with respect to minimizing the geometric mean of the nutrient concentration, is the only species that can invade every other species. Further support for the choice of this particular objective function is provided in Appendix A.2 through numerical competition experiments in the two-species system (Figures A.1-A.4).

Without constraint (trade-offs) the optimal fitness response would be to spend an infinite amount of effort growing at every possible environmental condition in the range, an unrealistic scenario. Rather, a constraint of constant area under the fitness curve is enforced here in order to normalize total fitness over the range of environmental conditions for all possible responses (Levins, 1968; Nagylaki, 1975; Slatkin and Lande, 1976; Lynch and Gabriel, 1987; Gilchrist, 1995). This particular constraint is often referred to as a specialist/generalist trade-off (Levins, 1968; Lynch and Gabriel, 1987; Gilchrist, 1995).

Specifically, for a given frequency of environmental fluctuation, ω, the optimization problem is to minimize the geometric mean of S, i.e., effectively to minimize

$$\int_{t_L}^{t_L+R/\omega} \ln S(t)dt \tag{2.4}$$

subject to the constraint

$$\int_{T_{\rm lo}}^{T_{\rm hi}} f(T)dT = \beta, \tag{2.5}$$

where R/ω is the duration of the environmental cycle, β is a constant, and t_L is any large time (large enough that the nutrient and species concentrations are indistinguishable from one cycle to the next). Note that $\beta = 1$ can be assumed, without loss of generality, as changing β amounts to rescaling time.

A defining characteristic of the fitness response function is its width. Let $E := \{T : f(T) > \alpha\}$, where $\alpha > 0$ is a predetermined threshold such that T values where $f(T) < \alpha$ are understood to be outside of the organism's fundamental niche, and the niche width of the organism is defined as the measure, $\mu(E)$, of the set E. If E contains only finitely many discrete values, $\mu(E) = 0$, otherwise $\mu(E)$ is the total length of the environmental intervals over which $f(T) > \alpha$. In the upcoming analyses we will demonstrate that the niche width of $f(T)$ is heavily influenced by the combination of the frequency, ω, with which the environment varies and the environmental density function, $W(T)$. $W(\tau)d\tau$ is defined as the measure of the total time in one environmental cycle that $\tau \leq T(t) \leq \tau + d\tau$, where $d\tau$ is a differential representing an infinitesimally small change in the environment. Informally,

$W(T)dT$ can be thought of as a measure of how much time the environment spends at each condition over one cycle.

Optimization of Fitness Response

Optimization results are presented for environmental fluctuations that occur with high, low, or intermediate frequency. For the limiting high and low frequency regimes (ω is significantly larger or smaller, respectively, than the effective growth rate), asymptotic approximations for the the model in (2.2) are obtained prior to the optimization step, i.e., nutrient concentration S and organismal biomass x are approximated and then a calculus of variations approach is used to calculate the optimal fitness responses. For intermediate frequencies, the optimal fitness responses are computed numerically. For illustrative purposes all results in this section are represented graphically using an assumed periodic environmental history function, $T(t) = 60 + 3\cos(\omega t)$, where ω represents the number of environmental cycles per time unit. A sinusoidal environmental history function provides a convenient way to investigate fluctuations that occur cyclically over specified time scales (e.g., daily, seasonally, and yearly). Note that the results presented in this section are intended to be general though, holding for any type of temporal environmental variation.

High Frequency Optimization Results

For high frequency fluctuations, $\omega \gg \frac{rS^0}{K_s}\bar{f} - D$, the 0th-order asymptotic approximation for nutrient concentration in the substrate is found to be

$$S_0 = \frac{DK_s}{r\bar{f}} \tag{2.6}$$
$$x_0 = Y(S^0 - S_0),$$

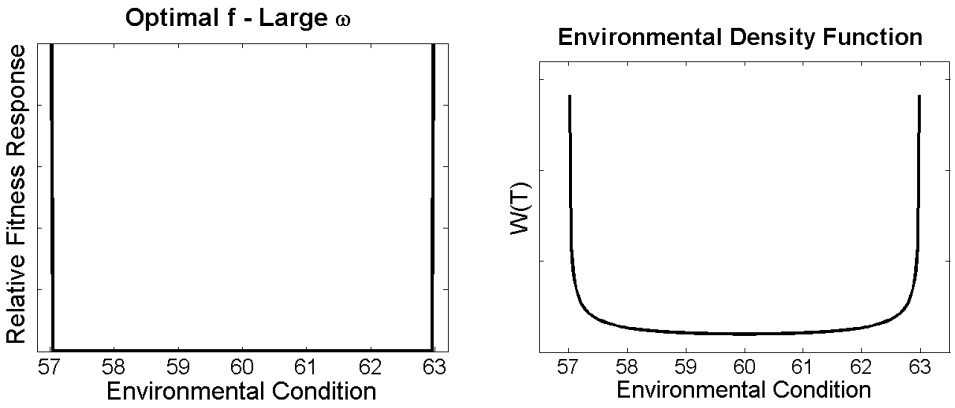
where

$$\bar{f} = \frac{\omega}{R}\int_{t_L}^{t_L+R/\omega} f(t)dt.$$

Two time-scale asymptotics (Holmes 1995) are used here, as well as in the low frequency regime, to calculate the approximations – the details are provided in Appendix A.3. The optimization problem stated in (2.4) and (2.5) is then solved with S_0 from (2.6) approximating S in the objective function. Details of the optimization calculation are included in Appendix A.4 where it is shown that an optimal f is

$$f(T) = \frac{1}{N}\sum_{\tau \in M}\delta(T-\tau), \tag{2.7}$$

where $M := \{\tau \in [T_{\text{lo}}, T_{\text{hi}}] : W(\tau) \geq W(T)\ \forall\ T \in [T_{\text{lo}}, T_{\text{hi}}]\}$, N is the number of elements in M, and δ is the Dirac delta function. Recall that $T(t)$ is the environmental history function and $W(T)$ is the corresponding environmental density function, as defined at the end of previous section. Note that the optimal fitness response for high-frequency environmental fluctuations depends exclusively on the environmental density function.



(A) (B)

Figure 2.1. (A) Optimal fitness response function for high frequency environmental oscillations. $T(t) = 60 + 3\cos(\omega t)$ in this calculation and the measurement units are arbitrary. (B) Environmental density function $W(T)$ for this computation.

As a simple example, for environmental history $T(t) = 60 + 3\cos(\omega t)$, Figure 2.1A shows an optimal f. Note that the time series $T(t)$ used for the computation in Figure 2.1A is atypical in that it results in two modes and hence the resulting f can have two delta functions. More generally, the model prediction is that in the high-frequency regime all growth should occur at the environmental condition(s) that are most common, i.e., at the mode(s) of $W(T)$ (see plot of $W(T)$ in Figure 2.1B). In other words, an extreme specialist(s) with effectively a niche width of zero is optimal, suggesting that high-frequency environmental fluctuations do not play a significant role in determining fundamental niche character – the environment simply changes before the organism can react.

Low Frequency Optimization

For the low frequency case, $\omega \ll \frac{rS^0}{K_s}\bar{f} - D$, the 0th-order asymptotic approximation is

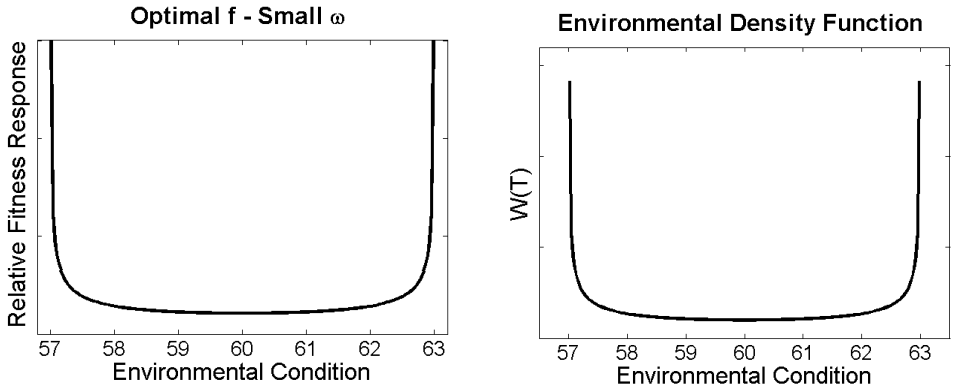
$$\begin{aligned} S_0(t) &= \frac{DK_s}{rf(t)} \\ x_0(t) &= Y\left(S^0 - \frac{DK_s}{rf(t)}\right). \end{aligned} \tag{2.8}$$

Approximation details and a plot comparing the solution from the ordinary differential equations (ODE) in (2.2) and the 0th-order approximation from (2.8) are provided in Appendix A.3, Figure A.5. The optimization problem, this time with S_0 from (2.8) replacing S in the objective function from (2.4), results in the solution

$$f(T) = \frac{W(T)}{\int_{T_{\text{lo}}}^{T_{\text{hi}}} W(T)dT}. \tag{2.9}$$

Calculation details are provided in Appendix A.4. Observe, as in the high frequency case, that the optimal f is dependent only on the environmental density function $W(T)$; in fact, f in this case is a scaled version of W. For the case that $T(t) = 60 + 3\cos(\omega t)$ a plot of the optimal f is shown in Figure 2.2A. Unlike the high frequency case, here there is a range of environmental conditions for which the organism is well-suited to grow. In other words, low frequency fluctuations in the environment lead to the formation of a fundamental niche that has a clearly defined, non-zero niche width that reflects only the fraction of time the environment spends at each condition (compare Figure 2.2A and Figure 2.2B). Note that for this

example the symmetric nature of the supplied environmental history yields two conditions that occur equally most often (see Figure 2.2B), possibly suggesting that two species may be the way to optimally utilize the available nutrient resources for this particular choice of $T(t)$. This is an example where environmental variation may lead to the coexistence of two interchangeable species in a setting with only a single limiting nutrient (for example see Stewart and Levin 1973). Alternatively, this result can also be interpreted as phenotypic variance within a functional group being the optimal strategy in a fluctuating environment (Norberg et al., 2001).



(A) (B)

Figure 2.2. (A) Optimal fitness response function for low frequency environmental oscillations. $T(t) = 60 + 3\cos(\omega t)$ in this calculation. (B) Environmental density function $W(T)$ for the environmental history in (A).

Intermediate Frequency Optimization

Asymptotics for equation (2.2) in the ω-limiting cases allowed for analytic approximations of the proposed constrained optimization problem, but for the intermediate frequency results presented in this section, numerical solution is required. Computations use a periodic environmental history function, with period R/ω, and, as before, the geometric mean of the nutrient concentration is minimized over one period after the chemostat had been operating for t_L time units. Solutions were obtained using a sequential quadratic programming (SQP) technique (Powell, 1978), which is a generalization of Newton's method (details in Appendix A.4), and are shown in Figure 2.3, where $T(t) = 60 + 3\cos\omega t$ is again the environmental history function. Figure 2.3 shows the transition of the optimal fitness response from the low frequency regime to the high frequency regime. Observe that intermediate fluctuations in the environment, as was the case in the low frequency regime, result in an optimal fitness response that is non-zero over a range of values, i.e., intermediate frequencies of temporal fluctuations can also play a role in characterizing the fundamental niche.

In the high and low frequency cases, results depend on environmental history only through the environmental density function. To test whether the same is true at intermediate frequencies, numerical solutions were computed using two different environmental functions (T_1 and T_2 shown in Figure 2.4A) that share the same density function ($W(T)$ is shown in Figure 2.4B; note that it is the same density

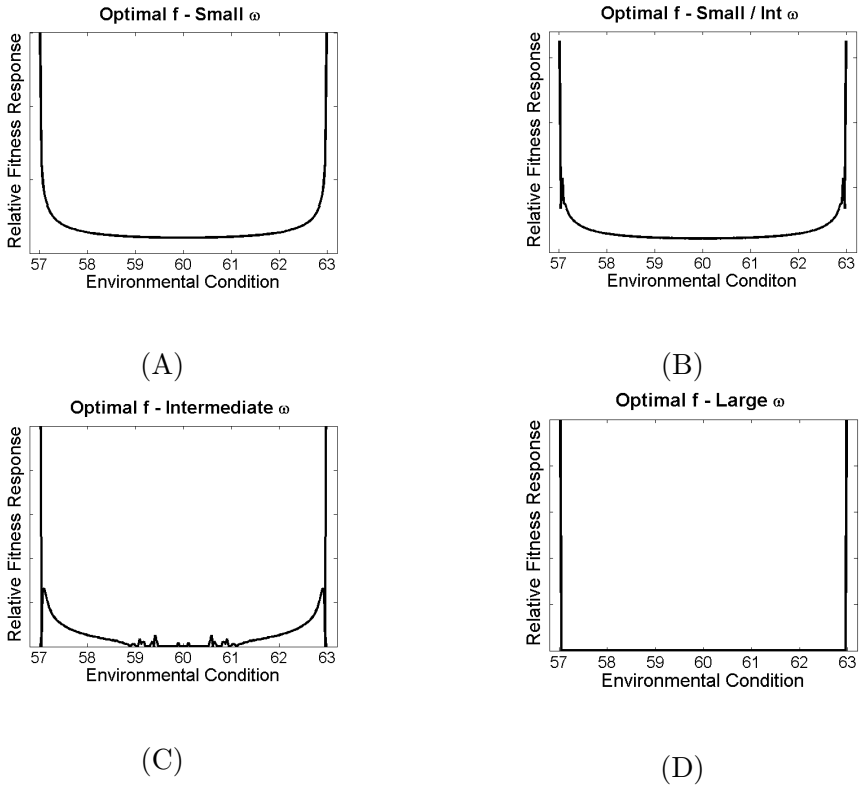
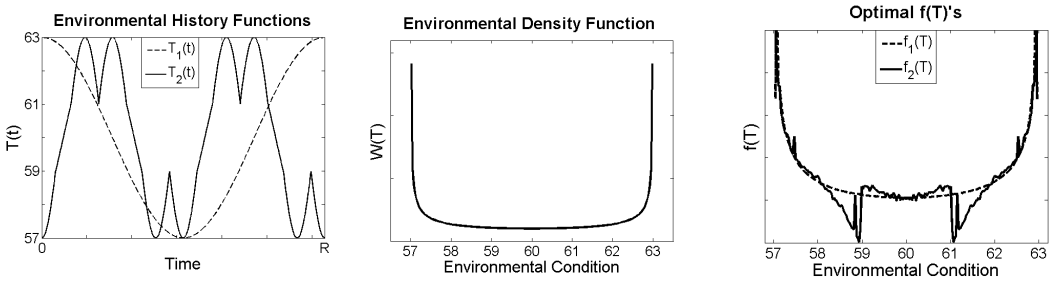


Figure 2.3. (A)-(D) show numerical solutions for several values of ω, where ω increases from very small to very large (with respect to the effective growth rate) as one moves from left to right and top to bottom; the ratios of effective growth rate to ω are 100, 10, 1, .1 respectively. The environmental range is split into $n = 200$ intervals in each plot and $T(t) = 60 + 3\cos(\omega t)$. For $n > 200$ the qualitative nature of the optimal fitness response does not change.

function as in Figure 2.1B and Figure 2.2B). See Appendix A.5 for the definition of $T_2(t)$. In Figure 2.4C the optimal fitness responses (f_1 and f_2 corresponding to T_1 and T_2, respectively) for a specific frequency of intermediate environmental fluctuation (the ratio of the effective growth rate to ω is 10 in this case) are plotted and reveal an interesting observation. Two different optimal fitness responses were obtained for environmental histories with identical density functions, showing that



(A) (B) (C)

Figure 2.4. (A) $T_1(t)$ (dashed) and $T_2(t)$ (solid) that have the same $W(T)$. (B) $W(T)$ for $T_1(t)$ and $T_2(t)$. (C) Optimal fitness responses $f_1(T)$ (dashed) and $f_2(T)$ (solid) corresponding to $T_1(t)$ and $T_2(t)$ for the case when the ratio of the effective growth rate of the organism to the frequency of the environmental fluctuation is 10.

for intermediate cases the optimal fitness response is dependent not only on $W(T)$, but also the form of $T(t)$, i.e., not only the density but also the order in which the environmental conditions are encountered matters. Low and high frequency asymptotics were confirmed numerically when the same optimal $f(T)$ was computed for both T_1 and T_2 (Figure 2.4A) in the limiting cases. Note that for intermediate frequencies, and only for intermediate frequencies, the optimal fitness f can depend on the full environmental history $T(t)$ (not just the density $W(T)$). Said another way, at intermediate frequencies, optimal response depends on the environmental schedule so that, effectively, clocking behavior is being predicted as an emergent behavior in response to environmental variability. As an example, in the particular case illustrated in Figure 2.4, observe that both optimal fitness responses in Figure 2.4C behave in a manner similar to that of the density function in Figure 2.4B, with

the exception of the noticeable intervals of sudden inhibited growth. Immediately after the most common conditions, 57 and 63 for this example, f_1 has an interval of inhibited growth, albeit briefly. On the other hand, f_2 has intervals of subdued growth at 59 and then again at 61, where it should be observed that the steepness of slope of T_2 is much less severe between 59 and 61 compared to when the environment is between 57 and 59. We suggest that this behavior may be due to recovery time and/or preparation time (recall that the chemostat system has built-in memory). Recovery time is the time it takes, after a period of heavy nutrient consumption, for the chemostat to replenish nutrient concentrations to a level that can again support growth. Preparation time is defined as time the organism is exerting minimal effort to grow in order to conserve nutrient for when conditions are more favorable, and the predictability of the assumed periodic environment may allow the organism to know what is coming next and respond by allocating time to prepare. Either way, a biological clock has been introduced. Of course, in reality, the on-off behavior from a recovery/preparation period would likely have a cost penalty and so might be less likely to occur in actuality. Such a cost is not included in the present model; including this cost might, in some cases, make introduction of a clock less efficient.

In summary, as in the limiting frequency cases, the optimal fitness responses for intermediately fluctuating environments (with respect to the effective growth rate) appear to be influenced by the density function, but unlike the limiting cases, the optimal behavior in the intermediate regime also appears to have some additional

dependence on the form of $T(t)$ (e.g., Figure 2.4C). With intermediate frequencies there is, on the one hand, time to react to environmental changes but, on the other, not so much time that the system can effectively forget previous environments. For such instances a circadian clock-like mechanism may be beneficial. It has already been documented that clocking behaviors are "widely believed to facilitate adaptation to the environment (Pittendrigh, 1993; Kondo and Ishiura, 2000)," and the behavior observed in the numerical output shown in Figure 2.4C, in combination with the results from the high and low frequency regimes, suggests that the existence of clock-like mechanisms in certain organisms may have evolved as a response to intermediate fluctuations in the environment.

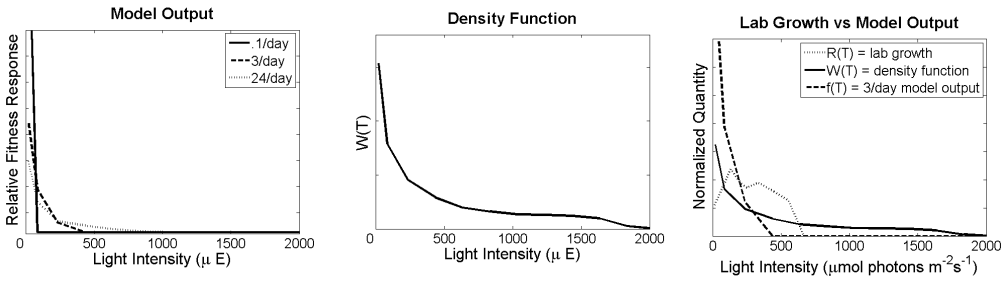
Results From Field Study

To test model predictions we compare model output to laboratory measurements of the fundamental light niche of a particular phototrophic organism isolated from Mushroom Spring, YNP. Over a two-year period, from 13 January 2011 to 13 January 2013, light intensity data were collected (Appendix Figure A.6; only one year shown) near the source pool of Mushroom Spring by averaging light measurements that were taken every five minutes over each 60 minute period. Light was measured with an LI-190SA quantum sensor, which measures photosynthetic active radiation (PAR) in the 400 - 700 nm waveband. Measurement units are μmol photons $m^{-2}s^{-1}$

and data were recorded with the LI-1400 Datalogger (all light equipment from LI-COR (Lincoln, Nebr)). The light data were then input to the model in the form of an environmental history function $T(t)$, and numerical simulations were performed to compute the optimal fitness response with respect to the *in situ* light conditions. From 11 August 2011 to 14 September 2011 the datalogger had a battery malfunction and from 2 November 2011 to 15 December 2011 the light sensor became buried in the snow. Therefore, only results for the year of data that was uninterrupted (13 January 2012 to 13 January 2013) are shown, though when both years of data were input into the model, with missing data from 2011 approximated with data from 2012, the results were nearly identical.

The light data from 13 January 2012 to 13 January 2013 were cycled for 10 simulation years (with all dark times omitted) and the optimal fitness responses were computed by minimizing the geometric mean of the nutrient concentration over year ten. Figure 2.5A shows the model output for the optimal fitness responses plotted for various supposed effective growth rates (in doublings/day). In relation to the earlier discussion, supposing a large effective growth rate is effectively the same as setting a low-frequency environmental variation (substantial growth occurs before the environment changes significantly), whereas a small effective growth rate is effectively the same as high frequency environmental variation (significant environmental change occurs before substantial growth occurs). The model predicts that if the effective doubling rate falls below approximately 0.1 doublings per/day (solid

curve in Figure 2.5A) the optimal fitness response of the organism is to grow only at the times when the most common light intensity occurs. If the effective growth rate is 1 doubling per/day or higher (dashed and dotted curves in Figure 2.5A), the organism's optimal behavior is still to focus its growing efforts at the more common light intensities, but also to spend some effort growing over a range of other light intensities. The model output in Figure 2.5A agrees with the earlier results in the



(A) (B) (C)

Figure 2.5. (A) Model output for a low number of doublings/day, 0.1/day (solid), intermediate number, 3/day (dashed), and for a high number, 24/day (dotted). (B) Environmental density function for empirical light data. (C) Growth of *Synechococcus* strain 65AY6-Li with respect to light ($R(T)$, dotted), density of empirical light data (W(T), solid), and the model prediction ($f(T)$, dashed). $R(T)$ has units of doublings/day, whereas $W(T)$ and $f(T)$ are unit-less quantities providing qualitative information, therefore, for comparison purposes, the three curves were normalized with respect to area under the curve. The growth experiment was performed in duplicate and the average growth rate of the duplicates during exponential growth phase is shown in the figure. Any light condition that resulted in an exponential growth phase of less than 48 hours was recorded as zero growth.

sense that the environmental density function shown in Figure 2.5B appears to play an important role in determining the optimal fitness responses. Note the empirical

light density function is quite different from the bimodal density function of the periodic sinusoidal example (Figure 2.4B); seasonal fluctuations in light intensity produce a unimodal distribution as the lower light intensities are the only intensities that occur year-round.

Next we compared the optimal fitness responses from Figure 2.5A to the laboratory-determined light adaptation of a phototrophic *Synechococcus* strain (65AY6Li) that was isolated from the top layers of the Mushroom Spring microbial mat and is representative of an abundant surface population (Becraft et al., 2011), designated *psaA* ecotype A1. A surface population was chosen for comparison because the light data included in the model were collected near the surface of the microbial mat. Note that a chemostat would be expected to be a better modeling platform for surface populations versus subsurface populations. The strain was isolated using a dilution and filtration protocol (Allewalt et al., 2006), and originated from an isolated microcolony from a filter on a high-dilution plate. The colony was then suspended in liquid growth medium and grown at 52°C in approximately 50 μmol photons $m^{-2}s^{-1}$ of white fluorescent light. DNA extraction using a bead beating protocol and Titanium-454 pyrosequencing (Research and Testing Lab, Lubbock, TX) were used to confirm the genotype. A phenotypic analysis was then performed with respect to downwelling irradiance by growing the isolates in batch cultures at 10 different light intensities (14, 56, 139, 200, 250, 333, 444, 556, 667, 770; all units are μmol photons $m^{-2}s^{-1}$), under continuous light, in duplicate, and with a starting

density of 10^6 cells/mL. The batch cultures were grown in between two light sources of equal intensity and spectral composition in order to minimize spectral changes that occur with depth in a medium. Samples were collected every twelve hours and counted with a BD-FACSCanto flow cytometer (BD Biosciences (San Jose, California)). Growth rates were determined by fitting a log-linear regression line to the exponential growth phase and calculating the slope.

It has been empirically shown (Allewalt et al., 2006; Kilian et al., 2007; Chapters 4 and 5) that *Synechococcus* spp. from this system characteristically have a growth rate of approximately 1-3 doublings/day under various laboratory conditions. Figure 2.5C shows a comparison of 3 doublings/day model predictions versus measured data from the growth experiment. Light intensity at the field site reaches downwelling irradiances of approximately 2000 μmol photons $m^{-2}s^{-1}$ in the summer months (Appendix Figure A.6), yet under the laboratory conditions imposed in the empirical growth study, the organisms were unable to grow at such extremities, and the model results support these empirical findings. The model results can be explained by the relatively infrequent occurrence of high light intensities over the course of a year. Light intensities over 1000 μmol photons $m^{-2}s^{-1}$ are only measured during the months of March through September, and at most for only a few hours each day. From experiments with laboratory strains of *Synechococcus* spp. closely related to the strain discussed here, it has been observed that under continuous

high-light conditions the cultures can only be sustained for short periods of time before perishing, possibly due to phototoxicity (Kilian et al., 2007; Chapters 4 and 5). This behaviorial response suggests that these organisms may have adapted certain mechanisms that enable them to withstand short periods of high light (such as a few hours per day) and the growth-inhibiting consequences that come with them (such as CO_2 limitation and the harmful effects of reactive oxygen species (Jensen et al., 2010)). This observation also provides additional support that prolonged exposure to an environmental condition, such as a certain light regime, may be necessary in order to have a niche-characterizing effect (e.g., conditions that change slowly or intermediately with respect to the effective growth rate are important, such as the day/night cycle in this case, and conditions that change quickly are not, such as light intensity on a partially cloudy day). System behavior is also consistent with clock predictions at intermediate frequencies, specifically, it should be noted that the *Synechococcus* spp. from Mushroom Spring are known to possess circadian clock genes (Baca et al., 2010; M. Olsen, unpublished data) and operate on a diel cycle, metabolizing CO_2 through photosynthesis during the day, and fermenting and fixing N_2 at night (Steunou et al., 2006). Furthermore, the model results and the results from the experiment both suggest optimal growth to occur during low light conditions and that the upper limit of sustained growth is in the $500 - 600$ μmol photons $m^{-2}s^{-1}$ range, which happens to be approximately equal to the maximum light intensity in December (the time of the year when light is most limited).

Differences in details in Figure 2.5C between the model predictions and the laboratory results are to be expected since only time dependence of the light data was included in the model. In particular, the model predicts a spike in $f(T)$ at a lower light intensity than the results from the growth experiments suggest, probably as a consequence of the fact that the model does not consider the growth limiting consequences of low photon levels that occur during extremely low light conditions. That is, in reality growth must stop when light intensity does. There are other physical effects that have been neglected as well: carbon dioxide levels (which have been shown to have a major effect on light tolerances of this species (Chapter 4)), spectra of light (the white fluorescent light spectrum used in the laboratory is not an exact match to the sunlight spectrum (Reese et al., 2011)), and temperature (the strain shown in Figure 2.5C was isolated from a sample collected at 65°C, but was grown at 52°C in the lab) are other variables that may play a role in determining the fundamental light niche of *Synechococcus* spp. in Mushroom Spring, but were not included in the model. This is deliberate; the goal here is to explore the consequences of temporal variation, in and of itself, on niche character. Nevertheless, the match between the predicted and measured niche is striking, suggesting that the fundamental light niche of these organisms is strongly influenced by the daily fluctuations in light.

Discussion/Conclusion

Our aim is to study the effects temporal environmental variations have on optimal behavior by combining theoretical and empirical approaches. First, a mathematical model was constructed and analyzed to develop a theoretical framework; second, light data and organisms were collected from a specific field site, the light data were incorporated into the model, the organisms were grown in the laboratory, and then model predictions were compared to the growth rate of the organisms. A significant feature in the modeling approach that we presented was that an arbitrary initial form for the niche function was not imposed (as done in other related literature, e.g., Levins, 1968; Slatkin and Lande, 1976; Lynch and Gabriel, 1987; Huey and Kingsolver, 1993; Lynch and Lande, 1993; Gilchrist, 1995, 2000; Norberg et al., 2001; Gomulkiewicz and Houle, 2009; Chevin et al., 2010); rather, all results emerged from environmental information alone, that is, the niche function was allowed to determine itself in response to environmental variation. As a result, we are able to offer explanations for certain phenomena that would not be evident if fixed-form niche functions are used.

The results from the model analyses support and extend comments made by Hutchinson (1961) several decades ago: some frequencies of temporally varying environmental fluctuations (relative to the effective growth rate of the organism) appear to play a role in structuring the fundamental niche (low and intermediate),

and others do not (high). For example, for a microorganism with an intrinsic doubling time on the order of once per day, high frequency fluctuations in light that occur on a partially cloudy day may not have a niche-structuring role; on the other hand, the less frequently repeating day/night and seasonal cycles may be very important. Further, results from the intermediate frequency regime suggest that it is exactly these frequencies that could have contributed to the emergence of clock-like mechanisms that are found in organisms of all complexities (Kondo and Ishiura, 2000). It is only at these intermediate times, where the time scale of fluctuation is comparable to the time scale of growth, that it is possible for organisms to couple to environmental variation. At higher frequencies, organisms do not have time to respond, while at lower frequencies, organisms are, rather, not able to wait long enough for the environment to respond.

This study is most closely related to those of Levins (1968), Lynch and Gabriel (1987), and Gilchrist (1995) in the sense that an optimization strategy to predict fundamental niche character in a variable environment was used in each case. In comparison to their results, we predicted that maximum fitness would occur at the most common condition; this agrees with Levins' (1968) results but disagrees with Lynch and Gabriel (1987), as they found maximum fitness would evolve to the average condition after several environmental cycles. Levins (1968) also predicted the optimal strategy in a fine-grained environment (high frequency fluctuations) would be a specialist and in a coarse-grained environment (low frequency fluctuations)

would be a generalist. Later, Gilchrist's (1995) work produced similar results to those of Levins', that is, significant within-generation temporal variation (high frequency fluctuations) and constant environments both favor specialists, while among-generation temporal variation (slow to intermediate fluctuations) favors generalists. Alternatively, Lynch and Gabriel (1987) predicted that any temporal variance in the environment selects for a wider fundamental niche than a constant environment would select for, which agrees with what we predicted in some cases but not all – particularly the limiting high frequency case where we predicted that the optimal fundamental niche is not determined by temporal variation. If put into the specialist/generalist context of Levins and Gilchrist, the predictions presented here agree with their results, noting once again that we did not assume a specific form for the niche function as they did. But by allowing a relaxed form for the niche function we were able to go beyond these previous studies in significant ways. In particular, we were able to identify possible ecological instances where speciation may arise (see Figure 2.3 and Figure 2.4C), such as when the most common environmental condition is not unique, or when the environmental history is multi-modal over one cycle. More stikingly, the relaxed niche form also lead to the hypothesis that it is exactly intermediate environmental fluctuations, such as the diel cycle of light, that may lead to the development of clock-like mechanisms.

In summary, we present a model that strips out influences other than environmental fluctuation in and of itself. The resulting predictions are, first, that when the

frequency of environmental fluctuation is higher than the doubling rate of the inhabiting organism, the organism does not have time to react and hence the optimal behavior is to spend all growing effort at the most common condition(s). Second, when the frequency of the environmental fluctuation is lower than the growth rate the system equilibrates, the organism reacts, and the optimal response is to partition growing efforts over the range of conditions in a way that is proportional to the environmental density function. In fact, the model predicts that the width and amplitude of the fundamental niche will scale with the width and amplitude of the environmental density function in the low-frequency regime. Third, in the intermediate frequency regime the environment varies on a time scale comparable to that of the doubling rate of the organism, and the model predicts this compatibility enables the organism to interact and ultimately conform to its environment. In conclusion, the results presented here suggest that temporal environmental variation, and more specifically, the frequency of that variation, may have important and far-reaching ecological consequences.

CHAPTER 3

COMPETITIVE ABILITY OF THE OPTIMAL SPECIES

Introduction

In Chapter 2 a constrained optimization procedure was applied to a temporally varying, one-species chemostat model in order to calculate the optimal fitness response of the inhabiting organism. A critical step in any optimization procedure is selecting the objective function, and in Chapter 2, the geometric mean of substrate over one environmental period was minimized (Haldane and Jayakar, 1963; Gilchrist, 1995). Recall that for this study the ideal optimal species is one that would outcompete any other species in a two-species chemostat system. In Appendix A numerical evidence suggested that the geometric mean did indeed produce the optimal species, but numerical calculations can only provide a glimpse of the true system behavior. Here, the competitive ability of the hypothesized optimal species, as so determined by minimizing the geometric mean of substrate, is investigated via a non-autonomous stability analysis on a two-species chemostat system. Specifically I ask: how optimal is the optimal species?

The theory of competitive exclusion (Gause, 1934) states that, at equilibrium, two species cannot coexist on a single limiting nutrient, i.e., only one species can occupy a niche. Temporal variability has been shown to introduce extra niche space

and allow for coexistence if the two species possess certain trade-offs (Smith and Waltman, 1995; Lenas and Pavlou, 1995; Wolkowicz and Zhao, 1998; Litchman and Klausmeier, 2001; Descamps-Julien and Gonzalez, 2005). For example, if the species have complementary fitness responses over different temperature ranges, and the temperature is fluctuating with respect to time, then there are regions of parameter space where coexistence can occur (Descamps-Julien and Gonzalez, 2005).

Invasibility has been proposed as a necessary condition for coexistence in a homogeneous spatial environment that is experiencing temporal variation, i.e., each species must be able to invade the other at different times (Smith and Waltman, 1995; Siepielski and McPeek, 2010 and the references within). Here, the invasibility properties of the hypothesized optimal species in a temporally varying two-species chemostat system are studied from both directions. That is, if only the hypothesized optimal species is in the chemostat and then another species is added at a low concentration, can the optimal species always fight off invasion? And vice versa, if only a sub-optimal species is in the system and then a relatively low concentration of the hypothesized optimal species is added, can the optimal species always invade? In the final section of this chapter a two-species chemostat model is numerically investigated to study the interaction between an invader species and the resident species after invasion has occurred. The purpose of this investigation is to provide a hypothetical example of how an arbitrary species may evolve to the hypothesized optimal species through a series of mutation and periodic selection events.

Two Species Chemostat Model

The two species chemostat model with temporally varying environmental input is

$$\begin{aligned}
\frac{dS}{dt} &= D(S^0 - S) - \frac{r}{K_s Y} S\left(x_1 f_1(T(t)) + x_2 f_2(T(t))\right) \\
\frac{dx_1}{dt} &= x_1\left(\frac{rS}{K_s} f_1(T(t)) - D\right) \\
\frac{dx_2}{dt} &= x_2\left(\frac{rS}{K_s} f_2(T(t)) - D\right),
\end{aligned} \tag{3.1}$$

where f_1 and f_2 are the fitness responses of the two species, and the other parameters are defined in Table 2.1. The system is reduced to a more mathematically tractable planar system by adding the first equation in (3.1) to $\frac{1}{Y}$ times the sum of the second and third equations. By applying the same approach followed in Appendix A.1, after $O(D^{-1})$ time, it can be shown that to an exponentially small approximation

$$S = S^0 - \frac{1}{Y}\left(x_1 + x_2\right). \tag{3.2}$$

Next, substituting (3.2) into the second and third equations of (3.1), and disregarding the first equation, yields the planar system

$$\begin{aligned}
\frac{dx_1}{dt} &= x_1\left(\frac{r}{K_s}\left(S^0 - \frac{1}{Y}\left(x_1 + x_2\right)\right) f_1 - D\right) \\
\frac{dx_2}{dt} &= x_2\left(\frac{r}{K_s}\left(S^0 - \frac{1}{Y}\left(x_1 + x_2\right)\right) f_2 - D\right).
\end{aligned} \tag{3.3}$$

Linearization About a Known Solution

The goal of this chapter is to investigate the invasibility properties of the optimal species that was determined by minimizing the geometric mean of substrate in the one-species model. To do so, I will utilize the fact that the nutrient and species solutions to the one-species system also solve equations (3.3), with the second species being identically zero. Invasibility can then be investigated by perturbing these solutions and studying the long term behavior of (3.3). This motivates the following calculation, that starts by linearizing (3.3) about a general known solution of the form $(\bar{x}_1, \bar{x}_2)$. Let

$$\begin{aligned} x_1 &= \bar{x}_1 + \tilde{x}_1 \\ x_2 &= \bar{x}_2 + \tilde{x}_2 \\ S &= \bar{S}_1 + \tilde{S}_1, \end{aligned} \tag{3.4}$$

where $(\bar{x}_1, \bar{x}_2)$ is a known solution of (3.3), $(\bar{S}, \bar{x}_1, \bar{x}_2)$ is a known solution to (3.1), and $\tilde{x}_1$, $\tilde{x}_2$, and $\tilde{S}_1$ are the species and nutrient perturbations applied to the known solutions, respectively. To derive the perturbation differential equations, the perturbed species equations given in (3.4) are substituted into (3.3) and rearranged as shown here.

$$
\begin{aligned}
\dot{x_1} &= \dot{\bar{x}}_1 + \dot{\tilde{x}}_1 \qquad (3.5)\\
&= (\bar{x}_1 + \tilde{x}_1)\left(\frac{r}{K_s}\left(S^0 - \frac{1}{Y}(\bar{x}_1 + \bar{x}_2 + \tilde{x}_1 + \tilde{x}_2)\right) f_1 - D\right)\\
&= \bar{x}_1\left(\frac{r}{K_s}\left(S^0 - \frac{1}{Y}(\bar{x}_1 + \bar{x}_2)\right) f_j - D\right) - \bar{x}_1 \frac{r}{K_s Y}(\tilde{x}_1 + \tilde{x}_2) f_1 +\\
&\quad \tilde{x}_1\left(\frac{r}{K_s}\left(S^0 - \frac{1}{Y}(\bar{x}_1 + \bar{x}_2)\right) f_1 - D\right) - \tilde{x}_1 \frac{r}{K_s}(\tilde{x}_1 + \tilde{x}_2) f_1\\
\dot{\bar{x}}_1 + \dot{\tilde{x}}_1 &= \dot{\bar{x}}_1 - \bar{x}_1 \frac{r}{K_s Y}(\tilde{x}_1 + \tilde{x}_2) f_1 +\\
&\quad \tilde{x_1}\left(\frac{r}{K_s}\left(S^0 - \frac{1}{Y}(\bar{x}_1 + \bar{x}_2)\right) f_1 - D\right) - \tilde{x}_1 \frac{r}{K_s}(\tilde{x}_1 + \tilde{x}_2) f_1.
\end{aligned}
$$

Solving for $\dot{\tilde{x}}_1$ results in

$$
\dot{\tilde{x}}_1 = \frac{r}{K_s}\left(S^0 - \frac{1}{Y}(2\bar{x}_1 + \bar{x}_2) f_1 - D\right)\tilde{x}_1 - \frac{r}{K_s Y} f_1 \bar{x}_1 \tilde{x}_2 - \frac{r}{K_s Y} f_1 \tilde{x}_1 (\tilde{x}_1 + \tilde{x}_2),
$$

and, similarly,

$$
\dot{\tilde{x}}_2 = -\frac{r}{K_s Y} f_2 \bar{x}_2 \tilde{x}_1 + \frac{r}{K_s}\left(S^0 - \frac{1}{Y}(\bar{x}_1 + 2\bar{x}_2) f_2 - D\right)\tilde{x}_2 - \frac{r}{K_s Y} f_2 \tilde{x}_2 (\tilde{x}_1 + \tilde{x}_2).
$$

Thus

$$
\begin{aligned}
\begin{pmatrix} \dot{\tilde{x}}_1 \\ \dot{\tilde{x}}_2 \end{pmatrix} &= \begin{pmatrix} -D + \frac{r}{K_s} f_1 \left(S^0 - \frac{1}{Y}(2\bar{x_1} + \bar{x_2})\right) & -\frac{r}{K_s Y} f_1 \bar{x_1} \\ -\frac{r}{K_s Y} f_2 \bar{x_2} - D & +\frac{r}{K_s} f_2 \left(S^0 - \frac{1}{Y}(\bar{x_1} + 2\bar{x_2})\right) \end{pmatrix} \qquad (3.6)\\
&\quad \cdot \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} - \frac{r}{K_s Y}(\tilde{x}_1 + \tilde{x}_2) \begin{pmatrix} f_1 \tilde{x}_1 \\ f_2 \tilde{x}_2 \end{pmatrix}.
\end{aligned}
$$

Again, since invasibility properties are of primary interest, only known boundary solutions of the form $(y(t), 0)$ and $(0, y(t))$ will be considered here, where $y(t)$ is the

species solution from the one-species system. The asymptotic approximations for the limiting low- and high-frequency regimes, that were calculated in Chapter 2, will be used to approximate $\bar{S}_1$ and $y(t)$ in the following analyses.

Stability Analysis

A non-autonomous stability analysis, taken from dynamical systems and Floquet theory, is now applied (Chicone, 2006; Rasmussen, 2007) to (3.6). Let $(\bar{x}_1, \bar{x}_2) = (y(t), 0)$. Then (3.6) can be written as

$$\begin{aligned}\begin{pmatrix}\dot{\tilde{x}}_1\\ \dot{\tilde{x}}_2\end{pmatrix} &= \begin{pmatrix}-D + \frac{r}{K_s} f_1\left(S^0 - \frac{2y(t)}{Y}\right) & -\frac{r}{K_s Y} f_1 y(t)\\ 0 & -D + \frac{r}{K_s} f_2\left(S^0 - \frac{y(t)}{Y}\right)\end{pmatrix}\begin{pmatrix}\tilde{x}_1\\ \tilde{x}_2\end{pmatrix} \qquad (3.7)\\ &\quad - \frac{r}{K_s Y}\left(\tilde{x}_1 + \tilde{x}_2\right)\begin{pmatrix}f_1\tilde{x}_1\\ f_2\tilde{x}_2\end{pmatrix}\\ &= \mathbf{A}(t)\begin{pmatrix}\tilde{x}_1\\ \tilde{x}_2\end{pmatrix} + \mathbf{p}\left(t, \begin{pmatrix}\tilde{x}_1\\ \tilde{x}_2\end{pmatrix}\right). \qquad (3.8)\end{aligned}$$

The first step in the stability analysis is to construct a fundamental matrix solution of the variational equation

$$\begin{pmatrix}\dot{\tilde{x}}_1\\ \dot{\tilde{x}}_2\end{pmatrix} = \mathbf{A}(t)\begin{pmatrix}\tilde{x}_1\\ \tilde{x}_2\end{pmatrix}. \qquad (3.9)$$

The upper triangular form of $\mathbf{A}(t)$ allows decoupling of the variables, and (3.9) can be solved explicitly, yielding

$$\tilde{x}_1 = e^{\int_0^t \left[-D+\frac{r}{K_s} f_1\left(S^0-\frac{2y(B)}{Y}\right)\right]dB}$$
$$\cdot \left[-\frac{rC_2}{K_sY}\int_0^t \left(e^{-\frac{r}{K_s}\int_0^B \left[f_1\left(S^0-\frac{2y(\alpha)}{Y}\right)-f_2\left(S^0-\frac{y(\alpha)}{Y}\right)\right]d\alpha} f_1 y(B)\right) dB + C_1\right]$$
$$\tilde{x}_2 = C_2 e^{\int_0^t \left[\frac{r}{K_s} f_2\left(S^0-\frac{y(B)}{Y}\right)-D\right]dB},$$

with C_1 and C_2 constants determined by initial conditions. The fundamental matrix solution $\mathbf{\Phi}(t)$ of (3.9) is then constructed by solving (3.9) as two initial value problems with initial conditions $\tilde{\mathbf{x}}(0) = \begin{pmatrix}1 & 0\end{pmatrix}^T$ and $\tilde{\mathbf{x}}(0) = \begin{pmatrix}0 & 1\end{pmatrix}^T$, respectively, yielding

$$\mathbf{\Phi_{1,1}}(t) = e^{\int_0^t \left[-D+\frac{r}{K_s} f_1\left(S^0-\frac{2y(B)}{Y}\right)\right]dB} \tag{3.10}$$
$$\mathbf{\Phi_{1,2}}(t) = e^{\int_0^t \left(-D+\frac{r}{K_s} f_1\left(S^0-\frac{2y(B)}{Y}\right)\right)dB}$$
$$\cdot \left[-\frac{r}{K_sY}\int_0^t \left(e^{-\frac{r}{K_s}\int_0^B \left[f_1\left(S^0-\frac{2y(\alpha)}{Y}\right)-f_2\left(S^0-\frac{y(\alpha)}{Y}\right)\right]d\alpha} f_1 y(B)\right) dB\right]$$
$$\mathbf{\Phi_{2,1}}(t) = 0$$
$$\mathbf{\Phi_{2,2}}(t) = e^{\int_0^t \left(\frac{r}{K_s} f_2\left(S^0-\frac{y(B)}{Y}\right)-D\right)dB}.$$

The monodromy matrix, $\mathbf{M}(t) := \mathbf{\Phi}^{-1}(0)\mathbf{\Phi}(Q)$, where Q is the period of $\mathbf{A}$, is then formulated. $\mathbf{\Phi}^{-1}(0) = I$ so the characteristic multipliers of the monodromy matrix are simply the eigenvalues, E_1 and E_2, of $\mathbf{\Phi}(Q)$. When $Q = \dfrac{2\pi}{\omega}$

$$\begin{pmatrix}E_1\\E_2\end{pmatrix} = \begin{pmatrix}e^{\int_0^{2\pi/\omega}\left[-D+\frac{r}{K_s} f_1\left(S^0-\frac{2y(t)}{Y}\right)\right]dt}\\ e^{\int_0^{2\pi/\omega}\left[-D+\frac{r}{K_s} f_2\left(S^0-\frac{y(t)}{Y}\right)\right]dt}\end{pmatrix}. \tag{3.11}$$

The exponents of E_1 and E_2 are referred to as Floquet exponents, denoted by λ_1

and λ_2, and they are

$$\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix} = \begin{pmatrix}\int_0^{2\pi/\omega}\left[-D+\frac{r}{K_s}f_1\left(S^0-2\frac{y(t)}{Y}\right)\right]dt\\ \int_0^{2\pi/\omega}\left[-D+\frac{r}{K_s}f_2\left(S^0-\frac{y(t)}{Y}\right)\right]dt\end{pmatrix}. \tag{3.12}$$

If λ_1 and λ_2 are both negative, this would imply $(y(t), 0)$ is a stable known solution with respect to the nonlinear system in (3.8). Here we investigate the signs of λ_1 and λ_2 through a series of claims and proofs.

Claim 1. If $y(t) > 0 \ \forall \ t$ and periodic, with period $Q = \frac{2\pi}{\omega}$, then $\lambda_1 < 0$ for all f_1 satisfying the total fitness constraint given in (2.5).

Proof of Claim 1. $S^0 - \dfrac{2y(t)}{Y} = 2\left(S^0 - \dfrac{y(t)}{Y}\right) - S^0 = 2\bar{S}_1 - S^0$. The last equality is a direct consequence of $\left(\bar{S}_1, y(t)\right)$ being a solution of the one-species chemostat system. Hence

$$\begin{aligned}\lambda_1 &= \int_0^Q \left[-D+\frac{r}{K_s}f_1\left(S^0-\frac{2y(t)}{Y}\right)\right]dt\\ &= \int_0^Q \frac{r}{K_s}f_1\left(\bar{S}_1 - S^0\right)dt + \int_0^Q\left(\frac{r}{K_s}f_1\bar{S}_1 - D\right)dt\\ &= \int_0^Q \frac{r}{K_s}f_1\left(\bar{S}_1 - S^0\right)dt. \end{aligned}\tag{3.13}$$

The last step can be verified by substituting $y(t)$ in for x in the species equation given in (2.2), dividing both sides of the species equation by $y(t)$, and then integrating both sides over one period, which is shown next.

$$\begin{aligned}
\frac{\dot{y(t)}}{y(t)} &= \frac{r}{K_s}\bar{S}_1 f_1 - D \\
\int_0^Q \frac{\dot{y(t)}}{y(t)}dt &= \int_0^Q \left[\frac{r}{K_s}\bar{S}_1 f_1 - D\right] dt \\
\ln(y(Q)) - \ln(y(0)) &= \int_0^Q \left[\frac{r}{K_s}\bar{S}_1 f_1 - D\right] dt \\
0 &= \int_0^Q \left[\frac{r}{K_s}\bar{S}_1 f_1 - D\right] dt.
\end{aligned}$$

The Q-periodicity of $y(t)$ (in long time $t > t_L$) (established in Appendix B.1) justifies the last step. Since $\bar{S}_1 < S^0$ when $y(t) > 0$, this implies the integrand in (3.13) is negative for all t. q.e.d.

Note that Claim 1 is true in a very robust way, that is, it is true for every environmental frequency and for all f_1 satisfying the constraint. Also note that an eigenvector corresponding to E_1 is $\mathbf{v}_1 = [1\ \ 0]^T$, and since $\lambda_1 < 0$ this implies $\mathbf{v}_1$ is in the span of the stable subspace. If it turns out to be the only eigenvalue that is negative, then the stable subspace is span $\{\mathbf{v}_1\}$.

Claim 2. Consider the limiting high frequency case, with $\bar{S}_1$ approximated with its asymptotic approximation derived in Chapter 2, $\bar{S}_1 = \dfrac{DK_s}{r\bar{f}_1(t)}$. If f_1 is replaced in the asymptotic approximation of $\bar{S}_1$ with the optimal fitness f_{opt}, obtained from solving the optimization problem stated in (2.4) and (2.5), then $\lambda_2 = \int_0^Q \left[-D + \dfrac{r}{K_s} f_2 \bar{S}_1\right] dt \leq 0$ for all f_2 satisfying $\int_{T_{\text{lo}}}^{T_{\text{hi}}} f_2(T)dT = \beta$.

Proof of Claim 2. Recall that f_{opt} is not unique if $W(T)$ is constant. In fact, it

was shown in Appendix A.4 that when $W(T)$ (the environmental density function) is constant, any fitness response satisfying the constraint is optimal. Therefore, two cases must be considered.

Case 1. $W(T)$ is not constant.

Substituting the asymptotic expansion in for S_1, with the optimal fitness f_{opt} in place of f_1 results in

$$\begin{aligned} \lambda_2 &= D \int_0^Q \left(\frac{f_2}{\bar{f}_{opt}} - 1 \right) dt \\ &= D \left(\frac{Q \bar{f}_2}{\bar{f}_{opt}} \right) - DQ. \end{aligned} \tag{3.14}$$

Since f_{opt} minimizes $\int_0^Q \ln \left(\frac{DK_s}{r\bar{f}} \right) dt$ over all f, this implies $\bar{f}_{opt} > \bar{f}_2$, and hence $\lambda_2 < 0$.

Case 2. $W(T)$ is constant.

From (3.14) and the fact that $\bar{f}_{opt} = \bar{f}_2$ when $W(T)$ is constant, it is easily observed that $\lambda_2 = 0$. q.e.d.

The proofs of Claim 1 and Claim 2 together imply that $(y(t), 0)$ is asymptotically stable with respect to (3.8) in the high-frequency limiting case if $W(T)$ is not constant. When $W(T)$ is constant, this analysis is inconclusive.

Claim 3. Next, consider the limiting low frequency case, with $\bar{S}_1$ this time approximated with the low-frequency asymptotic approximation derived in Chapter 2 (equation (2.8)), $\bar{S}_1 = \frac{DK_s}{rf_1(t)}$. If f_1 is replaced with the optimal fitness f_{opt} in the

asymptotic approximation of $\bar{S}_1$, then $\lambda_2 = \int_0^P \left[-D + \frac{r}{K_s} f_2 \bar{S}_1\right] dt$ is identically 0 for all f_2 satisfying $\int_{T_{\rm lo}}^{T_{\rm hi}} f_2(T)dT = \beta$.

Proof of Claim 3. Recall from the proof of Claim 1 that $\bar{S}_1 = S^0 - \frac{y(t)}{Y}$. Applying this to the λ_2 equation from (3.12) and then substituting $\bar{S}_1 = \frac{DK_s}{r f_{opt}(t)}$ into λ_2 yields

$$\begin{aligned}
\lambda_2 &= \int_0^Q \left(-D + \frac{r}{K_s} f_2 \bar{S}_1\right) dt \\
&= D \int_0^Q \left(\frac{f_2}{f_{opt}} - 1\right) dt \\
&= D \int_{T_{\rm lo}}^{T_{\rm hi}} \left(\frac{f_2(T)}{f_{opt}(T)} - 1\right) W(T) dT
\end{aligned} \tag{3.15}$$

The last step shown above is a consequence of the change of variables from t to T, which introduces the environmental density function $W(T)$ into the integrand. This particular change of variable was thoroughly discussed in Appendix A, Section A.4. Recall that the optimization calculation in the limiting low-frequency case resulted in

$$f_{opt}(T) = \beta \frac{W(T)}{\int_{T_{\rm lo}}^{T_{\rm hi}} W(T)}. \tag{3.16}$$

Substituting (3.16) into (3.15) yields

$$\lambda_2 = D \int_{T_{\rm lo}}^{T_{\rm hi}} \left(\frac{\int_{T_{\rm lo}}^{T_{\rm hi}} W(T)dT}{\beta} \frac{f_2}{W(T)} - 1\right) W(T) dT. \tag{3.17}$$

Distributing $W(T)$ through in the integrand of (3.17) and applying the integral constraint to f_2 cancels all terms, completing the proof of Claim 3. q.e.d.

The fact that $\lambda_2 = 0$ means that further investigation is required to learn anything about the stability of the full two-species system in the low-frequency limiting case. That is, there are not any conclusions to infer about the ability of the hypothesized optimal species to fight off invasion from all other species. What can be shown, and is done in claims 4 and 5 below, is that the hypothesized optimal species is the only species that can invade any other.

Claim 4. If f_1 is optimal, then $(0, y(t))$, where $y(t)$ is now a species that has fitness response $f_2 \neq f_{opt}$, is unstable to perturbation in the limiting low- frequency case.

Proof of Claim 4. Let $f_1 = f_{opt}$. Then

$$\int_0^Q \ln\left(\frac{DK_s}{rf_2}\right) dt > \int_0^Q \ln\left(\frac{DK_s}{rf_{opt}}\right) dt$$

for any f_2 satisfying $\int_{T_{\text{lo}}}^{T_{\text{hi}}} f_2(T)dT = \beta$. Moving all terms to one side and using basic log rules results in

$$\begin{gathered}\int_0^Q \ln\left(\frac{f_{opt}}{f_2}\right) dt > 0 \Rightarrow \\ \frac{1}{Q}\int_0^Q \ln\left(\frac{f_{opt}}{f_2}\right) dt > 0 \Rightarrow \\ \exp\left(\frac{1}{Q}\int_0^Q \ln\left(\frac{f_{opt}}{f_2}\right) dt\right) > 1.\end{gathered}$$

Since the arithmetic mean is greater than or equal to the geometric mean (Alzer,

1996; Xia et al., 1999)

$$\frac{1}{Q}\int_0^Q \frac{f_{opt}}{f_2}dt \geq \exp\left(\frac{1}{Q}\int_0^Q \ln\left(\frac{f_{opt}}{f_2}\right)dt\right) > 1 \Rightarrow \tag{3.18}$$

$$\int_0^Q \frac{f_{opt}}{f_2}dt > Q \Rightarrow$$

$$\int_0^Q \left(\frac{f_{opt}}{f_2} - 1\right)dt > 0 \Rightarrow$$

$$D\int_0^Q \left(\frac{f_{opt}}{f_2} - 1\right)dt > 0.$$

The left-hand side of the last inequality is exactly one of the Floquet exponents when $(0, y(t))$ is the known solution used in the linearization. Since λ_2 is positive for all f_2 when $f_1 = f_{opt}$, $(0, y(t))$ is unstable to perturbation. q.e.d.

This result implies that a species with optimal fitness (with respect to minimizing the geometric mean of substrate) can invade any other species in the limiting low-frequency regime. Observe that this result also holds for the limiting high-frequency regime (replace f_1 with $\bar{f}_1$ and f_2 and $\bar{f}_2$). Next, it is now shown that the optimal species is the only species that can invade any other species.

Claim 5. If $f_1 \neq f_{opt}$, then $(0, y(t))$ with $y(t)$ having fitness response $f_1(1-\epsilon) + \epsilon f_{opt}$, with $0 < \epsilon < 1$, is stable to invasion by a species with fitness f_1 in the limiting low-frequency regime.

Proof of Claim 5. This proof involves long division, Jensen's inequality, and the fact that was proven in claim 3, i.e.,

$$\int_0^Q \left(\frac{f_1}{f_{opt}} - 1\right)dt = 0 \tag{3.19}$$

$\forall\ f_1$ satisfying the same constraint as f_{opt} (only true in the limiting low-frequency case). In this argument I will use the easily observed extension of (3.19), $\int_0^Q \frac{f_1}{f_{opt}} dt = Q$. By again noting that $S^0 - \frac{y(t)}{Y} = \bar{S}_1 = \frac{DK_s}{rf}$, the second Floquet exponent in (3.12) can be rewritten as

$$\lambda_2 = D \int_0^Q \left(\frac{f_1}{f_1 (1-\epsilon) + \epsilon f_{opt}} - 1 \right) dt.$$

First, long division is applied to the first term in the integrand, restricting $0 < \epsilon < 1$, to get

$$\begin{aligned} \frac{f_1}{f_1 (1-\epsilon) + \epsilon f_{opt}} &= \frac{1}{1-\epsilon} - \frac{\frac{\epsilon}{1-\epsilon} f_{opt}}{f_1 (1-\epsilon) + \epsilon f_{opt}} \\ &= \frac{1}{1-\epsilon} \left(1 - \frac{\epsilon f_{opt}}{f_1 (1-\epsilon) + \epsilon f_{opt}} \right), \end{aligned} \tag{3.20}$$

which implies another way to write λ_2 is

$$\lambda_2 = D \int_0^Q \left(\frac{1}{1-\epsilon} \left(1 - \frac{\epsilon f_{opt}}{f_1 (1-\epsilon) + \epsilon f_{opt}} \right) - 1 \right) dt. \tag{3.21}$$

The approach taken here will show that (3.21) is negative. Note once again that this proof only holds for the limiting low-frequency case. Recall from Chapter 2 (equation (2.9)) that

$$f_{opt} = \beta \frac{W(T)}{\int_{T_L}^{T_H} W(T) dT}. \tag{3.22}$$

Assuming the given environmental history $T(t)$ is sufficiently smooth, which implies $W(T) > 0\ \forall\ T$, which in turn implies $f_{opt}(T)$ will be bounded away from 0 for all

T, and hence all t. This is important to note because it is possible that $\exists$ at least one t such that $f_1(T(t)) = 0$.

For notational convenience, let the second term in the parenthesis in (3.20) be defined as Z(t), that is

$$Z(t) := \frac{\epsilon f_{opt}}{f_1\left(1-\epsilon\right) + \epsilon f_{opt}}.$$

Applying Jensen's inequality, $\phi\left(\int_\Omega Y(t)dt\right) \leq \int_\Omega \left(\phi \circ Y\right) dt$ (or said another way, using the fact that the arithmetic mean is greater than or equal to the harmonic mean), to the average of $Z(t)$ over one period yields

$$\begin{aligned}
\frac{1}{Q}\int_0^Q Z(t)dt &= \frac{1}{Q}\int_0^Q \frac{\epsilon f_{opt}}{f_1\left(1-\epsilon\right) + \epsilon f_{opt}}dt && (3.23)\\
&\geq \frac{1}{\frac{1}{Q}\int_0^Q \frac{f_1(1-\epsilon)+\epsilon f_{opt}}{\epsilon f_{opt}}dt} \\
&= \frac{1}{\frac{1}{Q}\left[\left(\frac{1-\epsilon}{\epsilon}\right)\int_0^Q \frac{f_1}{f_{opt}}dt + Q\right]} && (3.24)\\
&= \frac{1}{\frac{1}{Q}\left[\left(\frac{1-\epsilon}{\epsilon}\right)Q + Q\right]} && (3.25)\\
&= \frac{1}{\left[\left(\frac{1-\epsilon}{\epsilon}\right)+1\right]} \\
&= \epsilon,
\end{aligned}$$

where the step from (3.24) to (3.25) is validated by substituting (3.22) into (3.24) and noting that $\int_{T_L}^{T_H} W(T)dT = Q$. Next, starting from the result of the above

computation

$$\begin{aligned}
\int_0^Q \frac{\epsilon f_{opt}}{f_1(1-\epsilon)+\epsilon f_{opt}}dt \geq Q\epsilon \Rightarrow \\
-\int_0^Q \frac{\epsilon f_{opt}}{f_1(1-\epsilon)+\epsilon f_{opt}}dt \leq -Q\epsilon \Rightarrow \\
Q-\int_0^Q \frac{\epsilon f_{opt}}{f_1(1-\epsilon)+\epsilon f_{opt}}dt \leq Q-Q\epsilon \Rightarrow \\
\frac{1}{1-\epsilon}\left(Q-\int_0^Q \frac{\epsilon f_{opt}}{f_1(1-\epsilon)+\epsilon f_{opt}}dt\right) \leq Q \Rightarrow \\
\frac{1}{1-\epsilon}\left(Q-\int_0^Q \frac{\epsilon f_{opt}}{f_1(1-\epsilon)+\epsilon f_{opt}}dt\right) - Q \leq 0. \qquad (3.26)
\end{aligned}$$

Observe that the left-hand side of (3.26) is precisely $\frac{1}{D}\lambda_2$ (as stated in (3.21)), which implies $\lambda_2 \leq 0$. Note that Jensen's inequality, with a convex ϕ ($\phi(x) = 1/x$ here), implies that equality is only true if the integrand is constant. The integrand in the step where Jensen's was applied cannot be constant ($f_1 \neq f_{opt}$ by assumption, and both f_1 and f_{opt} cannot be constant, unequal, and satisfy the same constraint). Therefore

$$\lambda_2 < 0,$$

for $0 < \epsilon < 1$. q.e.d.

The last inequality implies both Floquet exponents are negative in this case, which in turn implies $(0, y(t))$ is stable. Therefore, if $f_1 \neq f_{opt}$ is the fitness response of species x_1, there always exists at least one species that x_1 cannot invade. Together

claims 4 and 5 show that, in the limiting low-frequency regime, the species with fitness f_{opt} is the only species that can invade any other species.

The stability results from the limiting low-frequency case are summarized in the phase diagram shown in Figure 3.1. The combined results of claims 1 and 3 are represented by the blue circle on the x_1 axis, where the x_1 component is the one-species periodic solution (in the limiting low-frequency case) with $f_1 = f_{opt}$. Since claim 3 did not provide conclusive evidence of stability (a zero Floquet exponent), an arrow has not been inserted in a neighborhood of this point (in the first quadrant). The combined results from claims 1 and 4 are represented by the red circle on the x_2 axis, where the x_2 component is the one species periodic solution with $f_2 \neq f_{opt}$, and the arrow leading away from this circle denotes the instability of this solution. The arrows pointing away from the red circle at the origin depict the instability of the trivial solution, which can be observed by studying equations (3.3). The boundary of the solution space consists of the positive x_1 and x_2 axes, and the line $S^0 - \frac{1}{Y}(x_1 + x_2) = 0$ (Figure 3.1, red line). The arrows on the red line represent the fact that both differential equations shown in (3.3) are negative on this line. The green circle represents an interior solution, which may or may not exist. To confirm that minimizing the geometric mean of substrate in the one-species chemostat does indeed result in the species that can outcompete any other species in a two-species chemostat, all interior solutions would have to be unstable and the solution represented by the blue circle would have to be stable.

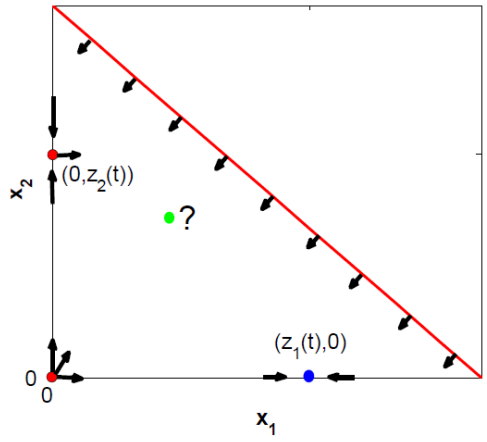


Figure 3.1. Partial phase diagram, combining the stability results with observations made from analyzing (3.3). Red dots represent known unstable solutions, the blue dot represents the known solution that led to the zero Floquet exponent described in claim 3, and the green dot represents other possible interior solutions. The red line, and the x_1 and x_2 axes represent the boundaries of the solution space. The local stability analysis performed here does not provide any information regarding the stability of the possible interior solutions.

Evolution of Arbitrary Species to Optimal Species

In the previous section a stability analysis approach was taken to investigate the invasibility properties of both the hypothesized optimal species and other, sub-optimal species. However, the analysis that was presented did not provide any information regarding the interaction between the invader and resident species if invasion was successful. In this section an example is presented that demonstrates how a sequence of invasion events could lead to an arbitrary resident species evolving into the hypothesized optimal species. This investigation is performed through a

sequence of numerical competition experiments in a two-species chemostat that is operating under limiting low-frequency temporal variations.

For this example a Gaussian fitness response function, f_0, was assigned to the initial resident species, in which the Gaussian satisfied the total fitness constraint and had mean equal to that of the average environmental condition. The invading species was assigned the fitness response f_1, which was defined as $f_1 = f_0 + \delta(-f_0 + f_{opt})$, with $\delta = 1/N$, where N is the number of sequential steps needed to reach the optimal species. Note that f_1 still satisfies the total fitness constraint and it represents a fitness response that is shifted in the optimal direction with magnitude δ. The two-species chemostat equations (3.1) were then numerically simulated by providing the resident species with an initial condition that was 10^6 times that of the initial condition of the invading species. This ratio of initial conditions was selected because it is approximately equal to the reported mutation rate per gene per generation in *E. coli* (Drake, 2009). The numerical results suggest that the invading species competitively excludes the resident species after $\sim$100 generations (environmental cycles) (see Figure 3.2A). This procedure was repeated by then competing the former invader, and new resident, with a new invader that had fitness response f_2 defined as $f_2 = f_1 + \delta(-f_0 + f_{opt})$. This again resulted in the invader competitively excluding the resident, with the number of generations to reach exclusion approximately twice that of the previous experiment. This process was iterated $1/\delta$ times, i.e., until the invader was the hypothesized optimal species. In each step the invader species had

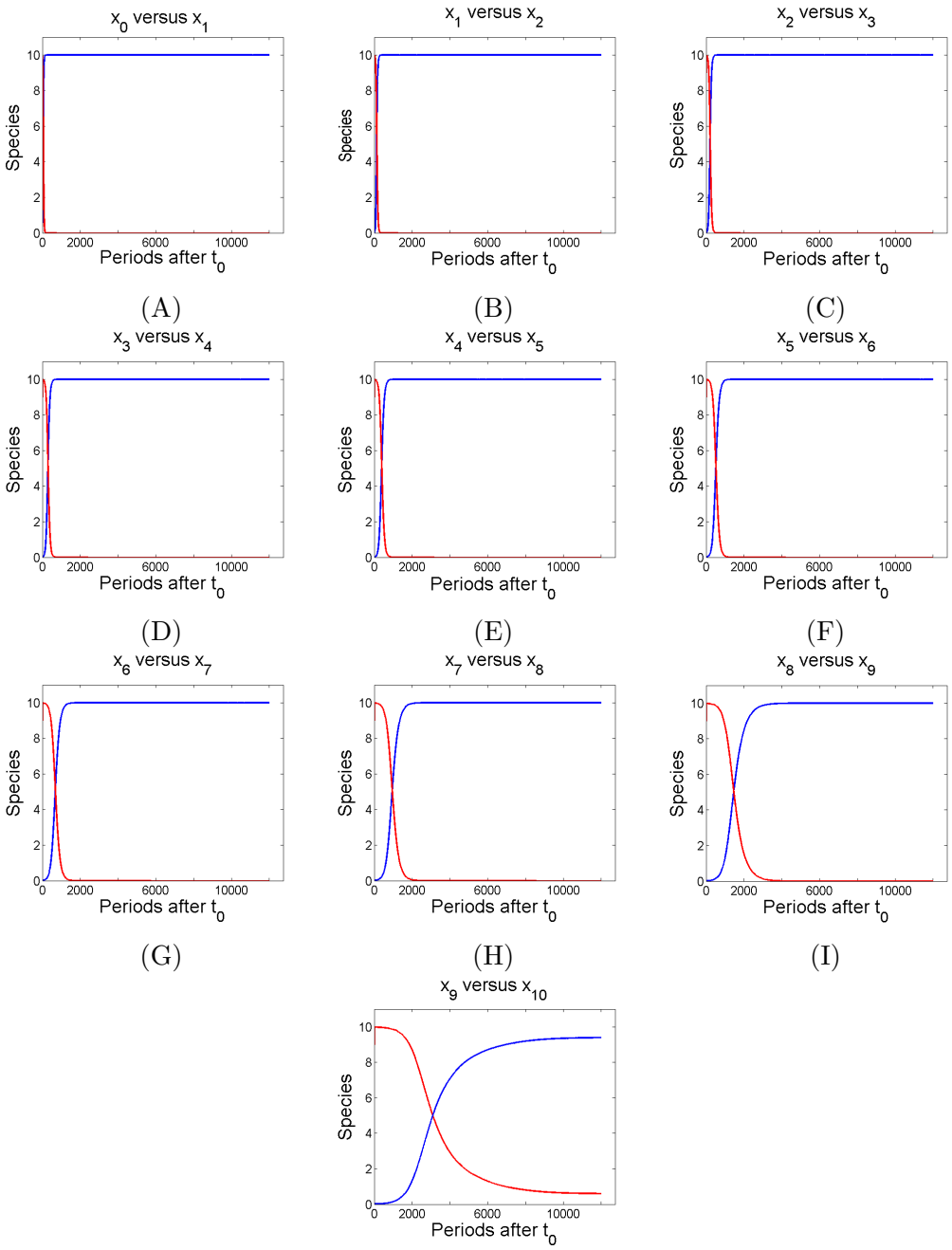
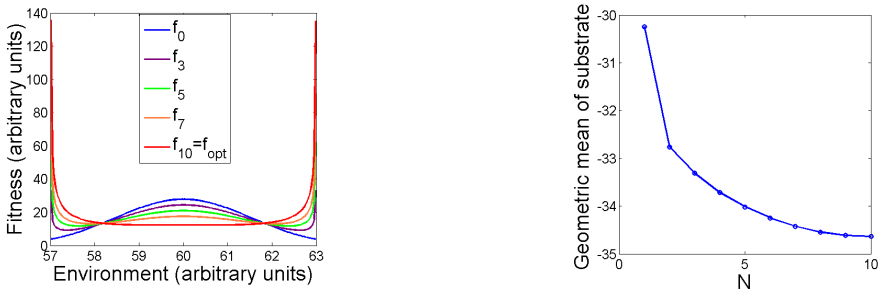


Figure 3.2. Numerical competition experiments between resident and invader species. Blue species is the invader species and red is the resident species. Top left is the original resident with Gaussian fitness (red). Moving from left to right and top to bottom the former invader becomes the new resident and the invader species evolves toward the optimal species.

fitness response f_k and the resident species that had fitness f_{k-1}, where

$$f_k =: f_{k-1} + \delta(-f_0 + f_{opt}),\ k = 1...N. \tag{3.27}$$

Each time the invader competitively excluded the resident, and each time, the number of generations (periods) to reach exclusion increased. The results for all competition experiments with species x_k versus x_{k-1} are shown in Figure 3.2 for $k = 1$ to $k = N$, with $\delta = 1/N$. In Figure 3.3A several of the evolving fitness response functions, f_k, are shown. In Figure 3.3B, the geometric mean of substrate in the one species model (the objective function used in the optimization procedure of Chapter 2) is plotted for species with fitness responses f_0 through f_N. Note that as the



(A) (B)

Figure 3.3. (A) Plots of selected fitness responses, starting from a Gaussian fitness curve, f_0, and moving by integer multiples of δ towards the fitness of the optimal species f_{opt} in the low frequency regime. (B) Geometric mean of substrate for species with fitness f_1 through f_N.

hypothesized optimal species is approached, the value of the objective function exponentially decays, which corresponds to the increasing time to exclusion that is shown in the successive competition experiments. One explanation for the observed correspondence is that the evolutionary process slows as the optimum is approached. Collectively, this sequence of numerical experiments provides a hypothetical way that a resident species could evolve towards the optimal species through a series of mutation and periodic selection events, which mimics the Stable Ecotype Model of species and speciation (Cohan and Perry, 2007) that is discussed in Chapter 5.

Summary

Although the stability results did not provide conclusive evidence that the hypothesized optimal species would always outcompete any other species in all frequency regimes, some meaningful results were still obtained. First, x_{opt} (a species with the hypothesized optimal fitness response f_{opt}) can outcompete all other species in the high frequency regime and second, x_{opt} is the only species that can invade all other species in the limiting low-frequency regime. The stability analysis applied here was inconclusive with respect to determining x_{opt}'s ability to fight off invasion from any other species in the limiting low-frequency regime, which is a consequence of one of the Floquet exponents of the monodromy matrix always being exactly 0. This result is interesting and worthy of further investigation. Initial alternative

techniques to study the long-term behavior are underway and preliminary findings are included in Appendix B.2.

Also, a two-species chemostat model was studied in this chapter to present a theoretical-based example of bacterial evolution. Although the results were based on a single example, it was shown that an extremely rare, but more fit population can competitively exclude the numerically dominant population. The more fit population was defined by moving the fitness response function of the dominant population in the direction of the optimal species, and initial conditions were set to reflect that of a mutant and dominant species. A series of numerical competition experiments between the invader species (mutant) and resident species (dominant) were then computed and showed the evolution of a species with Gaussian fitness to a species with the hypothesized optimal fitness.

CHAPTER 4

EVIDENCE OF CLOSELY RELATED *SYNECHOCOCCUS* SPECIES INHABITING THE MICROBIAL MATS OF MUSHROOM SPRING, YELLOWSTONE NATIONAL PARK

Contribution of Authors and Co-Authors

Manuscript in Chapter 4

Author: Shane Nowack

Contributions: Obtained partial funding (Montana Space Grant Consortium), collected samples from the field site, isolated the organisms in the laboratory, performed the experiments, and wrote the manuscript.

Co-Author: Millie T. Olsen

Contributions: Assisted in cultivation and performed molecular procedures.

Co-Author: Eric D. Becraft

Contributions: Performed molecular procedures and analyzed sequence data.

Co-Author: Donald A. Bryant

Contributions: Discussed results and edited the manuscript.

Co-Author: David M. Ward

Contributions: Partially provided funding, assisted with experimental design, discussed the results and edited the manuscript at all stages.

Manuscript Information Page

Shane Nowack, Millie T. Olsen, Eric D. Becraft, Donald A. Bryant, and David M. Ward
Applied and Environmental Microbiology
Status of Manuscript:
_x_ Prepared for submission to to a peer-reviewed journal
___ Officially submitted to a peer-reviewed journal
___ Accepted by a peer-reviewed journal
___ Published in a peer-reviewed journal

Abstract

Closely-related thermophilic *Synechococcus* strains with identical or nearly identical 16S rRNA sequences, but with sequences at the more highly-resolving *psaA* locus that are representative of different putative ecological species populations, were isolated from the microbial mats found in the effluent channels of Mushroom Spring, Yellowstone National Park (YNP). High-throughput Ti454-barcode sequencing was used to confirm that the isolates were predominated by a single *psaA* sequence. Differences in growth rate as a function of light intensity were observed among the strains and reflected the light environments of the vertical positions in the mat where the associated species populations were found. The strain with the highest light tolerance had a *psaA* genotype corresponding to the species that was found in greatest abundance near the surface. The other strains exhibited lower light tolerances, and had *psaA* genotypes that were representative of predominant subsurface species populations in the mat. Additionally, it is shown that the light responses of the isolated strains were strongly influenced by the availability and form of dissolved inorganic carbon provided, and the temperature at which the experiments were conducted. The cells representative of the different strains were also shown to have similar morphologies under low-light conditions, but under high-light conditions, differences in morphological acclimation among the strains were observed.

Together, the results presented here suggest that (i) several closely-related yet distinct, ecological species, filling different light niches, exist in the Mushroom Spring microbial mats, (ii) the fundamental light niche is interconnected with other environmental parameters, and (iii) molecular resolution beyond the 16S rRNA locus may be required in order to identify the ecologically distinct populations in microbial communities.

Introduction

The primary producers of the microbial mats found in the effluent channels of two chemically similar alkaline siliceous hot springs (pH ∼8.2), Mushroom Spring and Octopus Spring in the Lower Geyser Basin of Yellowstone National Park (YNP) (Papke et al., 2003), are members of the genus *Synechococcus* (Brock, 1978; Ward et al., 2012), with the predominant organisms at temperature sites of 50°C and hotter belonging to the 16S rRNA-defined *Synechococcus* A/B clade (Ferris and Ward, 1997). Molecular analyses of mat samples collected along temperature and light gradients in Mushroom Spring have revealed genetic diversity at the 16S rRNA locus and internal transcribed spacer (ITS) region that separates the 16S and 23S rRNA genes. For example, seasonally stable distributions of 16S rRNA genotypes A″, A′, A, B′, and B were progressively detected at temperatures ranging from 72°C to 50°C along the thermal gradients in both Mushroom Spring and Octopus Spring (Ferris and Ward, 1997; Ward et al., 2006). Sequence diversity has also

been observed along the vertical gradient at various temperature sites in Mushroom Spring. At a 60°C site it was reported that B′-like 16S rRNA sequences were found throughout the top 1 mm layer of the mat, whereas A-like sequences were only found between depths of 400 and 800 μm (Ramsing et al., 2000). At a 68°C site, minimal sequence diversity was detected at the 16S rRNA locus along the vertical gradient, but sequence variation was observed when the more highly divergent ITS region was analyzed (Ferris et al., 2003). Observations of genetic diversity along these environmental gradients have led to hypotheses regarding the existence of closely-related *Synechococcus* spp. that may have evolved distinct ecological adaptations to temperature and light (Ward, 1998; Ward and Cohan, 2005).

In addition to the molecular studies, oxygen microsensor data collected from the mat during a disturbance experiment in Octopus Spring (Ferris et al., 1997) showed a bimodal occurrence of oxygenic photosynthetic activity as the mat recovered. The first peak occurred near the surface of the mat, as expected, but a second peak was observed at a depth of $\sim$1 mm in the mat, which suggested the existence of both low-light and high-light adapted *Synechococcus* populations.

To test these adaptation hypotheses, laboratory isolates were cultivated that had genotypes corresponding to the predominant 16S rRNA and ITS sequences found in Octopus Spring (Allewalt et al., 2006). The cultivated strains had distinct temperature adaptations, e.g., the A-like isolates exhibited a higher upper-temperature limit

than the B′-like isolates. Cultivated strains of *Synechococcus* spp. with distinct temperature adaptations were also reported in previous studies of Oregon hot springs (Peary and Castenholz, 1964; Miller and Castenholz, 2000). The significance of the temperature adaptations observed in the Octopus Spring study was that the isolates had 16S rRNA and ITS genotypes that corresponded to the predominant sequences found *in situ*, and the laboratory-observed temperature adaptations corresponded with the distributions of these sequence types found along the thermal gradient.

Allewalt (2004) studied growth rates of *Synechococcus* spp. A, B and B′ isolates at light intensities between 10 and 385 μmol photons/m^2/sec, but results did not provide clear and reproducible evidence of differences in adaptation to light. Only small differences were observed in the amount of light required to support net photosynthesis of the A, B and B′ strains when temperatures approached the upper limit (Allewalt et al., 2006). Light was not isolated as a variable in these studies because saturating levels of CO_2 were not provided. Oxygenic phototrophs utilize light to fix CO_2, so that consequently, it is possible that the observed light responses were due to CO_2 limitation. In a later study of a subculture derived from the B′ isolate investigated in Allewalt et al. (2006) (Kilian et al., 2007) CO_2 limitation was prevented by bubbling 3% CO_2 in air during the experiment. Under these conditions it was shown that this isolate could grow in continuous light environments at an intensity of 200 but not 400 μmol photons/m^2/sec, far below downwelling irradiances of $\sim$2000 μmol photons/m^2/sec and scalar irradiance intensities of up

to 3500 μmol photons/m^2/sec, which have been measured *in situ* in midsummer (Chapter 2). Due to light scatter, scalar irradiance measurements are a more appropriate measure of the *in situ* light reality, especially with depth in the mat (Kühl et al., 1994). The relative intolerance of the isolate studied by Kilian et al. (2007) to higher light intensities was surprising in that the Ramsing et al. (2000) study reported the 16S rRNA B′-like sequence to be predominant throughout the top 1 mm of the microbial mat, especially at and near the mat surface. It had therefore been hypothesized that an isolate with this genotype would be able to tolerate high light levels.

A possible explanation for the failure of Allewalt (2004) to observe reproducible light-adaptation results might relate to culture impurity (i.e., the presence of multiple *Synechococcus* strains with different light adaptations), and the inability to recognize impurity because the highly conserved 16S rRNA sequence was used to discern differently adapted species that can only be recognized by using more rapidly evolving genetic loci (see below). This is especially problematic since the isolates studied by Allewalt et al. (2006) originated from 10^2- to 10^5-fold dilutions of the original mat sample, which was estimated to contain between 10^8 and 10^{10} cells/mL (Bauld and Brock, 1974; Ferris et al., 1996). Furthermore, when the serially diluted mat samples were plated, colony counts decreased from thousands to zero in a single ten-fold dilution. Isolates obtained from low-dilution environmental samples often are not representative of the predominant populations found *in situ* (Santegoeds et

al., 1996). The observed plating irregularities prompted Allewalt (2004) to conduct experiments designed to improve the culture medium, but none of the supplements tested improved colony recovery on high-dilution plates. The mismatch between the lower light tolerance of the B′ isolate studied by Kilian et al. (2007) and the high light intensities occurring in nature might also be due to the selection of a relatively rare, low-light adapted B′-like *Synechococcus* strain, especially since the initial cultivation of the strain by Allewalt et al. (2006) was done at a low light intensity (downwelling irradiance of ∼60 μmol photons/m^2/sec).

Recent analyses have examined sequence variation in protein-encoding loci, which have evolved more rapidly than 16S rRNA and ITS loci and thus offer greater molecular resolution. A theory-based evolutionary simulation model, the Ecotype Simulation algorithm (Koeppel et al., 2008), was used to predict ecologically distinct species from such sequence data. Results from these studies suggested that the 16S rRNA locus and ITS region may not offer enough molecular resolution to identify the true, fundamental species-like units in the Mushroom Spring and Octopus Spring systems (Becraft et al., 2011; submitted; Melendrez et al., 2011). In particular, mat samples collected from Mushroom Spring exhibited sequence variation at the *psaA* locus (a photosystem I reaction center subunit gene essential for photosynthesis and unique in this community to *Synechococcus* spp.) that was characteristic of what would be expected of several ecologically distinct populations (Cohan and Perry, 2007). Specifically, the Ecotype Simulation algorithm has predicted as many as 15

different *psaA* putative ecological species (or putative ecotypes, PEs) within the 16S rRNA-defined A clade (see Figure 4.1A), and 18 different PEs within the 16S rRNA-defined B′ clade (see Chapter 5, Figure 5.2).

Moreover, *psaA* PE distributions along the vertical gradient in the 60°C Mushroom Spring mat (Becraft et al.; submitted) have suggested that different *Synechococcus* PEs may have different adaptations to light (see Figure 4.1B for distributions and Figure 4.1C for *in situ* light attenuation profile). For example, PE A1 was shown to be more abundant near the surface compared to the other predominant A-like PEs (A4, A6, and A14), which were found almost exclusively in deeper mat layers. The vertical distributions at 60°C also revealed a B′-like PE, PE B′9, in relatively high abundance nearer to the surface than the A-like PEs, suggesting that this PE may be well-adapted to higher light intensities at this temperature. These results also led to an interest in determining both the purity and distribution, with respect to the *psaA* locus, of the A and B′ 16S rRNA isolates discussed in Allewalt et al. (2006), which might provide an explanation for why the B′-like isolate did not have the hypothesized light response in the earlier experiments (Allewalt et al., 2006; Kilian et al., 2007).

In the study presented here we test the hypotheses that strains representative of different *Synechococcus* PEs have different adaptations to light, and that the light responses are dependent on temperature and the availability of CO_2. Isolates were cultivated from Mushroom Spring that have *psaA* genotypes corresponding to the

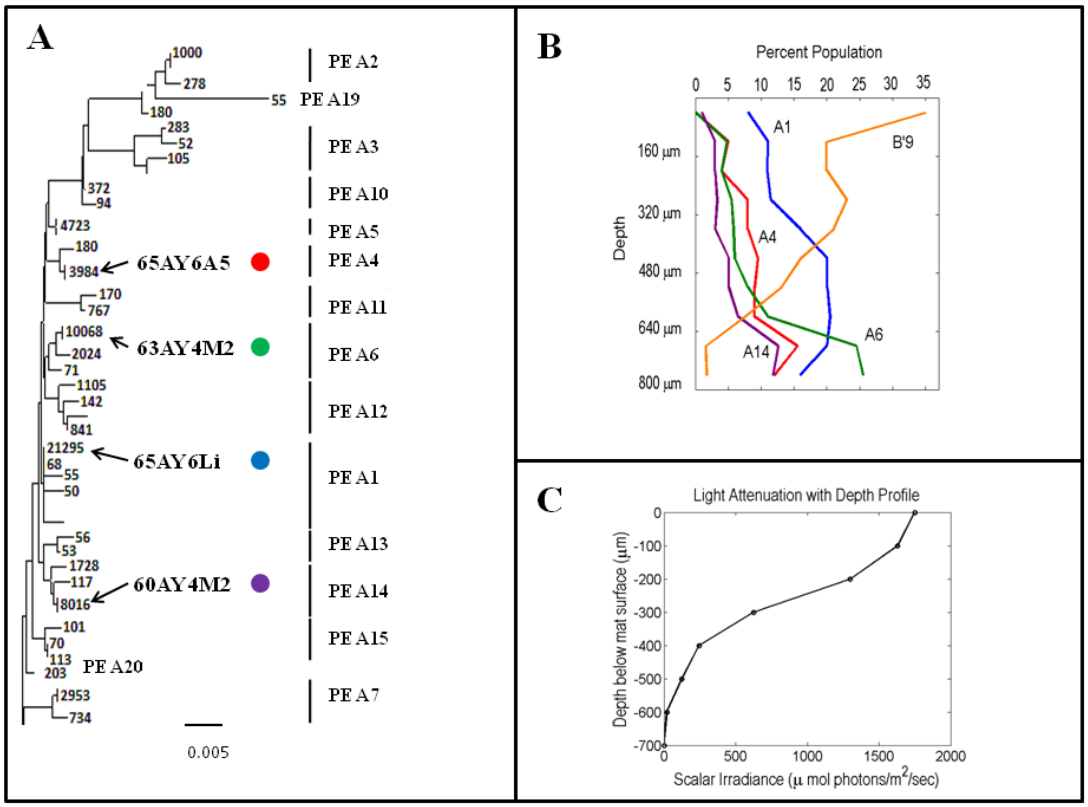


Figure 4.1. Relationship between *Synechococcus* strains and putative ecotypes (PEs) and their vertical distributions relative to light intensities in the Mushroom Spring 60°C mat. (A) Neighbor-joining tree containing all high-frequency *Synechococcus* A-like *psaA* sequences (i.e., sequences occurring at least 50 times across the 96 DNA samples analyzed, including dominant variants) identified in Ti454-barcode analysis of ∼60 to 68°C region of the Mushroom Spring microbial mat. Black vertical bars represent PEs predicted by Ecotype Simulation. Numbers indicate the total occurrences of a particular sequence observed across all DNA samples. Colored PE designations represent predominant populations discussed in the main text, and colors correspond across Figures 4.1, 4.4, 4.5, and 4.6. Arrows point to the dominant variant sequence of the PE to which cultivated strains correspond identically. Scale bar represents 0.005 substitutions per site. (B) Relative abundance of predominant *psaA* putative ecotypes in ∼160 μm subsections along the vertical gradient in the ∼60°C region of the Mushroom Spring mat. (C) Vertical scalar light intensity profile for the same region sampled for barcode analyses (modified from Becraft et al, submitted).

predominant A-like *psaA* PEs found *in situ*. Growth rates of the isolates were measured over the range of scalar light intensities that have been recorded at Mushroom Spring (0 - 3000 μmol photons/m^2/sec), at two fixed temperatures (52 and 60°C), and with saturating levels of CO_2 provided. The light adaptation results were then compared to the vertical distributions of the corresponding PEs found *in situ* – the goal being to identify ecological adaptations that could explain the observed diversity and distribution of PEs found in the natural environment. In addition to the increased molecular resolution provided by analyzing the highly-resolved *psaA* locus, improvements in experimental methods from those that were used in previous studies were made (Allewalt et al., 2006; Kilian et al., 2007). First, a new cultivation protocol was developed in order to obtain isolates from more highly diluted mat samples; hence, cultivated strains were more likely to be representative of the predominant organisms. Second, high-throughput Ti454-barcode sequencing methods were used to determine the *psaA* genotypes of the isolates. This also provided deep sampling of the genetic structure of isolate populations to ensure that they were predominated by only one *Synechococcus* strain. Third, light adaptations were studied using a growth chamber that is capable of achieving light intensities comparable to those found in nature. Fourth, in order to isolate light as a variable, saturating levels of CO_2 were provided. CO_2 availability is likely to fluctuate during the daytime, since photosynthesis consumes CO_2, which is acidic, shifting the equilibrium of dissolved inorganic carbon (DIC) species towards bicarbonate and

carbonate, and hence, increasing pH. Evidence of this occurring in the mat is that intense mid-day photosynthesis raises pH values to >9.4 (compared to a pH of 7.2 before sunrise) (Revsbech and Ward, 1984). Therefore light responses of two strains were also determined under conditions of CO_2 limitation, because these conditions are probably more natural since photosynthesis links light and CO_2 demand. Finally, post-experiment samples were collected for each isolate after incubation at the experimental light conditions in order to ensure that the *psaA* genotype had not shifted during exposure to different light intensities.

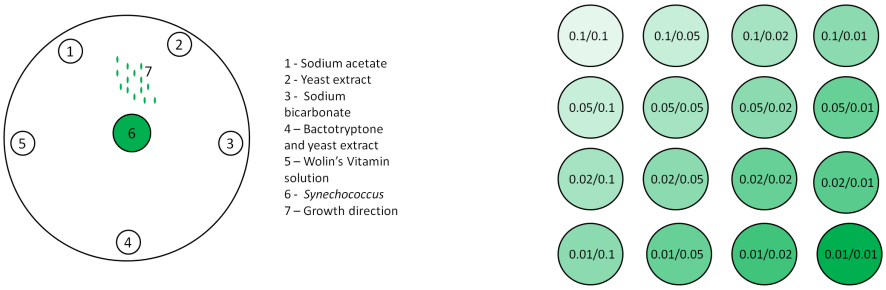
Materials and Methods

Nutrition Experiments

Initially, attempts were made to improve the culture medium that had been used by Allewalt et al. (2006). The first experiment was designed to test whether nutrients evaluated by Allewalt et al. (2006) might be required at lower concentrations than had been used in that study, or that a combination of nutrient deficiency and quorum sensing (Miller and Bassler, 2001) was involved. 1.5% (w/v) Gelrite powder (Research Products International, Chicago, IL) was added to medium DH (Castenholz, 1969) (with pH set to 8.2 to reflect *in situ* reality) and then autoclaved. After allowing the mixture to cool to 50°C in a water bath, a 25 mL volume was inoculated with 10^6 cells of a batch culture of the *Synechococcus* A-like strain (JA-3-3Ab) obtained by Allewalt et al. (2006), and then poured into a 60 mm diameter

Petri dish. After the medium was allowed to solidify, five equally spaced cylindrical wells (#2 cork borer; 5 mm diameter) were bored along the circumference of the plate, and one additional well was bored in the geometric center of the plate using a sterile #4 cork borer (8 mm diameter). Nutrients not found in medium DH but suspected to possibly have a stimulatory effect were included in separate wells at a concentration of 1% (w/v). These included yeast extract, sodium acetate trihydrate, sodium bicarbonate, a mixture of tryptone and yeast extract in equal proportions (comprising the 1% (w/v)), and Wolin's vitamin solution (Wolin et al., 1964) as depicted in Figure 4.2A. To provide any necessary quorum sensing compounds a #4 core of the *Synechococcus* culture described above, that was densely growing in medium DH on solid Gelrite medium, was placed in the central well to provide a source of quorum-sensing compounds. The plate was then incubated at 52°C and a scalar irradiance of 50 μmol photons/m^2/sec of white fluorescent light and observed daily to determine which nutrients stimulated colony development.

In response to the results from this first experiment, a second experiment was conducted to evaluate the concentrations of nutrients that most stimulated growth. Gelrite plates (1.5% (w/v)) were poured with various concentrations of sodium acetate trihydrate and yeast extract added to medium DH (16 different conditions in total, see cartoon in Figure 4.2B, ranging from a concentration of 0.01% to 0.1% (w/v)). Each plate was inoculated with 10^6 cells of the same culture that was used



(A) (B)

Figure 4.2. Schematic summarizing results of nutrition experiments. (A) Nutrition experiment in which nutrients were placed in wells on the circumference of the plate, with a dense *Synechococcus* culture in the central well. Optimal growth of *Synechococcus* inoculated in low density throughout the plate is depicted by green rods between the sodium acetate and yeast extract wells. (B) Nutrition experiment in which different concentrations of sodium acetate and yeast extract were provided (numbers inside plates are percent sodium acetate/yeast extract (w/v)) to a low-density *Synechococcus* culture inoculated throughout the plate. Density of green is proportional to growth of *Synechococcus*.

in the first experiment. The condition that resulted in the best growth (most observable biomass) was then tested against the basal medium DH. This was done by inoculating both types of media with a batch culture that had been serially diluted tenfold to extinction. The plates were then incubated for several weeks under the above-described conditions.

Sample Collection

Samples for cultivation were collected from Mushroom Spring, YNP, at temperature sites of 60, 63, and 65°C on 7 September 2010 using a #4 cork borer (8 mm diameter). The top green layer, ∼1 mm in thickness, of each mat sample was then removed with a razor blade, placed in a 1.5 mL Eppendorf tube (Eppendorf

Canada, CAN), and then returned to the lab in a thermos containing Mushroom Spring source water (cooled to the temperature at which the sample was collected). Time from sampling to the lab was approximately two hours, with the temperature in the thermos approximately 7-10°C degrees cooler upon arrival at the lab than what had been measured at the field site.

Microscopic Counts

Each sample was homogenized in 10 mL of autoclaved Mushroom Spring water. A 10 μL subsample of the diluted 63°C mat sample was analyzed to obtain an estimate of the typical number of *Synechococcus* cells found in the top green layer of a #4 mat core. A Bright-line hemacytometer (Hausser Scientific, Horsham, PA) and microscope (Zeiss Axioskop 2 plus with an HBO 100 UV lamp) at 40X magnification were used to obtain cell counts. *Synechococcus* cells were identified by the property that they autofluoresce red with a rhodamine filter, and they are the only organisms in this community that do so (Ward et al., 1998, 2006).

Cultivation

The protocol for obtaining *Synechococcus* cultures and ensuring their predominance by single cyanobacterial strains is summarized in Appendix C, Figure C.1.

Gelrite Dilutions. Plates were prepared with Gelrite powder added to medium DHAY (Castenholz's medium D plus 0.1 g/L of sodium acetate trihydrate, 0.1 g/L

of yeast extract, and 5 mM HEPES buffer; pH set to 8.2) at a concentration of 1.5% (w/v). The mixture was then aliquoted into individual 50 mL glass bottles, at 20 mL per bottle, and then autoclaved. After cooling to 50°C in a water bath, media were inoculated with 5 mL of a prepared dilution of a known *Synechococcus* cell density, and then poured into 60 mm diameter Petri dishes in a laminar flow hood. After allowing the Gelrite to solidify, the plates were placed in sealed Ziploc bags with wetted paper towels and then incubated at 52°C under 50 μmol photons/m^2/sec of white fluorescent light. Plates were inspected every day over a one-month period for colony growth. Any isolated colonies that grew on high-dilution plates (diluted at least five orders of magnitude from the original mat sample) were picked with a sterile toothpick and placed in 2 mL of liquid medium DHAY.

Liquid Cultures. Colonies were originally suspended in 2 mL liquid medium DHAY in 20 mL glass test tubes with caps that were loosely fastened to allow gas exchange. After growing to stationary phase, the culture was scaled up by addition of 18 mL medium DHAY, then, after growth of this culture to stationary phase, by transferring to 80 mL fresh medium DHAY. This procedure prevented bottle-necking from occurring during transfer, i.e., at each step the higher volume was obtained by adding enough medium to the lower-volume culture to achieve the desired total volume. Incubation conditions were as described above and remained constant throughout this process.

Molecular Analyses

DNA Extraction. 1.5 mL of liquid culture were pelleted by centrifugation for 3 minutes at 4800 x g and re-suspended to a volume of 200 μL. DNA was then extracted and purified from the concentrated biomass using the Fastprep Cell Disrupter (Bio101 Savant Instruments, New York) for cell lysis and the FastDNA Spin kit (Molecular Biosciences) for purification of DNA, per the manufacturer's instructions.

DNA Amplification. Polymerase chain reaction (PCR) amplification was performed on extracted DNA to prepare for gene sequencing. For each 1 μL DNA used, 30.75 μL sterile, deionized water, 5 μL buffer II (100 mM Tris-HCl, pH 8.3, 500 mM KCl), 5 μL of MgCl buffer (25 mM), 5 μL bovine serum albumen, 1 μL dNTPs (nucleotide tri-phosphates, 10 mM), and 1 μL of each primer (50 ng/μL concentration) for the *psaA* gene were added (for A/B′-like *psaA* primers see Becraft et al., 2011). Then 0.25 μL Taq polymerase (5 units/μL) were pipetted into an Eppendorf tube and mixed using a vortex mixer to make a master mix, with 49 μL of the master mix and 1 μL of each DNA sample (concentration between 10-300 ng/μL) aliquoted into PCR tubes. After briefly centrifuging on a tabletop microcentrifuge to ensure that all ingredients were mixed, PCR was performed using an Applied Biosystems 2700 thermocycler with the program of 45 seconds at 92°C, 45 seconds at 53°C and

90 seconds at 72°C. Amplification was verified by electrophoresis using the protocol described in Becraft et al. (2011).

Sanger Sequencing and Ecotype Demarcation. Sanger sequencing was done at the Idaho State University Molecular Biosciences Core Facility. Ecotypes were determined by aligning *psaA* culture sequences to the environmental variation (Becraft et al., submitted). Cultures with a clean sequence (i.e., no ambiguous base calls and low background signal) that were 100% identical to a dominant variant (DV) representative of a predominant mat population were diluted to extinction in liquid medium DHAY, and the resulting highest-dilution subculture was prepared for high-throughput, Ti454-barcode sequencing (see methods below). Batch cultures with poor Sanger sequences (ambiguous base calls and/or high background signal) were re-diluted to extinction on plates, and the cultivation process was repeated by picking well-isolated colonies from high-dilution plates (see steps 2-5 in Appendix C, Figure C.1).

Ti454-barcode Sequencing. Cultures with clean Sanger sequences were grown to late exponential growth phase (between 2 x 10^7 and 5 x 10^7 cells/mL) and then 60 mL were pelleted for 30 minutes at 1000 x g in a Sorvall RC 5B Plus centrifuge. Cell pellets were flash frozen in liquid nitrogen and stored in a -80°C freezer before extracting DNA. In order to have enough DNA for both Ti454-barcode sequencing and genome sequencing, assuming that the purity tests at the Ti454 level were passed,

an enzymatic extraction approach was performed. Each frozen pellet was thawed and resuspended to 1 mL total volume in medium DHAY, then DNA was extracted by lysozyme/proteinase K lysis followed by a phenol/chloroform/isoamyl alcohol extraction (see DNA extraction protocol in Appendix C). RNA and other impurities were removed using the RNase I treatment protocol required by the Department of Energy Joint Genome Institute at http://my.jgi.doe.gov/general/protocols.html. DNA was quantified using a NanoDrop Spectrophotometer ND-1000 (NanoDrop Technologies, Wilmington, DE) and PCR amplified as described above to ensure that the DNA could be amplified. To evaluate heterotrophic contaminants, Ti454-barcode sequencing of 16S rRNA gene segments was performed according to the methodology posted on The Research and Testing Laboratory website (www.researchandtesting.com). Sequencing of *psaA* gene segments was performed as in Becraft et al. (submitted) to assess the complexity of the cultivated *Synechococcus* population.

To achieve an extra level of confidence regarding culture purity, all cultures were diluted to extinction in liquid medium DHAY and the most highly diluted subculture was sequenced again with Ti454-barcode technology. A culture was considered to be predominated by a single cyanobacterial strain if it contained a single DV *psaA* sequence, as well as closely related, less-abundant genetic variants with 1 or 2 randomly distributed nucleotide substitutions compared to the DV sequence, which

might have arisen during cultivation or represent sequencing error. This sequence-based criterion was also used to test the purity of the two *Synechococcus* strains (JA-3-3Ab and JA-2-3B′a (2-13)) that were cultivated by Allewalt et al.(2006), and whose genomes were obtained (Bhaya et al., 2007). The purity of an axenic *Synechococcus* strain (CIW-10) obtained by Kilian et al. (2007), that was derived from JA-2-3B′a (2-13), was also determined using this method.

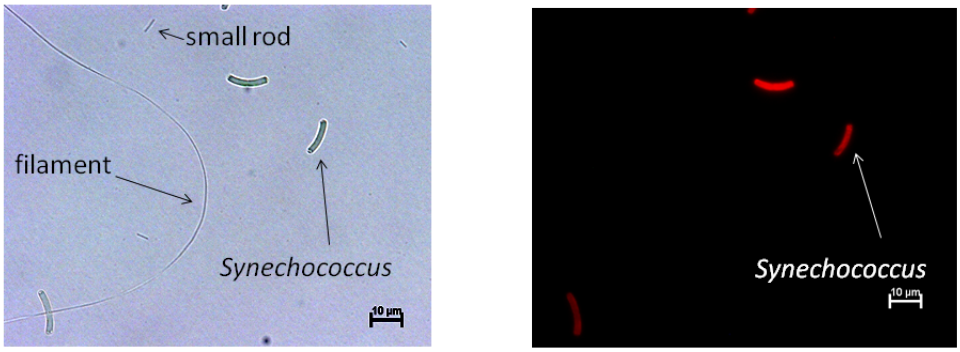
Growth Experiments

All growth experiments were conducted in an illuminated growth chamber that consisted of an aquarium of dimensions 152.4 cm x 15.2 cm x 38.1 cm that served as a water bath. The temperature was controlled with a PolyScience circulator, Series 7000 (Niles, Illinois). The chamber was illuminated with two identical ATI 60″ 6 x 80 W SunPower T5 high-output fluorescent fixtures (Denver, CO), one on each side of the aquarium. The fluorescent bulbs, 12 total, were of type F80W-T5-841-ECO (General Electric, Fairfield, CN). The growth medium used was medium DHAY. Prior to inoculation batch cultures were pre-grown to late exponential phase (to minimize lag phase), or to a density of approximately 2 x 10^7 cells/mL, at 52°C and a scalar irradiance of 50 μmol photons/m^2/sec of white fluorescent light. 100 mL total volume of each batch culture, set to a predetermined cell density, were aliquoted into 175 mL glass P/T culture tubes (BELLCO) and then capped with silicon sponge closures (Sigma-Aldrich) to allow for gas exchange. Different light

conditions, spanning the range of light intensities observed in nature, were achieved by applying various layers of neutral-density filter covering (GAM products, Los Angeles, CA) around the culture tubes. Light intensity was measured using a scalar irradiance probe, model QSL2100 from Biospherical Instruments (San Diego, CA), by submersing the probe in 100 mL of water prior to sterilizing the culture tubes. Light intensity and pH were measured before and after the experiment to ensure that these parameters had remained stable during the experiment. 1 mL samples were taken every 12 hours and frozen in a -80°C freezer after being fixed with glutaraldehyde (0.125% final concentration). Cell counts were obtained using flow cytometry (see details below). Each sample was filtered through a 70 μm screen-cap filter (Fisher Scientific) before analysis on the cell counter in order to eliminate clogging of the flow cell. Growth rates were determined by estimating log-linear slopes during exponential growth phase.

Quantification of Heterotrophic Contaminants. To investigate the possible influence of the growth of heterotrophic contaminants on *Synechococcus* growth, a preliminary experiment was conducted using the protocol described above with the following additional specifications. Temperature was set to 52°C. Batch cultures of one of the cultivated strains were grown, in duplicate, at scalar irradiances of 25, 125, 250, and 600 μmol photons/m^2/sec of white fluorescent light. The initial cell density of the *Synechococcus* population was set to 5.5 x 10^5 cells/mL and a BD-FACSAria

II flow cytometer and BD counting beads (BD Biosciences) were used to quantify cell densities. *Synechococcus* cells are typically 8-10 μm in length and microscopy revealed that they were noticeably longer than any of the heterotrophs that were detected in the batch culture and also able to pass through the screen-cap filter (see Figure 4.3). The *Synechococcus* cells also contain chlorophyll *a*, a pigment that is excited by the SYTO 17-A laser on the FACSAria II flow cytometer, and none of the heterotrophic contaminants have this pigment. Therefore, plots of forward scatter (a measure of cell size) versus SYTO 17-A signal (a measure of autofluorescence), and also forward scatter versus side-scatter (cell complexity) were analyzed in order to distinguish the heterotrophic cells from the *Synechococcus* cells (see Appendix C, Figure C.2).



(A) (B)

Figure 4.3. Microscopic images of the *Synechococcus* strain representative of PE A1 and heterotrophic contaminants. (A) Differential interference contrast microscopy. (B) Image from (A) using fluorescence microscopy. Scale bar is 10 μm.

Light Responses at 52 and 60°C. The light responses of three strains that are representative of three of the predominant A-like PEs in Mushroom Spring (PE A1, PE A4, and PE A14) were measured at 52 and 60°C. The growth-experiment protocol described above was followed with these additional specifications. To isolate light as a variable, a sterilized, cotton-plugged 7 mm glass tube was connected to a gas cylinder containing 6% CO_2 in air (GENDCO, Bozeman, MT) and used to bubble the cultures at a rate of approximately 1 bubble/second. To negate acidification effects, 26 mM $NaHCO_3$ was added to the growth medium (before autoclaving). The concentration of $NaHCO_3$ was determined in a preliminary experiment by bubbling liquid medium DHAY with 6% CO_2 in air at ∼1/bubble second and gradually adding more $NaHCO_3$ until the pH stabilized. These methods were applied at two different temperatures, 52 and 60°C. The initial cell density of the *Synechococcus* population was set to 1 x 10^6 cells/mL in all experiments.

A BD-FACSCanto flow cytometer (BD Biosciences) and CountBright Absolute counting beads (Invitrogen) were used to quantify cell densities. The results from the 60°C experiments were compared to the 60°C vertical distributions of the PEs that these strains represent (Figure 4.1B). The results from the 52 and 60°C experiments were compared to one another in order to investigate the effects of temperature on the fundamental light niche.

Dissolved Inorganic Carbon Effects on Light Response. To highlight the interconnectivity between light responses and the form of available DIC, the growth rates with respect to light intensity of two strains were measured when three different forms of DIC were provided. The corresponding genome sequences of the two strains chosen for this experiment revealed that one of the strains possessed an extra putative carbonic anhydrase gene that the other strain did not possess (M. Olsen and D.M. Ward, unpublished). This gene is known to produce an enzyme that catalyzes the bi-directional interconversion of CO_2 and water to bicarbonate. Hence, we hypothesized that the two strains would respond differently to different forms of DIC. The protocol described in the growth experiment subsection was followed with these additional specifications. The temperature was set to 52°C for all three experiments. In the first experiment, 6% CO_2 in air was provided, as described in the previous subsection. In the second experiment, the cultures were not bubbled with 6% CO_2, but instead filter-sterilized $NaHCO_3$ was added to medium DHAY after autoclaving (to 12 mM final concentration), creating a condition in which bicarbonate was the primary DIC source since medium DHAY is devoid of DIC. In the third experiment, the culture was not bubbled with 6% CO_2 and no additional bicarbonate was added, creating a condition in which CO_2 was limited by diffusion from the air above the culture.

Post-experiment Validation of *psaA* Ecotype

When cultures had completed exponential growth phase (cell density between 2 x 10^7 and 5 x 10^7 cells/mL), cells were harvested by centrifugation for 30 minutes at 1000 x g. The cell pellets were flash-frozen in liquid nitrogen and then stored in a -80°C freezer, and at a later time, DNA was extracted from the cell pellet. Sanger sequences were then obtained as described above to verify that the *psaA* genotype had not shifted during the experiment due to the selection of a rare *Synechococcus* variant with a different light adaptation.

Cell Morphology

To observe the cell morphologies of the heterotrophic contaminants and to possibly identify any morphological differences among the three strains, differential interference contrast and fluorescence microscopy were performed using a Nikon Eclipse 80i microscope with a Nikon Intensilight C-HGFl UV lamp. Near the end of exponential growth phase (cell densities between 2 x 10^7 and 5 x 10^7 cells/mL) of the experiments that were conducted at 60°C, photomicrographs of cells (Nikon DS-Ri1 camera) grown under low light(25 μmol photons/m^2/sec) or high light (600 μmol photons/m^2/sec) were obtained for each strain. Composite images were also obtained for each strain by combining cells that were grown under the two light conditions into a single sample (at an equal proportion). The purpose of obtaining the composite images was to seek evidence of morphological acclimation to light.

These same samples were also analyzed on the BD-FACSCanto flow cytometer. Cell size (forward scatter), complexity (side scatter), and autofluoresence were measured and compared to the microscopic observations.

Results

Nutrition Experiments

Preliminary experiments were conducted to evaluate whether the culture medium could be improved in terms of its ability to recover *Synechococcus* isolates representative of the most predominant PE populations. In a nutrient supplementation experiment, optimal growth of *Synechococcus* sp. strain JA-3-3Ab (*psaA* PE A1) occurred near the central well that contained a dense culture of *Synechococcus* cells, but mostly in the direction of wells containing acetate and yeast extract (Figure 4.2A). This suggested that these nutrients were stimulatory, at least at low concentration, and that other nutrients or quorum-sensing compounds might also be required. In a subsequent experiment in which sodium acetate and yeast extract were provided in a range of concentrations, the best growth occurred when both sodium acetate and yeast extract were added at a concentration of 0.01% (w/v) to medium DHAY (Figure 4.2B). The culture formed colonies on this medium when diluted 10^7-fold as compared to only 10^4-fold in medium DH. The findings from these two experiments resulted in a new protocol for isolation of *Synechococcus* spp. from Mushroom Spring. Specifically, compared to the former protocol (Allewalt et

al, 2006), medium DH was replaced with medium DHAY and floating filters were replaced with solid Gelrite media at a 1.5% (w/v) concentration.

Cultivation of *Synechococcus* Isolates from Mat Samples

Microscope counts revealed that ~1.8 x 10^8 *Synechococcus* cells were present in the top 1 mm green layer of a 63°C mat sample that was approximately 0.05 mL in volume. This is equivalent to ~3.6 x 10^9 cells/mL and is within an order of magnitude of previous *Synechococcus* cell counts that were obtained from Octopus Spring mat samples (Bauld and Brock, 1974; Ferris et al., 1996). Growth of *Synechococcus* colonies on the incubated plates was observed as early as day three for the low-dilution plates, and as late as day 21 for the high-dilution plates. The plated dilutions resulted in colony growth on all plates out to the 10^7-fold dilution for the 60°C sample, and out to the 10^8-fold dilution for the 63°C and 65°C samples. Hundreds of well-isolated colonies containing cells with the typical unicellular morphology of *Synechococcus* were picked from high-dilution plates, suspended in 2 mL liquid medium DHAY, and then gradually scaled up to larger volumes in preparation for sequencing.

Molecular and Morphological Descriptions of Cultures

Sanger and Ti-454 barcode sequencing revealed four cultures with *psaA* sequences representative of predominant A-like PEs that are differently distributed

in the vertical profile at 60°C: A1, A4, A6 and A14 (Figure 4.1B), which were assigned strain names 65AY6Li, 65AY6A5, 63AY4M2 and 60AY4M2, respectively. Barcode analyses showed that all were dominated by the DV expected for these PEs and their associated singleton variants (Table 4.1). For simplicity, these strains will be referred to by the *psaA* PEs that they represent. Barcode analyses also revealed that *Synechococcus* strain JA-3-3Ab previously cultivated by Allewalt et al. (2006), which is representative of PE A1, was heavily dominated by the expected DV of PE A1 and associated random singleton variants; a single sequence corresponding to PE A14 was detected (representing less than 0.1% of the variants in the culture). In contrast, previously cultivated *Synechococcus* strain JA-2-3B′a (2-13), which is not representative of a high-abundance *psaA* PE found *in situ* (Becraft et al, 2011; submitted), contained multiple high-frequency sequences representative of three different PEs (B′19, B′24, and B′11). Similar observations were made regarding the impurity of an axenic strain (CIW-10) that was derived from JA-2-3B′a (2-13). In fact, the barcode data revealed that the CIW-10 strain was dominated by PE B′24, which was detected in the parent culture (JA-2-3B′a (2-13)) but was not the dominant PE. Thus, while strain JA-3-3Ab appears to be predominantly unicyanobacterial, the JA-2-3B′a (2-13) and CIW-10 cultures we examined are not unicyanobacterial and appear to contain multiple PEs.

Microscopic observation showed that these cultures also contained heterotrophic

Table 4.1. Summary of *psaA* Ti454-barcode sequencing analyses of *Synechococcus* cultures. Sequences with systematic errors were excluded.

			dominant variant sequences[a]		other sequences[b]	
Strain name	total sequences	PE	DVs	%	total	%
JA-3-3Ab	1116	all	931	83.4		
		A1	930	83.3	185	16.6
		A14	1	0.1		
JA-2-3B′a	2054	all	1690	82.3		
(2-13)		**B′19**	1385	67.4	364	17.8
		B′24	300	14.6		
		B′11	5	0.2		
CIW-10	1528	all	1077	70.5		
		B′24	438	28.7	413	27
		†	639	41.8	38	2.5
65AY6Li	980	**A1**	780	79.6	200	20.4
65AY6A5	1325	all	1106	83.5		
		A4	1105	83.4	219	16.5
		A1	1	0.1		
63AY4M2	2027	**A6**	1756	86.6	271	13.4
60AY4M2	920	**A14**	756	82.2	164	17.8

Putative ecotypes (PEs) in bold are the dominant PEs in the culture.

[a] Number of sequences that are identical to the dominant variant (DV; identical sequence that made up the plurality of a putative ecotype (PE) clade) of the PE listed in the same row.

[b] Other sequences that are closely related to the DV sequence.

†A dominant sequence variant in the culture that is distinct from PE B′24 by 17 single-nucleotide polymorphisms.

contaminants (Figure 4.3). In addition to rod-shaped, red-autofluorescing *Synechococcus* cells approximately 8-10 μm in length, the cultures contained smaller rod-shaped cells approximately 2-5 μm in length and filamentous cells greater than 50 μm in length (Figure 4.3A). 16S rRNA barcode analyses (Appendix C, Table C.1) suggested that *Meiothermus* spp. were the most abundant heterotophic contaminants, comprising 10.69 to 58.74% of the total number of sequences analyzed in each culture, and are most likely represented by the small rods shown in Figure 4.3A. It

should be noted that *Caldilinea aerophila*-like sequences, which were detected in low abundance (<3.1% of total sequences analyzed in any of the strains) (Appendix C, Table C.1). Isolates of *Caldilinea aerophila* are filamentous, and can grow to lengths of over 200 μm. The filament shown in Figure 4.3A may be a representative of this organism.

Adaptive Light Responses

Effect of Heterotrophic Growth on *Synechococcus* Light Responses. A preliminary experiment was conducted to investigate the possibility that the measured light responses of *Synechococcus* strains might be secondary to effects of the heterotrophic contaminants. By taking advantage of the unique fluorescence and light-scatter characteristics of these populations, the growth of both populations was tracked. The results for the PE A1 strain are shown in Figure 4.4. During the first 12 hours of the experiment, the doubling rate of the heterotrophic population was nearly identical at all light intensities (∼12-13 doublings/day); meanwhile, the *Synechococcus* population remained in lag phase (see Figure 4.4). After 24 hours the heterotrophic population began to decrease when the *Synechococcus* population began to increase, and this was followed by the *Synechococcus* population overtaking the culture. The growth rate of the *Synechococcus* population varied with the light intensity, whereas the growth rate of the heterotrophic population did not.

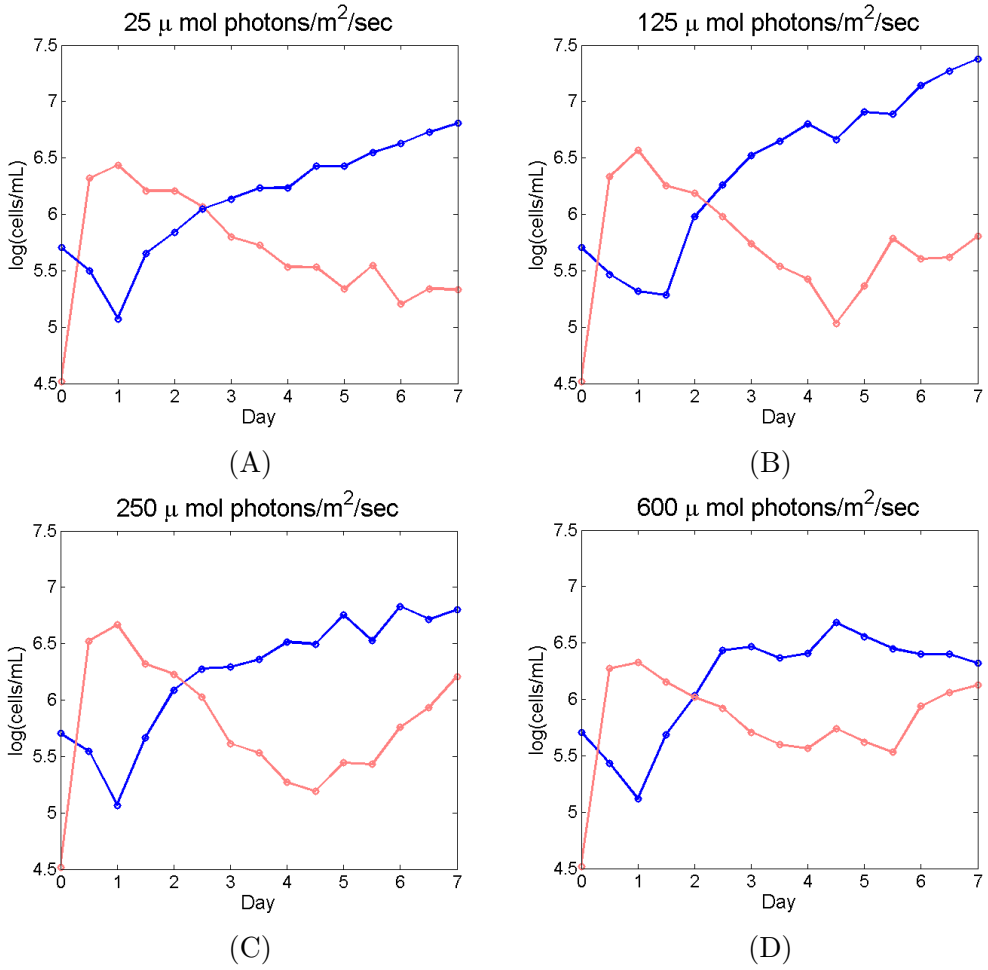
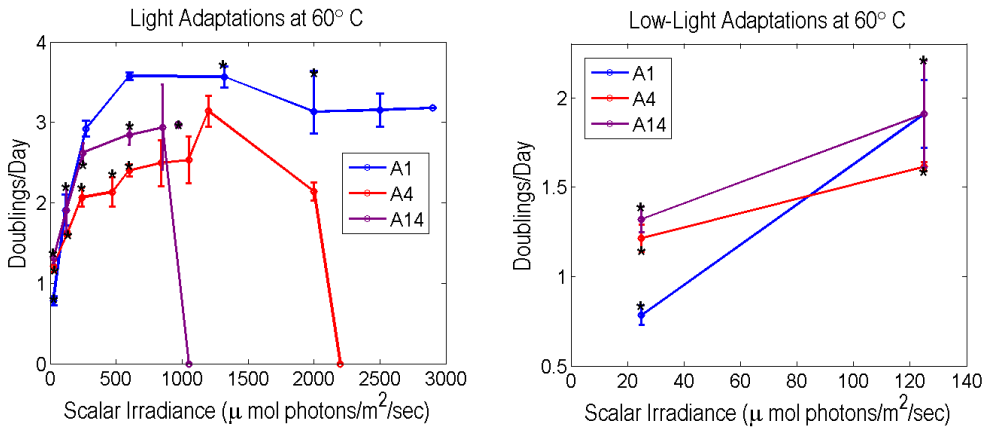


Figure 4.4. Heterotroph growth versus *Synechococcus* growth. Number of cells of the heterotrophic population (pink) and the *Synechococcus* 65AY6Li strain (PE A1) (blue) at four different light intensities, over a seven-day period at 52°C, and without any additional DIC provided to medium DHAY.

Light Responses of Different Strains at 60°C with 6% CO_2 in Air. The light adaptations of three *Synechococcus* strains are shown in Figure 4.5 and are summarized in Table 4.2. The highest scalar light intensity supporting growth for the PE A1 strain (Figure 4.5, blue) was at least 3000 μmol photons/m^2/sec at 60°C, the maximum intensity that could be achieved in the lab. The PE A4 strain (Figure



(A) (B)

Figure 4.5. Growth rates of *Synechococcus* isolates representative of predominant PEs as a function of light intensity when grown at 60°C and bubbled with 6% CO_2 in air. (A) All tested light intensities. (B) Low-light portion of (A) expanded. Error bars are range bars and asterisks represent the post-experiment samples for which *psaA* genotypes were confirmed.

4.5, red) was able to grow at 2200 μmol photons/m^2/sec, but not at 2500 μmol photons/m^2/sec. The PE A14 strain (Figure 4.5, purple) had the lowest light tolerance of all strains tested, and was able to grow at 850 μmol photons/m^2/sec, but consistent growth was not observed at 1050 μmol photons/m^2/sec. At that intensity the experimental results for the PE A14 strain were not repeatable – one of the replicates grew at this intensity after a long lag phase ($\sim$3.5 days compared to $\sim$1 day when grown at lower light intensities), but the other replicate did not. The growth rate was reported as zero in Figure 4.5 because the average growth rate in this instance does not accurately reflect the observed behavior. Neither replicate of the PE A14 strain grew at 1400 μmol photons/m^2/sec. A two-factor ANOVA anal-

Table 4.2. Summary of light responses of *Synechococcus* strains grown at 60 and 52°C and bubbled with 6% CO_2 in air.

Strain name	PE	Temperature/ dilution[a]	Upper light limit[b] 60°C	Optimal light intensity[b] 60°C	Upper light limit[b] 52°C	Optimal light intensity[b] 52°C
65AY6Li	A1	$65°/10^{-7}$	> 3000	600-1200	< 2300	125-1600
65AY6A5	A4	$65°/10^{-7}$	2200	1200	600	250
60AY4M2	A14	$60°/10^{-5}$	1050	600-850	1100	600

[a] Temperature of mat sample and dilution from which isolate originated.
[b] All units are scalar μmol photons/m^2/sec.

ysis confirmed that the PE A4 and PE A14 strains ($p < 0.05$) were better adapted to lower light intensities (25 μmol photons/m^2/sec) than the PE A1 strain (Figure 4.5B). On the otherhand, at 125 μmol photons/m^2/sec, significant differences in the growth rates of the three strains were not observed (see Appendix C, Table C.2 for times that were included in the analysis).

Sanger sequence analysis of post-experiment samples did not show any evidence of *psaA* PEs switching in the PE A1, A4, and A14 strains as a result of the imposed experimental conditions. On the other hand, post-experiment samples of the PE A6 strain did show a switch in genotype at two of the light intensities tested. Results for this mixed strain are presented in Appendix C, Figure C.3 and Table C.3.

Light Responses at Different Temperatures. Light responses of the three strains discussed above were also determined when these strains were grown at 52°C, and the results were compared to the 60°C results (see Figure 4.6 and Table 4.2). In Figure 4.7, surface plots of the results shown in Figure 4.6 are also provided.

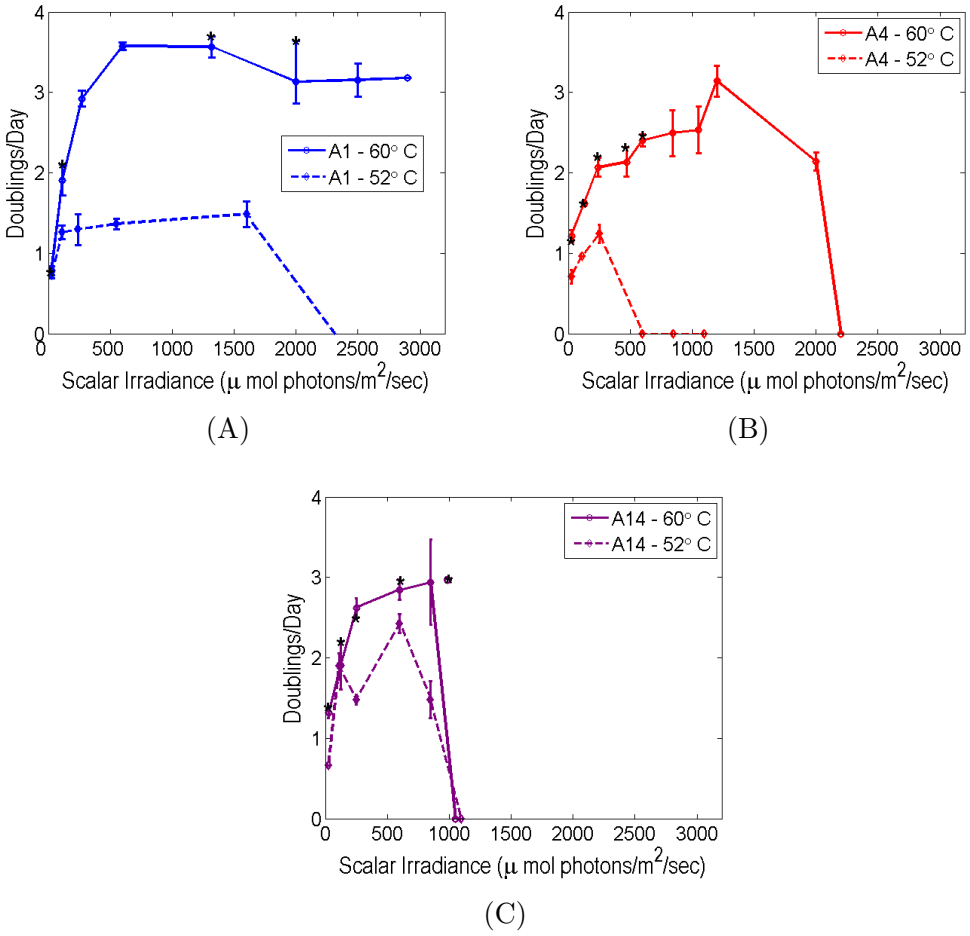
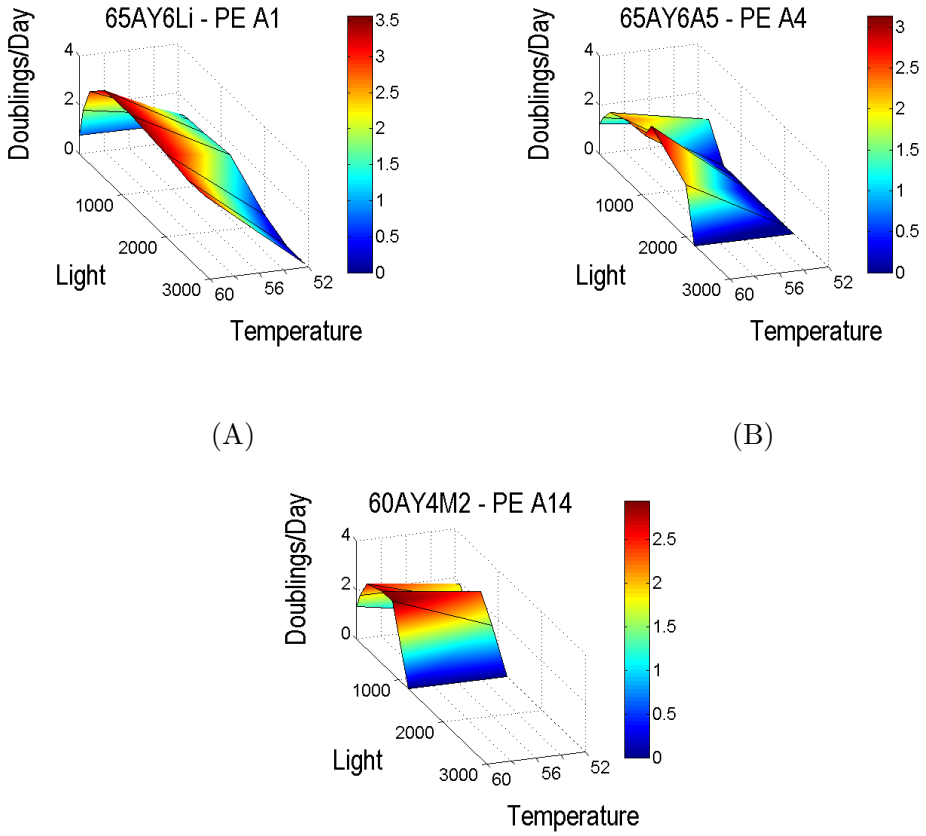


Figure 4.6. Growth rates of one strain of each of three dominant *Synechococcus* PEs (A1, A4, and A14) grown at 52 (dashed) and 60°C (solid) and bubbled with 6% CO_2 in air. (A) Strain representative of PE A1, (B) strain representative of PE A4, and (C) strain representative of PE A14.

When the temperature was increased from 52°C to 60°C all three strains exhibited higher growth rates at all light intensities. The PE A1 and PE A4 strains also exhibited a higher tolerance to light when incubated at 60°C. The growth rate at the optimal light intensity of the strain representative of PE A1 increased from 1.49 doublings/day to 3.6 doublings/day and the upper-light limit increased from



(A) (B)

(C)

Figure 4.7. Surface plots of the data presented in Figure 4.6 for each strain. (A) Strain representative of PE A1, (B) strain representative of PE A4, and (C) strain representative of PE A14. The units on the light axis are μmol photons/m^2/sec. The units on the temperature axis are degrees Celsius.

<2300 to >3000 μmol photons/m^2/sec (Figure 4.6A). Of the three strains, the strain representative of PE A4 showed the greatest improvement with respect to growth rate at the higher temperature, improving from the strain that had the lowest light tolerance at 52°C (it would grow at 450 but not at 600 μmol photons/m2/sec) to the strain that had second-highest light tolerance at 60°C (it would grow at 2200

but not 2500 μmol photons/m^2/sec; Figure 4.6B). The strain representative of PE A14 showed the least overall improvement of the three strains when the temperature was increased (Figure 4.6C). Its growth rate at the optimal intensity improved from 1.9 to 2.9 doublings/day at the higher temperature, but its upper-light limit did not increase with the increase in temperature.

Light Responses Under Different DIC Conditions. Because light and DIC utilization are linked, we examined the light responses of the PE A1 and PE A6 strains under conditions in which the supply of DIC was limited by diffusion, which may be more natural *in situ*, especially at high light intensity when photosynthesis is maximal. Under these conditions, the PE A1 strain had a higher growth rate than the PE A6 strain at all light intensities tested, as was the case when a continuous supply of CO_2 was provided, but the growth rates and light tolerances for both strains were much lower under the CO_2-limiting condition (compare solid and dashed lines of the same color in Figure 4.8 and see Table 4.3 for summary). Even at the extreme

Table 4.3. Summary of light responses of *Synechococcus* strains grown at 52°C with various forms of dissolved inorganic carbon (DIC) provided.

Strain name	*psaA* PE	6% CO_2 in air Upper-light limit[a]/ Optimal growth[b]	12 mM $NaHCO_3$ Upper-light limit/ Optimal growth	No DIC Upper-light limit/ Optimal growth
65AY6Li	A1	2300/1600	1050/900	1100/260
63AY4M2	A6	600-1500/550	550/125	550/125

[a] All units are scalar μmol photons/m^2/sec.
[b] Light intensity at which the strain grew the fastest.

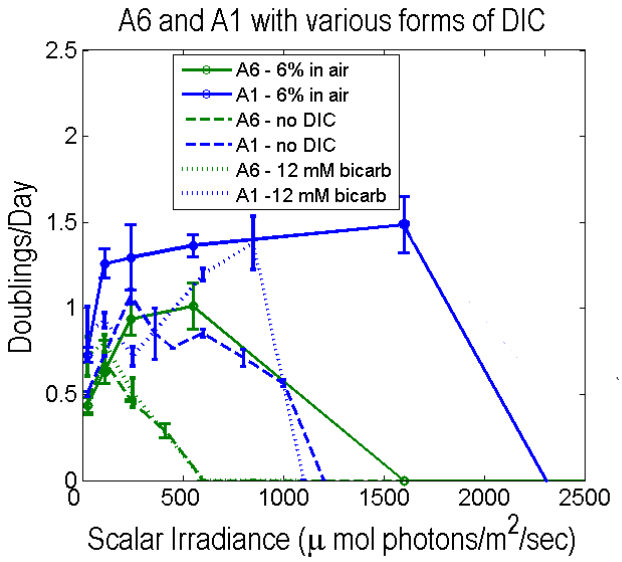


Figure 4.8. Growth rates of *Synechococcus* strains representative of PEs A1 and A6 at 52°C as a function of light intensity in medium DHAY under different DIC conditions: no additional dissolved inorganic carbon (DIC) (dotted), with 12 mM filter-sterilized $NaHCO_3$ added (dashed), with bubbling with 6% CO_2 in air (solid) and 26 mM $NaHCO_3$ to buffer pH.

high pHs measured in the mat (pH 9.4), the predominant DIC species should be bicarbonate (Stumm and Morgan, 2012). Thus, since the PE A1 strain possesses an extra putative carbonic anhydrase gene that is not found in the PE A6 strain, we also tested whether addition of bicarbonate led to a growth advantage under conditions in which CO_2 was limited by diffusion. By comparison of dotted and dashed lines of the same color in Figure 4.8, it can be seen that the growth rate of the PE A1 strain, but not the PE A6 strain, was increased by the addition of

bicarbonate at all light intensities tested except 250 μmol photons/m^2/sec. Post-experiment validation of *psaA* genotype has not yet been performed on the batch cultures that were grown under these conditions.

Acclimative Light Responses

Evidence of *Synechococcus* cells acclimated to different light conditions was provided by comparing photomicrographs and flow cytometry output of cells grown under low-light and high-light conditions. Cells grown under these light conditions could be resolved by microscopy and flow cytometry after mixing them in equal proportions. As shown for the PE A1 strain in Figure 4.9A and B, cells grown under low light (left column) exhibited higher chlorophyll *a* concentrations (based on autofluorescence levels) compared to cells grown under high light (middle column), and this difference can be observed in the images of the composite sample (right column).

The different chlorophyll *a* concentrations of cells grown at the two light intensities can also be observed in the flow cytometer output (Figure 4.9C), where the cells grown under low light (left column, green) have a higher fluorescence signal (PerCP-Cy5-5-A) than cells grown under high light (middle column, red). Cells grown at high light were also longer than cells grown at low light. Similar results were observed for strains of PEs A4 and A14, as shown for analyses of combined low-light and high-light samples in Figure 4.10. The chlorophyll *a* differences for

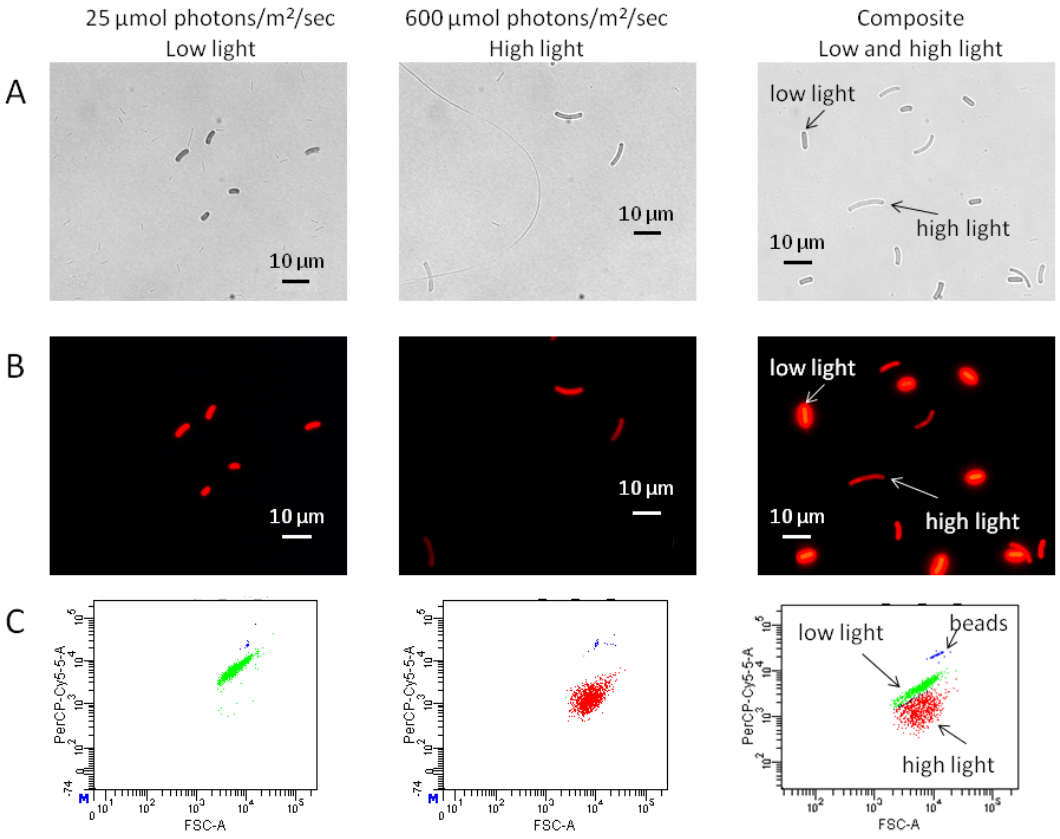


Figure 4.9. Microscopic and flow cytometric analyses of the *Synechococcus* PE A1 strain grown at 60°C and bubbled with 6% CO_2 in air, under a low-light (25 μmol photons/m^2/sec, left column) and a high-light condition (600 μmol photons/m^2/sec, middle column). (A) Differential interference contrast microscopy of samples that were collected at the end of exponential growth phase. Scale bar is 10 μm. (B) Images from (A) using fluorescence microscopy. (C) Scatter plots from BD-FACSCanto flow cytometer of forward scatter (cell size, horizontal axis) versus fluorescence signal (PerCP-Cy5-5-A, vertical axis) of samples shown in (A) and (B). Cells grown under low light are represented by the green data points and cells grown under high light are represented by the red data points. The blue data points represent the fluorescent counting beads that serve as a fluorescence control.

cells grown under the two light conditions were the most extreme for the PE A14 strain. This is depicted in Figures 4.10A and 4.10B, in which the high-light grown cells appear to be severely photo-bleached in 4.10A, and this corresponds to the

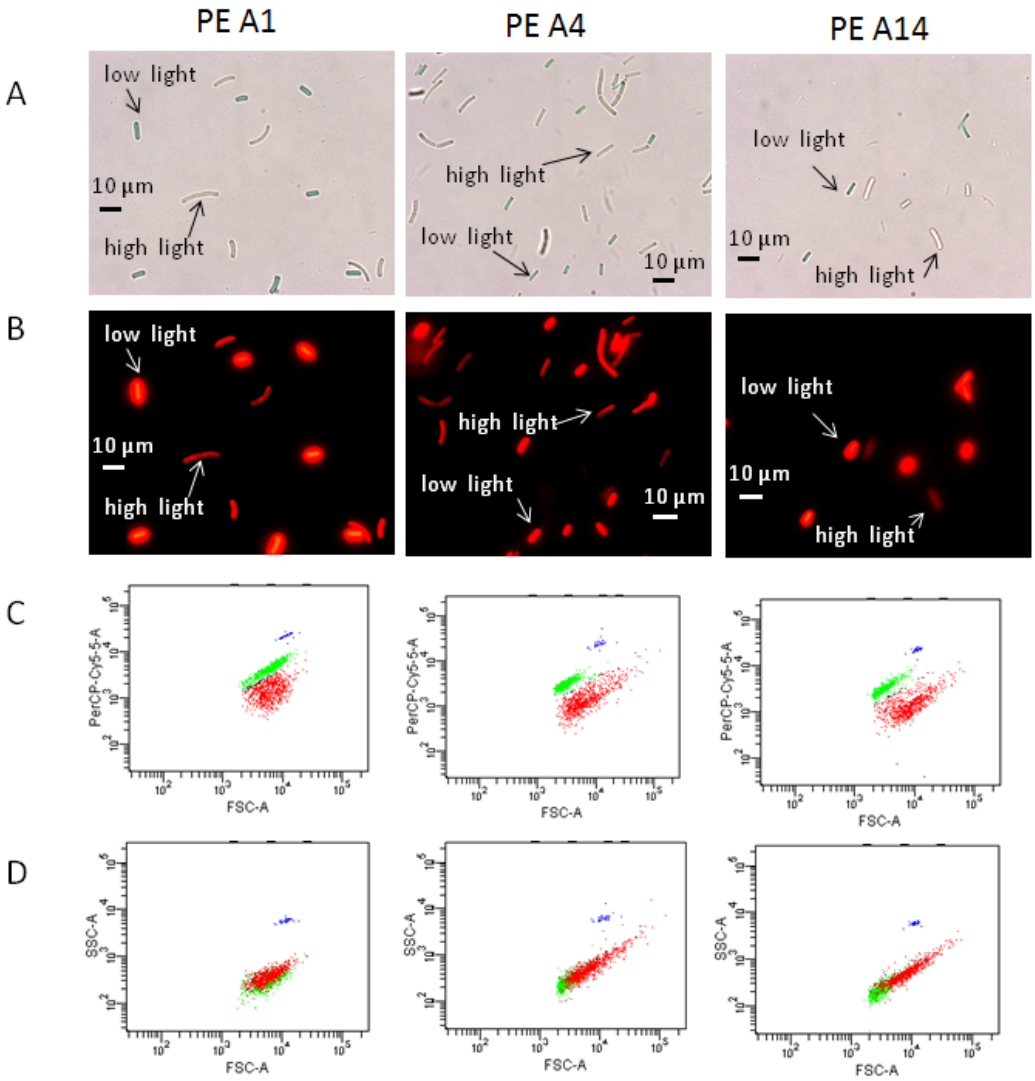


Figure 4.10. Microscopic and flow cytometric analyses of the strains representative of PE A1, PE A4, and PE A14 grown at 60°C and bubbled with 6% CO_2 in air, under a low-light (25 μmol photons/m^2/sec) and a high-light condition(600 μmol photons/m^2/sec). (A) Differential interference contrast microscopy of samples that were collected at the end of exponential growth phase. Scale bar is 10μm. (B) Images from (A) using fluorescence microscopy. (C) Scatter plots from BD-FACSCanto flow cytometer of forward scatter (cell size, horizontal axis) versus fluorescence signal (PerCP-Cy5-5-A, vertical axis) of samples shown in (A) and (B). Cells grown under low light are represented by the green data points and cells grown under high light are represented by the red data points. The blue data points represent the fluorescent counting beads that serve as a fluorescence control. (D) Scatter plots of forward scatter (horizontal axis) versus side scatter (complexity, vertical axis) samples shown in (A) and (B).

low autofluorescence signal of these cells in 4.10B. Cells from these strains also appear to elongate at higher light conditions, and this can be observed in both the microscopic and flow cytometer output. In the flow cytometer output this behavior can be identified by observing the right-shifted forward scatter signal (FSC-A) of the high-light cells (red) compared to the low-light cells (green) (Figures 4.10C and 4.10D). Microscopy also revealed that a large percentage PE A4 and A14 cells (roughly 10 and 25%, respectively), but not the PE A1 cells ($< 1\%$), appeared to be unable to separate from their parent cell after division (Appendix C, Figure C.4). We hypothesize that the unseparated cells may be the reason for the observed upward shift in side scatter for the PE A4 and PE A14 cells, but not for the PE A1 cells, that were grown under high light (Figure 4.10D). Statistical analyses for these observations are in progress.

Discussion

Cultivation Progress

The protocol applied here regularly produced colony growth from inocula that were three orders of magnitude more dilute than what had been achieved by Allewalt et al. (2006). Assuming that the direct cell count of the 63°C mat sample is representative of the other mat samples, <1% of the inoculated cells from 60°C, ∼5 - 20% from 63°C, and ∼50 - 100% from 65°C formed colonies on the plated dilution series. Additionally, at least at the *psaA* locus, there is strong evidence that the

isolates that were cultivated are representative of predominant *Synechococcus* PEs in Mushroom Spring (Becraft et al., 2011; submitted), with four of the five most abundant *psaA* PEs found *in situ* at 60°C being represented. PE B′9 is the fifth PE detected in significant abundance *in situ*, but a strain representative of this *psaA* PE has not yet been cultivated. The inability to cultivate isolates representative of PE B′9 could explain the disparity in colony forming units that was observed in dilutions of mat samples collected from different temperature sites. While PE B′9 is present at high relative abundance at 60°C, its relative abundance *in situ* decreases as temperature increases (Becraft et al., submitted), possibly explaining the better colony yield on plated dilutions of the 63°C and 65°C mat samples (which would have fewer and fewer PE B′9-like cells, respectively). As shown in Figure 4.1B, PE B′9 is also most abundant near the mat surface at 60°C, where it is exposed to high light, however low-light conditions were used during laboratory selection ($\sim$50 μmol photons/m^2/sec). In an earlier study, high-light adapted strains of *Prochlorococcus* spp. were unable to grow when low-light conditions were provided in culture (Moore et al., 1998). Hence, we hypothesize that cells representative of the PE B′9 population may have a higher low-light threshold than A-like cells, which may explain why incubating under low-light conditions has not yielded any B′9-like isolates (see Chapter 5).

The challenge of obtaining a pure, unicyanobacterial, laboratory isolate was also

documented here. Ti454-barcode data suggested the cultivated isolate representative of PE A6 was predominated by a single *psaA* sequence, but sequence data from samples collected at the end of the incubation period in the 60°C light response experiment provided evidence that the *psaA* genotype in this culture had switched under the experimental conditions. This could have been the result of a rare organism of another PE in the culture going undetected in the Ti454-barcode analysis, and then overtaking the batch culture when conditions were more favorable for this organism due to competitive exclusion (Santegoeds et al., 1996). 2027 sequences were included in the Ti454-barcode analysis of this PE A6 isolate (Table 4.1), but the DNA extracted from this strain was from a sample containing $\sim 10^8$ cells. Hence, cells representative of other PEs would not have been detected unless they comprised >0.002% of the *Synechococcus* cells in the culture. An alternative explanation is that cross-contamination with another strain in the laboratory occurred at some time after the initial barcode purity analysis was conducted. The DIC experiments performed at 52°C involving the PE A6 strain were done prior to the 60°C experiments. Therefore, samples collected at the end of the DIC experiments are currently being sequenced in an attempt to pinpoint the time of contamination, and to determine if the results from these experiments are valid. Also, at the time DNA was extracted for Ti454-barcode sequencing, freezer stocks of the PE A6 strain were prepared. To further investigate this contamination issue one of the PE A6 freezer-stocked samples has recently been revived so that Ti454-barcode sequencing

can be conducted before and after the strain has been exposed to different light intensities.

Another cultivation issue that remains unsolved is the inability to cultivate axenic strains of *Synechococcus* spp. from these hot spring systems. Ti454-barcode data (Table 4.1) suggested that the strains investigated here are predominated by a single *psaA* sequence, but, as in previous isolation attempts (Allewalt et al., 2006), heterotrophic contaminants, primarily *Meiothermus* spp., remain in abundance in the batch cultures. One hypothesis for why the new protocol has had such a positive effect on cultivation is that the added carbon sources (sodium acetate and yeast extract) are substrates for *Meiothermus* spp., which in turn feed the *Synechococcus* spp. with their byproducts. This cross-feeding hypothesis is not a novel one (Pfeiffer and Bonhoeffer, 2004; Ward et al., 2012; Weltzer and Miller, 2013) and analyses via flow cytometry, as shown in Figure 4.4, support this theory. We have shown that the heterotrophic contaminants do not have an effect on the reported light responses as the growth rates of the heterotrophic species over the first 12 hours after transfer were nearly constant over all light intensities tested (ranging from 12-13 doublings per day). Consequently, we predict that an axenic *Synechococcus* culture, in which the metabolites supplied by the heterotrophs are replaced with the appropriate nutrients in the growth medium, would have a similar light adaptation to that of the corresponding mixed culture. An alternative hypothesis to cross-feeding is that the heterotophs may consume byproducts of *Synechococcus*, such as oxygen, which

could become toxic at high levels (Sakamoto et al., 1998). To test these hypotheses, samples that were collected at different time points of the 25 μmol photons/m^2/sec experiment are currently being analyzed for metabolites.

Light-adapted Strains

The laboratory conditions provided in the 60°C experiment included 6% CO_2 bubbled in air and continuously supplied light. Constant temperatures, continuous light, and saturating levels of CO_2 were imposed to isolate light as a variable but these conditions are unnatural *in situ*. For this reason we consider the results of this experiment as describing the fundamental light niches (Hutchinson, 1957) at 60°C of the respective strains. The fundamental light niche of the PE A1 strain (Figure 4.5, blue) corresponds to that of a generalist species. The PE A1 strain appeared to be very fit at all light intensities between 600 and 3000 μmol photons/m^2/sec. The PE A4 strain (Figure 4.5, red) exhibited a narrower fundamental light niche and lower light tolerance than the PE A1 strain (the growth rate was ~60% of that of the PE A1 strain at light intensities between 125 and 2200 μmol photons/m^2/sec). Only at the lowest light intensity tested (25 μmol photons/m^2/sec) did the PE A4 strain have a significantly higher growth rate than the PE A1 strain ($p<0.05$). Conversely, relative to the PE A1 and A4 strains, the PE A14 strain (Figure 4.5, purple) exhibited a fundamental light niche that could be categorized as that of a specialist – having superior fitness with respect to the PE A4 strain in the 125-850 μmol photons/m^2/sec

range, but intolerant of intensities above 1050 μmol photons/m^2/sec. At the lowest light intensity tested (25 μmol photons/m^2/sec) the growth rate of the PE A14 strain was higher than the PE A1 strain but comparable to the PE A4 strain. At the 1050 μmol photons/m^2/sec condition, one of the replicates had a doubling rate of 2.77 cell divisions/day and the other replicate did not grow at all. Sequence data obtained from a post-experiment sample of the replicate that grew at 1050 μmol photons/m^2/sec confirmed that the *psaA* genotype had not shifted and was still PE A14-like. This finding suggests that the PE A14 strain is not contaminated with a *Synechococcus* population that is high-light adapted, and that the irreproducibility of this result may be due to an acclimative response. For all strains investigated in this study the cells used to inoculate cultures were pregrown at 50 μmol photons/m^2/sec before shifting to higher or lower light intensities, and it was observed that the acclimation period (lag phase) increased as the strains were shifted to intensities that approached their upper-light limit. In fact, one explanation for the replicates of the PE A14 strain responding differently near the upper-light limit is that, in one replicate, a few cells were able to acclimate before perishing, and in the other replicate this did not occur.

At higher light intensities (above 600 μmol photons/m^2/sec), cells from all strains appeared to contain less chlorophyll *a* and to elongate (Figure 4.9 and 4.10). Also, the high-light samples collected from the PE A4 and A14 strains had a larger number of cells that had divided, but were unable to separate from their parent cell

(Appendix C, Figure C.4). Approximately 10% of the PE A4 and 25% of the PE A14 cells exhibited this behavior, compared to <1% PE A1 cells. This cell arrangement possibly suggests a disruption in the cell-division cycle (Jacquet et al., 2001), which we hypothesize may be due to an upper-light limit being approached.

Collectively, these results suggest that strains representative of the predominant *psaA* putative *Synechococcus* ecological species in Mushroom Spring have distinct fundamental light niches. Moreover, the observed light niches correspond to *in situ* vertical distributions of the PEs (Becraft et al., 2011; submitted). That is, PE A1 is the most abundant A-like PE in the upper layers of the mat at 60°C and hence regularly receives the highest light intensities *in situ* (Becraft et al., submitted); correspondingly, the strain with the highest light tolerance among the three strains investigated was the strain representative of PE A1. Alternatively, *psaA* PEs A4 and A14 are predominantly found lower (vertically) in the mat at 60°C, and the strains representative of these PEs were shown to have lower-light intensity tolerances than the PE A1 strain, and to also have higher growth rates at a low-light intensity. These findings are comparable to distributions of *Prochlorococcus* ecotypes found with depth in the water column in marine environments. Using variation in the 16S rRNA gene it was found that cultivated *Prochlorococcus* strains could be divided into low-light and high-light ecotypes (Moore et al., 1998), and then in a later study (West and Scanlan, 1999), it was found that the lower light ecotypes were found

primarily in the deeper part of the photic zone, whereas the high-light ecotypes were primarily found nearer to the surface waters.

Of the three strains, the PE A1 strain had the highest growth rate at all light intensities (performance) above 125 μmol photons/m^2/sec, yet it also appeared to have the widest fundamental light niche (tolerance), with respect to scalar irradiance (Figure 4.5). One explanation for why a minimal performance/tolerance tradeoff is observed between the PE A1 strain and the other strains is that the *psaA* locus does not offer enough molecular resolution to demarcate all ecological species, and this observation is an artifact of the PE A1 strain being comprised of more than a single A1-like ecological species. That is, it grows better at most light intensities because multiple ecological populations with distinct light adaptations contribute to the observed fundamental light niche. It should be noted that Mushroom Spring vertical distributions at 63°C (Becraft et al., submitted) suggest that PE A1 has a bimodal distribution at this higher temperature, so that this lumping artifact could be true in nature as well as in culture.

Another explanation for the observed minimal tradeoff is that Figure 4.5A is misleading without considering the natural environmental history. Although Figure 4.5B depicts only approximately 5% of the light intensity range that is observed in nature, it has been shown that low-light conditions are much more common than high-light conditions (Nowack et al., submitted; Chapter 2). In fact, low light was predicted to be a better growing condition than high light in a mathematical model

based on *in situ* light data collected in the vicinity of Mushroom Spring (Nowack et al., submitted; Chapter 2). Low light intensities occur every day, but high light is only received a few hours a day and only over a few months of the year (and only on clear days). As a result, the amount of time, over the course of a year, that the environment experiences low light is significantly greater than the amount of time the environment spends at higher light conditions. Therefore, one hypothesis is that adaptations might be based on the commonality of light level, and in nature the light niche has been partitioned, with PEs A4 and A14 thriving under low-light conditions and PE A1 thriving under high-light conditions. If this is indeed the case, the results presented in Chapter 2 provide a possible explanation for why the growth rate of the PE A1 strain is significantly higher than the other strains at the higher light conditions – that is, in order for an organism to thrive at conditions that occur relatively infrequently, the organism must have enhanced fitness when those rare conditions are present.

Interconnectivity of the Dimensions of the Fundamental Light Niche

We reiterate that the laboratory conditions used to measure the fundamental light niches at 60°C do not attempt to reproduce the *in situ* environmental reality, but were used to isolate light as an experimental variable. The results presented

here suggest that in order to understand the true, realized light niche of the *Synechococcus* PEs found in nature, at least two other environmental conditions should be considered: temperature and DIC availability. First, temperature was revealed to have an effect on light response as well, which can be observed by comparing the light responses of the strains that were grown at different temperatures (Figures 4.6 and 4.7). The temperature results also suggest that the strains representative of PEs A1 and A4 may be adapted to temperatures of 60°C or higher and the strain representative of PE A14 may be approaching its optimal temperature at 60°C. It should be noted that the strains representative of PEs A1 and A4 were both cultivated from 65°C mat samples, and the strain representative of PE A14 was cultivated from a 60°C sample. Plans have been made to collect light-response data at other temperatures, specifically, at a lower temperature (∼45°C), an intermediate temperature (∼56°C), and higher temperatures (63°C, 65°C and 70°C). With these data in hand, surface plots as shown in Figure 4.7 for the two temperatures discussed here, can be used to illustrate the full effect that temperature has on the fundamental light niche of these three and other strains. And second, the significance of DIC availability is apparent in Figure 4.8 where it is shown that the PE A1 and A6 strains have very different light responses when saturating CO_2 levels are provided versus when they are not.

Comparative genomics analyses (M. Olsen and D.M. Ward, unpublished) have

revealed that the strains representative of PE A4 and PE A14 have putative light-harvesting antennae genes that the strain representative of PE A1 does not have. This observation has led to the hypothesis that light quality, in terms of the availability of certain wavelengths of light, may also be an important dimension of the fundamental light niche. Experiments to test this hypothesis, employing filters to reproduce the real light quality at different depths in the mat, have been proposed and hopefully will be conducted in the near future by other Ward lab members. Collectively, these findings suggest that the realized light niches of the abundant Mushroom Spring *Synechococcus* PEs might be quite different from the fundamental light niches of representative cultivated strains.

As mentioned above, a comparative genomics analyses of these strains (M. Olsen and D.M. Ward, unpublished) is currently underway. The gene contents of the strains are being analyzed and the goal is to link genes to functions, in order to possibly identify genes that are responsible for certain ecological specializations, such as light adaptations. Thus far the phenotypic analyses have directed the genomic analyses in some cases, but we envision this to be a bidirectional process, in which the comparative genomics drives future phenotypic studies and vice versa. Two examples have been highlighted here. First, the genomic data revealed that the PE A1 strain had an extra putative carbonic anhydrase gene that the PE A6 strain did not. This led to testing the hypothesis that the PE A1 strain, but not the PE A6 strain, would have a higher growth rate when $NaHCO_3$ was provided in the growth

medium (compared to when CO_2 availability was limited by diffusion). Second, the differences among strains with respect to putative light-harvesting genes has been noted, and future experiments to investigate wavelength adaptations have been planned. Another goal of the combined phenotypic/comparative genomic analyses is to contribute to identifying a universal microbial species definition. Is there a single gene whose diversity can be used to demarcate the fundamental species-like units in a given community, or is a multi-locus approach required (Retchless and Lawrence, 2007; Luo et al., 2011; Melendrez et al., in preparation)?

Concluding Remarks

In summary, evidence has been provided that supports the hypothesis that closely related *Synechococcus* spp. with distinct light adaptations exist within the Mushroom Spring microbial mat community, and that the *psaA* PEs may partition the light niche. The results also provided clear evidence that several environmental parameters comprise the fundamental light niche, and highlighted the consequences of isolating a single environmental variable during laboratory experiments. All of the strains discussed here are identical or nearly identical at the 16S rRNA locus, yet the observed light adaptations are not reflective of what is expected of ecologically interchangeable populations (Cohan and Perry, 2007). This provides an example of how increased molecular resolution can be used to discern ecological species that are lumped by more conserved molecular markers. Furthermore, by analyzing sequence

diversity at the *psaA* locus we found that the B′ 16S rRNA isolate from Allewalt et al. (2006) is not representative of a predominant *psaA* PE and also is not unicyanobacterial at the *psaA* locus. The suspicions regarding the B′ strain arose from the fact that the reported light adaptations (Kilian et al., 2007) did not match what the *in situ* vertical distributions (Becraft et al., submitted) suggested about its light niche. Therefore, we suggest that investigators using this B′ strain take these facts into consideration when designing experiments and interpreting results. In a future experiment we plan to compare phenotypic properties of strains within *psaA* PEs to understand if even more molecular resolution and/or a multi-locus approach is required in order to identify the true, fundamental species-like units that would allow predictions of an organism's phenotypic properties from its genotype.

Acknowledgements

We thank George Schaible for his assistance with the molecular work and Jennifer Weeding for performing the statistical analyses. U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Genomic Science Program 395 (GSP). This contribution originates from the GSP Foundational Scientific Focus Area (FSFA) at the Pacific Northwest National Laboratory (PNNL) under contract 112443. It was also supported by Montana Agricultural Experiment Station project 911352. Genomic sequencing was performed under Joint Genome Institute Community Sequencing Project NPUSR006316. This study

was conducted under Yellowstone National Park research permits YELL-0129 and 5494 (D.M.W.), and we appreciate the assistance from National Park Service personnel.

CHAPTER 5

CULTIVATION AND CHARACTERIZATION OF MULTIPLE STRAINS OF DIFFERENT PUTATIVE *SYNECHOCOCCUS* SPECIES

Funding from a Department of Energy Joint Genome Institute Community Sequencing Program was awarded to David Ward (Land Resources and Environmental Sciences Department, Montana State University) and colleagues to obtain 18 genome sequences of closely-related *Synechococcus* strains cultivated from Mushroom Spring, Yellowstone National Park (YNP). The overarching goal of this ongoing cultivation and genome sequencing project is to further test the hypothesis that the sequence diversity and distributions that have been observed *in situ* (Ferris and Ward, 1997; Ramsing et al., 2000; Ferris et al., 2003; Ward et al., 2006; Becraft et al., 2011; submitted) are due to the existence of distinct ecological species adapted to different environments (Ward, 1998; Ward and Cohan, 2005). To achieve this goal, it was proposed to cultivate three unicyanobacterial *Synechococcus* strains within each of six predominant putative ecological species populations (putative ecotypes, PEs). PEs were predicted by Ecotype Simulation (a theory-based evolutionary simulation algorithm (Koeppel et al., 2008)) from *psaA* sequence variation found in mat samples collected from Mushroom Spring (Becraft et al., 2011; submitted), and are hypothesized to represent the ecologically distinct populations, or species (Cohan and Perry, 2007), in this hot spring community. Obtaining replicate isolates

representative of the predominant *psaA* PEs found *in situ* and their corresponding genome sequences has allowed comparative phenotypic and genotypic analyses of strains within and among PEs to be conducted.

The first section of this chapter is devoted to describing the cultivation methods that were used to obtain the targeted isolates. In the second section, the light adaptations of strains within predicted species populations are compared in order to test the hypothesis that strains within *psaA* PEs exhibit the ecological interchangeability that is expected of species (Cohan and Perry, 2007).

Using Environmental Context to Obtain Isolates Representative of Predominant PEs

Introduction

Ti454-barcode sequencing revealed that in the 60°C Mushroom Spring microbial mat the predominant *psaA* PEs were A1, A4, A6, A14, and B′9 (see Figure 4.1B), therefore three cultivated strains representative of each of these PEs were targeted for cultivation. Strains representative of the 16S rRNA A′-like population were also of interest, as it has been shown to be a predominant mat population at higher temperatures (Ferris and Ward, 1997), and the Ward lab did not have a representative isolate with this 16S rRNA genotype. Cultivation of strains representative of the predominant PEs found in these hot spring systems has proven to be difficult (Allewalt, 2004; Allewalt et al., 2006; also see Chapter 4 introduction), and has been

an evolving process that has taken place over a period of approximately five years. My cultivation efforts will be discussed in this section, and present and future Ward lab members will continue these efforts to work towards satisfying the goals of the community sequencing project.

By collecting mat samples from 60, 63, and 65°C sites in Mushroom Spring and applying the Gelrite protocol with medium DHAY as described in Chapter 4, 11 of the 18 targeted strains were obtained. Four of these strains were described in Chapter 4 and the others will be described in this chapter. Although this first attempt was considered highly successful, only one representative each of an A′-like PE and PE A1, and no representatives of PE B′9, were obtained. This was not surprising as a single laboratory environment was being provided during selection (52°C in ∼50 μmol photons/m^2/sec of white fluorescent light), and *in situ* distributions have suggested that strains representative of the targeted PEs may have different adaptations to temperature and light (Allewalt et al., 2006; Becraft et al., 2011; submitted). Therefore, the cultivation attempts that followed were structured around exploiting the distribution data that had been collected from Mushroom Spring in the past. It was hypothesized that by allowing the species distribution data to guide the selection of the sampling sites and laboratory incubation conditions, isolates representative of the targeted PEs would be obtained.

Species distribution data from (i) mat samples collected along the thermal gradient in Mushroom Spring and Octopus Spring (Ferris and Ward, 1997; Ward et al.,

2006), (ii) mat samples collected along the vertical gradient at a 60°C site in Mushroom Spring (Figure 4.1B), and (iii) effluent-channel water samples collected along the thermal gradient above the mat, at different times of the year, were consulted (Figure 5.1) to guide cultivation efforts. The mat distribution data (Ferris and Ward,

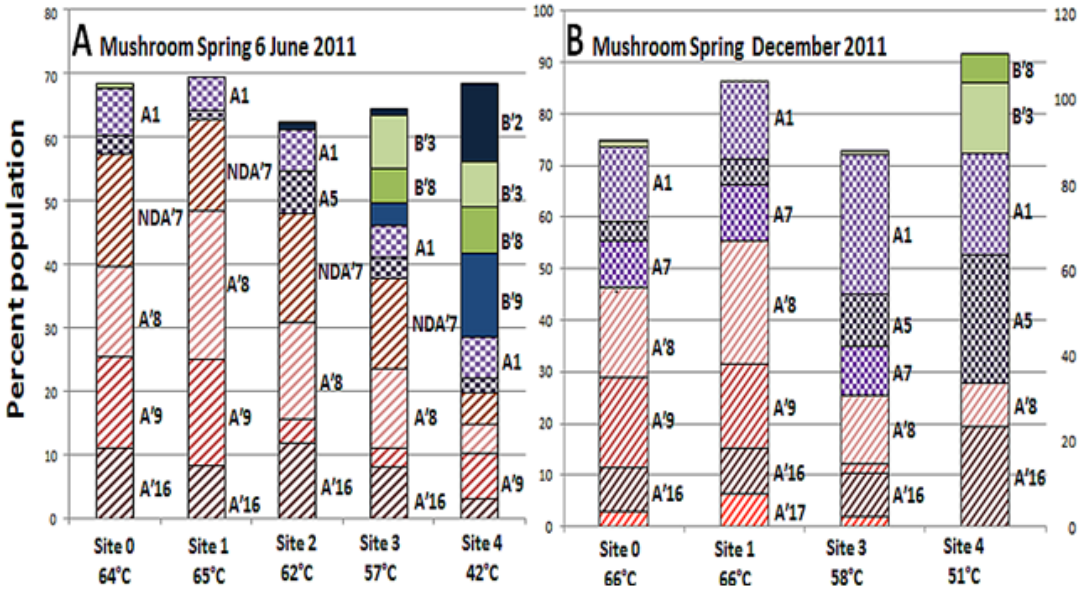


Figure 5.1. Percent population of relatively abundant putative ecotypes (PEs) in the water along the effluent flow path at Mushroom Spring on (A) 6 June and (B) 15 December 2011, based on Ti454-barcode sequencing of *psaA* gene segments. PE designations are labeled to the right of the representative bar. Missing data for site 2 in panel B resulted from failed sequencing reactions. Striped bars represent A′-like, dotted bars A-like, and solid bars B-like PEs (modified from Becraft (2014)).

1997; Ward et al., 2006) suggested that A′ and A-like PEs are better-adapted to higher temperatures ($\geq$ 60°C), and that the B′-like PEs are better-adapted to lower

temperatures ($\leq 60°C$). Yet, with the incubation temperature set to 52°C, which is more representative of the B′-like habitat, the initial cultivation attempt (described above) from mat samples did not yield any isolates representative of B′-like PEs. Since strains representative of PE B′9 were targeted, additional details regarding the distribution of this PE were sought. The vertical distribution data revealed that PE B′9 was most abundant at and near the mat surface (Figure 4.1B), which suggested that PE B′9 may be high-light adapted. Additionally, the results from the seasonal water collections suggested that PE B′9 was a predominant PE in the overflowing water in the summer, at temperatures below 57°C (Figure 5.1A). The inferences suggested by the species distribution data led to the formulation of the following hypothesis. By cultivating strains from mat and water samples collected from lower temperature sites (55 - 60°C) in the summer, and providing a higher light intensity during selection, isolates representative of PE B′9 may be obtained. It should be noted that *Synechococcus* cell numbers in the overflowing water are on the order of 10^3 cells/mL compared to $\sim 10^{10}$ cells/mL in the mat (S. Nowack, unpublished observations).

The seasonal water-distribution data (Figure 5.1) also revealed that A′-like PEs and PE A1 were found in high relative abundance at all temperature sites in both the summer and winter. The winter water samples appeared to be more heavily dominated by A′-like PEs and PE A1 and led to the the following hypothesis: cultivation of strains from high temperature water samples ($\sim 65°C$) that are collected

in the winter may lead to obtaining additional representatives of A′-like PEs and PE A1. Obtaining representatives of PE A1 were prioritized over obtaining representatives of A′-like PEs for two reasons. First, recent results from a multi-locus study suggested that molecular resolution at the *psaA* locus may not be sufficient to discern ecological species within PE A1 (Becraft, 2014). It was shown that the PE A1 strain obtained by Allewalt et al. (2006) and a PE A1 strain that was recently cultivated from Mushroom Spring (65AY6Li), were representative of different multi-locus PEs. Additionally, Becraft et al. (submitted) showed that PE Al exhibits a bimodal vertical distribution in the 63°C Mushroom Spring mat, further suggesting that more than a single ecological species may exist within *psaA* PE A1. Hence, if additional isolates representative of this PE were to be obtained, the phenotypic and genetic diversity within PE A1 could be further investigated. Second, the focus of the cultivation efforts was to obtain representatives of PEs that were predominant in the 60°C Mushroom Spring microbial mat, and A′-like PEs have been shown to have a higher temperature distribution (Ferris and Ward, 1997; Ward et al., 2006). This suggests that higher incubation temperatures, which have not yet been provided, may be required to obtain isolates representative of A′-like PEs.

In the process of testing the above-stated hypotheses, other isolates, representative of PEs that were not targeted, were also obtained. Strains representative of a less abundant B′-like PE, PE B′24, were of particular interest. In Chapter 4, the B′ strain that was isolated by Allewalt et al. (2006) was discussed and a warning

was given about the purity of this strain. One of the three *psaA* genotypes that was detected in the Ti454-barcode analysis of the B′ strain was a representative of PE B′24 (comprising ∼15% of the culture, see Table 4.1). In fact, an axenic strain (CIW-10) that was derived from the Allewalt et al. (2006) B′ strain, and discussed in Kilian et al. (2007), was shown to be dominated by this PE B′24 *psaA* variant(see Table 4.1). In the final section details are provided that describe the sampling and incubation conditions that yielded the strains representative of PE B′24.

Materials and Methods

Sampling. As summarized in Table 5.1 samples were collected from Mushroom Spring, Yellowstone National Park, on five separate occasions. On 7 September 2010 mat samples were collected from temperature sites of 60, 63, and 65°C as described in the methods section of Chapter 4. Following the same protocol, a mat sample was collected at a 55°C site on 2 October 2011, and at a 60°C site on 26 July 2012. Water samples were collected at a 55°C site on 2 October 2011, at a 55 and a 65°C site on 22 January 2013, and at a 60°C site on 19 September 2013. In each case 500 mL of effluent channel water were collected in a thermos, transported back to the lab, and processed for cultivation within three hours of collection.

Cultivation Methods. The Gelrite protocol described in Chapter 4 was applied to the samples collected on 7 September 2010. The mat samples collected on 2

Table 5.1. Summary of the isolates discussed in the cultivation section of this chapter and the sampling and incubation conditions that yielded them. In all cultivation attempts medium DHAY was used as the culture medium and 52°C was the incubation temperature.

Strain name	PE	Date of collection	Sample type / temperature	∼Plated cells / colonies[a]	Plate type	Light intensity of incubator[b]
65AY6Li	A1	09/07/10	mat / 65°C	100/50	Gelrite	50
65AY4H20-AG	A1	01/22/13	water / 65°C	200/20	filter	50
65AY4H20-D	A1	01/22/13	water / 65°C	200/20	filter	50
65AY4H20-AAE	A1	01/22/13	water / 65°C	200/20	filter	50
65AY4H20-J	A1	01/22/13	water / 65°C	200/20	filter	50
65AY3H20-A	A1	01/22/13	water / 65°C	2000/100	filter	50
65AY6A5	A4	09/07/10	mat / 65°C	100/50	Gelrite	50
65AY629	A4	09/07/10	mat / 65°C	100/50	Gelrite	50
65AY6-552-A4	A4	09/07/10	mat / 65°C	100/50	Gelrite	50
63AY4M2	A6	09/07/10	mat / 63°C	10^4/500	Gelrite	50
6-M65AY6	A6	09/07/10	mat / 65°C	100/50	Gelrite	50
65AY6-552-A6	A6	09/07/10	mat / 65°C	100/50	Gelrite	50
60AY4M2	A14	09/07/10	mat / 60°C	10^4/100	Gelrite	50
63AY4M1	A14	09/07/10	mat / 63°C	10^4/500	Gelrite	50
65AY640	A14	09/07/10	mat / 65°C	100/50	Gelrite	50
65AY6A′	A′8	09/07/10	mat / 65°C	100/50	Gelrite	50
55AY5-B	B′5	10/02/11	mat / 55°C	1000/100	filter	50
55AYH20	B′9	10/02/11	water / 55°C	2000/100	filter	50
60H20-5F	B′9	09/19/13	water / 60°C	2000/100	filter	160
68DH1S1[c]	B′12	09/10/08	mat / 68°C	10^8/500	filter	50
YM60AY1HL-3C	B′24	07/26/12	mat / 60°C	10^8/100	filter	450
YM60AY1HL-3D	B′24	07/26/12	mat / 60°C	10^8/100	filter	450
YM60AY1HL-4E	B′24	07/26/12	mat / 60°C	10^8/100	filter	450

[a] Number of cells plated from a diluted environmental sample compared to the number of colony forming units that formed on the plate from which the isolate originated.
[b] Units are scalar μmol photons/m^2/sec.

October 2011 and 26 July 2012 were processed using the dilution and filtration protocol described by Allewalt et al. (2006). The water samples were processed by filtering effluent-channel water containing 200, 20, 2, and 0.2 mL through 0.2 μm pore size polycarbonate filters (Millipore, Billerica, MA). Before filtering, the 2 and 0.2 mL water samples were suspended in 10 mL total volume of autoclaved

Mushroom Spring water that was collected from the source pool. Any well-isolated colonies that grew on plates were picked with a sterile toothpick and were suspended in 2 mL of liquid medium DHAY.

Selection Conditions. The incubation temperature was set to 52°C for all cultivation attempts. Three different light conditions were used to account for the vertical distribution patterns of the targeted PEs (Figure 4.1B). On 7 September 2010 the Gelrite plates that had been inoculated with the serially diluted mat samples were incubated under 50 μmol photons/m^2/sec of white fluorescent light. Two parallel dilution series were prepared and plated on filters for both the mat and the water samples collected on 2 October 2011. One was incubated at 50 μmol photons/m^2/sec and the the other was incubated at 450 μmol photons/m^2/sec. The mat samples collected on 26 July 2012 were diluted to extinction, plated on filters, and incubated at 450 μmol photons/m^2/sec. The filters inoculated with the diluted water samples collected on 22 January 2013 and 19 September 2013 were incubated at 50 μmol photons/m^2/sec and 160 μmol photons/m^2/sec, respectively. Sampling, cultivation, and selection conditions are summarized in Table 5.1.

Molecular Analyses. After colonies were picked and suspended in liquid medium DHAY, the methods described in Chapter 4 were followed to determine the *psaA* PE of the liquid culture and to assess culture purity. These methods are summarized in steps 1-7 of the flow chart shown in Appendix C, Figure C.1.

Results and Discussion

The details regarding the sampling methods, incubation conditions, and the dilutions that yielded each of isolates discussed in the following subsections are summarized in Table 5.1. Since the water samples had approximately seven orders of magnitude fewer cells/mL than the mat samples, for comparison purposes, the numbers of cells plated and the number of colonies forming on the plate are reported to describe the dilution. Figure 5.2 indicates the current status of each isolate in the cultivation, phenotyping and genomic sequencing pipeline. Table 5.2 reports the Ti454-barcode data that corresponds to each strain that has either had or is in the process of having its genome sequenced.

Isolates Obtained from the Initial Cultivation Attempt. In addition to the four isolates discussed in Chapter 4 (PE A1 (65AY6Li), PE A4 (65AY6A5), PE A6 (63AY4M2), and PE A14 (60AY4M2)), several other strains were obtained from the mat samples collected on 7 September 2010. Thus far, the genome sequences of these four strains and two additional strains representative of PE A14 (63AY4M1, 65AY640) have been determined and annotated (Figure 5.2, green). Two additional strains representative of PE A4 (65AY629, 65AY6-552-A4), two strains representative of PE A6 (65AY6-552-A6, 6-M65AY6) and one strain representative of PE A′8 (65AY6A′) were also cultivated from the mat samples collected on 7 September

Table 5.2. Summary of *psaA* Ti454-barcode sequencing analyses of *Synechococcus* cultures that either have had, or are in the process of having their genomes sequenced. Colors correspond to where in the genome sequencing pipeline the strains currently reside (see Figure 5.2). Sequences with systematic errors were excluded.

			dominant variant sequences[a]		other sequences[b]	
Strain name	total sequences	PE	DVs	%	total	%
JA-3-3Ab[c]	1116	all	931	83.4		
		A1	930	83.3	185	16.6
		A14	1	0.1		
65AY6Li	980	**A1**	780	79.6	200	20.4
65AY6A5	1325	all	1106	83.5		
		A4	1105	83.4	219	16.5
		A1	1	0.1		
65AY629	2264	all	1844	81.5		
		A4	1843	81.4	420	18.6
		B′-like	1	0.04		
63AY4M2[d]	2027	**A6**	1756	86.6	271	13.4
6-M65AY6	6327	all	5602	88.54		
		A6	5596	88.45	725	11.5
		B′-like	6	0.09		
65AY6-552-A6	1213	**A6**	1035	85.3	178	14.7
60AY4M2	920	**A14**	756	82.2	164	17.8
63AY4M1[e]	443	**A14**	374	84.4	69	15.6
65AY640	1007	**A14**	816	81	191	19
65AY6A′	532	**A′8**				
55AY5-B	1004	all	752	75		
		B′5	751	74.8	253	25.2
		A1	1	0.1		
68DH1S1	843	**B′12**	703	83.4	140	16.6
JA-2-3B′a	2054	all	1690	82.3		
(2-13)[c]		**B′19**	1385	67.4	364	17.8
		B′24	300	14.6		
		B′11	5	0.2		
CIW-10[f]	1528	all	1077	70.5		
		B′24	438	28.7	413	27
		†	639	41.8	38	2.5

Putative ecotypes (PEs) in bold are the dominant PEs in the culture.

[a] Number of sequences that are identical to the dominant variant (DV; identical sequence that made up the plurality of a putative ecotype (PE) clade) of the PE listed in the same row.

[b] Other sequences that are closely related to the DV sequence.

[c] Isolate obtained by Allewalt et al. (2006).

[d] The strain representative of PE A6 in which the *psaA* genotype switched under experimental conditions.

[e] The strain representative of PE A14 that has several heterozygous single nucleotide polymorphisms.

[f] Isolate obtained by Kilian et al. (2007) that was a direct descendant of the isolate obtained by Allewalt et al. (2006).

†A dominant sequence variant in the culture that is distinct from PE B′24 by 17 single-nucleotide polymorphisms.

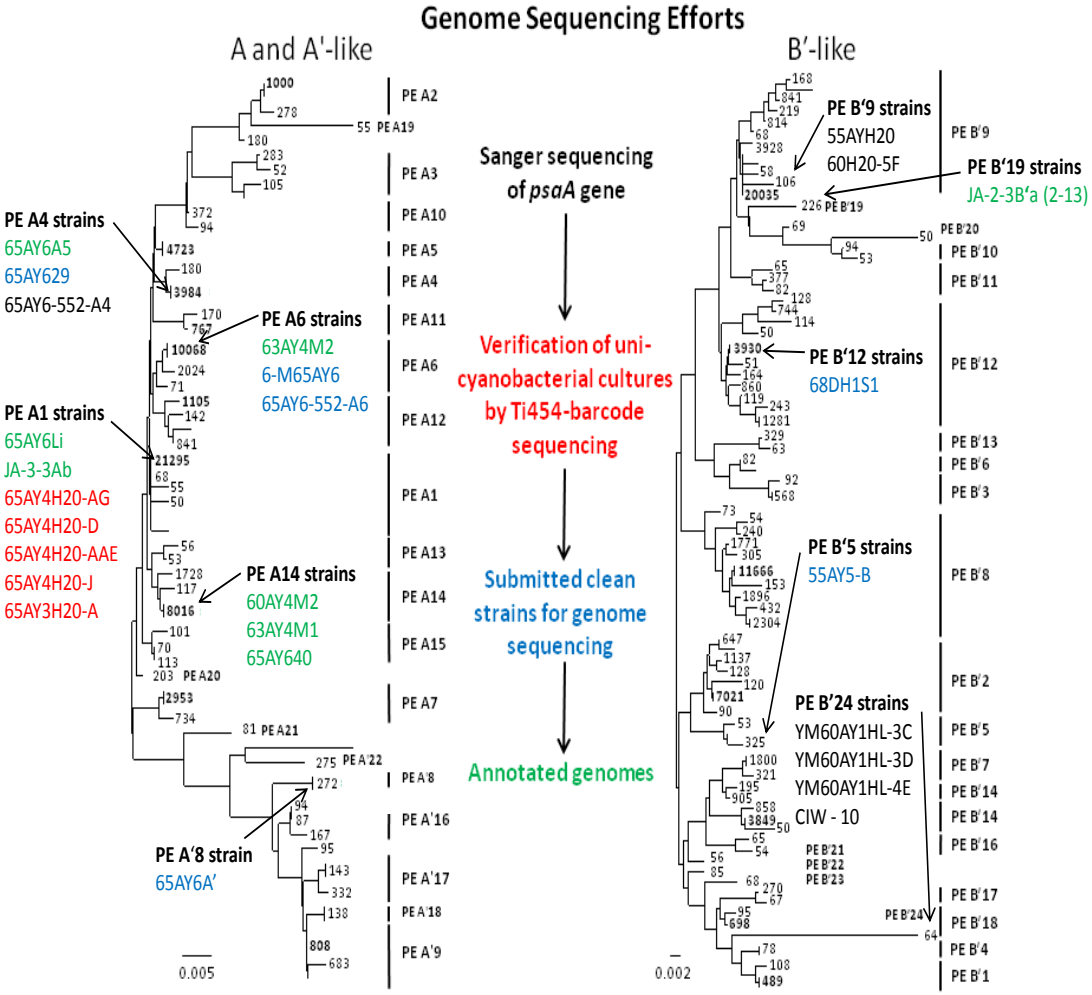


Figure 5.2. Phylogenetic affiliation of *Synechococcus* strains relative to putative ecotypes (PEs) predicted by Ecotype Simulation from high frequency (≥ 50 occurrences across 96 samples) *psaA* barcode sequences. Separate neighbor-joining phylogenies of A-like and A′-like (left), and B′-like (right) are separated by a scheme showing the analysis pipeline (center). PEs are demarcated with vertical bars. Color highlighting of strains indicates where in the pipeline the culture currently resides (modified from Becraft, 2014). With the exception of the strains representative of PE B′9 all isolates are 100% identical to the HFS that is most abundant within the PE to which it is classified.

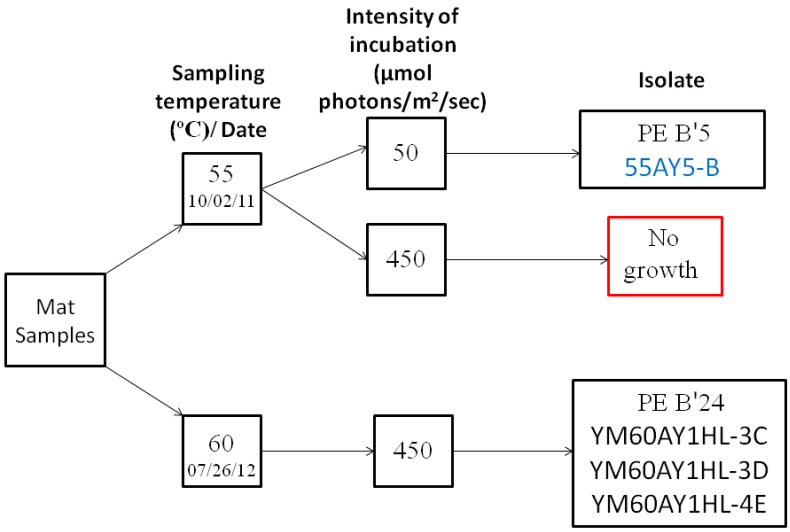
2010, but Ti454-barcode sequencing revealed that these strains were not predominated by a single cyanobacterium. These five strains were purified by the dilution and filtration protocol described in Allewalt et al. (2006), and re-sequenced using Ti454-barcode technology. The barcode data revealed that four of these strains, both of the strains representative of PE A6, one of the strains representative of PE A4 (65AY629), and the strain representative of PE A′8 had been purified (based on the sequence-based criterion established in Chapter 4, see Table 5.2). These four strains have been sent for genome sequencing but have not yet been annotated (Figure 5.2, blue). The data from this second round of barcode sequencing revealed that strain 65AY6-552-A4 (Figure 5.2, black) was still not pure, therefore the purification process was applied a second time to this strain, for which the Ti454-barcode results are still pending.

Collectively, the first cultivation attempt yielded three strains representative of each of three PEs (A4, A6, and A14), one strain representative of PE A1, one strain representative of PE A′8, and no strains representative of PE B′9. Each of these strains is 100% identical at the *psaA* locus to the most abundant high frequency sequence within the PE clade ($\geq$ 50 occurrences across 96 mat samples). See Table 5.1, Figure 5.2, and Table 5.2 for a summary of the details regarding these strains.

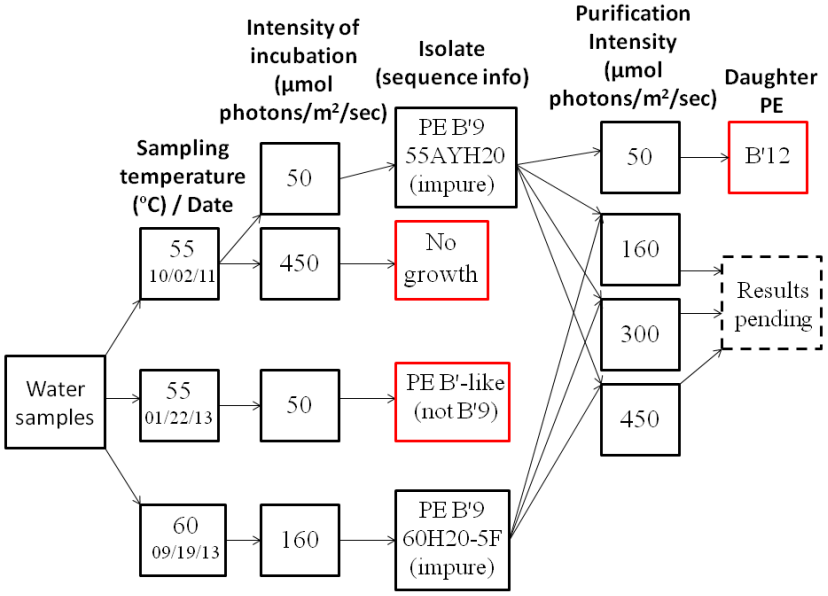
After the initial success with Gelrite-solidified media, I was unable to obtain

colony growth at any dilution using this cultivation protocol. In subsequent cultivation attempts the floating filter technique employed by Allewalt et al. (2006) was used, except that medium DHAY was substituted for medium DH.

Obtaining Isolates Representative of PE B′9. In an attempt to obtain strains representative of PE B′9, lower-temperature mat and water samples were collected for cultivation (≤ 60°C). On 2 October 2011 a 55°C mat sample was processed using the dilution and filtration protocol described in Allewalt et al. (2006). The details that follow are summarized in Figure 5.3A. The filters were incubated both at a low-light condition (50 μmol photons/m^2/sec) and a high-light condition (450 μmol photons/m^2/sec) in order to compare the PEs of strains that were cultivated under different incubation conditions. A strain representative of a less abundant B′-like population, PE B′5 (55AY5-B; Figure 5.2 and Table 5.2, blue), originated from a colony picked from a filter that was incubated at 50 μmol photons/m^2/sec (Figure 5.3A). The purity of this strain was confirmed via the Ti454-barcode criterion, and its genome is currently being sequenced. Colony growth was not observed on any of the filters that were incubated at 450 μmol photons/m^2/sec. On 26 July 2012 a second attempt to obtain isolates representative of PE B′9 from a diluted mat sample that was incubated at 450 μmol photons/m^2/sec was made. In this case, the mat sample was collected from a 60°C site. The rationale for repeating the incubation conditions was that this particular sample was collected closer to the



(A)



(B)

Figure 5.3. Flow chart depicting the isolation process for cultivation efforts targeting representatives of PE B′9. (A) Isolates originating from mat samples and (B) isolates originating from water samples. The red boxes indicate that isolation efforts along that path have been discontinued.

summer solstice, and the seasonal distributions had suggested that PE B′9 may be more abundant in the summer (Figure 5.1). Once again, isolates representative of PE B′9 were not obtained. In fact, isolates representative of a less-abundant B′-like PE, PE B′24, were the only isolates obtained from this collection (Figure 5.3A). These isolates are discussed in a later subsection.

As described in the introduction, PE B′9 was also found in high relative abundance in the water overflowing the mat at lower temperatures in the summer (<57°C; Figure 5.1A). Therefore, on the 2 October 2011 sampling trip, two parallel dilution series were also prepared from a 55°C water sample, and plated on filters. The isolation process that is described next is depicted in Figure 5.3B. One set of filters was incubated at 50 μmol photons/m^2/sec and the other set was incubated at 450 μmol photons/m^2/sec. While growth did not occur on any of the filters incubated under the high-light condition, several isolated colonies were picked from filters incubated under the low-light condition. One of the strains (55AYH20; Figure 5.2, black) had a Sanger sequence representative of PE B′9, but Ti454-barcode sequence data revealed that this strain was not predominated by a single cyanobacterium. Thus, this culture was re-diluted to extinction and the dilution series was plated on filters. Several colonies formed on high-dilution plates and the daughter colonies were suspended in liquid medium DHAY. After extracting DNA from the scaled-up volumes, Sanger sequencing revealed that all of the subcultures derived from the daughter colonies were representative of PE B′12, which confirmed that the original

culture was indeed not unicyanobacterial. All of the subcultures were incubated at 50 μmol photons/m^2/sec throughout this process, which is summarized in Figure 5.3B. One hypothesis for why the purification attempts resulted only in PE B′12-like subcultures is that the light intensity may have been too low to select for B′9-like organisms, possibly due to representatives of this PE having a higher low-light threshold (see the discussion in Chapter 4). To test this hypothesis, the original parent culture (55AYH20) was diluted to extinction and plated on filters a second time, and then incubated under higher light conditions (160, 300, and 450 μmol photons/m^2/sec, see Figure 5.3B) (these dilutions have been done in collaboration with Katrina Jackson, a Ward Lab undergraduate student). Colonies grew on filters incubated at all light intensities, and well-isolated colonies have been picked and scaled-up to higher volumes in liquid medium DHAY. Sanger sequencing is in progress. It should be noted that from a previous cultivation effort (not described here) the Ward lab had obtained an isolate representative of PE B′12 from a low-dilution mat sample and its genome is currently being sequenced (68DH1S1, Figure 5.2 and Table 5.2, blue). The species demarcation method, at the time this strain was submitted for genome sequencing, grouped PE B′12 and PE B′9 into a single PE. Later analyses suggested that PE B′12 is distinct from PE B′9 and not as predominant *in situ*, thus, additional isolates representative of PE B′12 have not been targeted.

The ability to cultivate a strain representative of PE B′9 from a water sample

suggested that additional cultivation efforts from lower-temperature water samples should be made. Therefore, on 22 January 2013 a water sample was collected at a 55°C site. Since the filters from the 2 October 2011 collection that were incubated under high-light conditions (450 μmol photons/m^2/sec) did not yield any colonies, the incubation condition was set to 50 μmol photons/m^2/sec. This low-light incubation condition did not result in obtaining any strains representative of PE B′9, and only B′-like isolates representative of less-abundant PEs were cultivated (see Figure 5.3B). Closer examination of the species distributions from the overflowing water (Figure 5.1) suggested that the PE B′9 population was not as abundant in December as it had been in June. Since this cultivation effort was made in January, as opposed to the previous cultivation attempt that had been made in October, it was hypothesized that the proposed high-light adapted PE B′9 population may be more abundant when more light is available. Therefore, on 19 September 2013 water was collected from a 60°C site, and an intermediate light intensity was provided as an incubation condition (160 μmol photons/m^2/sec) (see Figure 5.3B). Katrina Jackson has been responsible for all cultivation work related to these particular samples as well. Sanger sequence data revealed that all of the strains obtained from this collection were B′-like, and more importantly, that one strain was representative of PE B′9 (60H20-5F; Figure 5.2, black). Since the Sanger sequence for the strain representative of PE B′9 was of poor quality (high background signal) this strain was re-diluted to extinction, plated on filters and incubated at 160, 300, and 450

μmol photons/m^2/sec (see Figure 5.3B). Several colonies from high-dilution filters have been picked, suspended in liquid medium DHAY, and scaled up to higher volumes. Efforts to extract DNA for Sanger sequencing are currently in progress, and if the subcultures that were derived from daughter colonies have the same PE B′9 genotype as the parent culture, this will provide enough evidence to have several of the subcultures prepared for Ti454-barcode sequencing.

Obtaining Additional Isolates Representative of PE A1. The data shown in Figure 5.1 revealed that PE A1 was one of the most abundant A-like PEs in the water at temperature sites of 57°C and hotter, and that PE A1 was especially abundant in water during the winter. Therefore, a water sample was collected from a 65°C site on 22 January 2013 and processed for cultivation. The light intensity of incubation was set to 50 μmol photons/m^2/sec. Five strains representative of PE A1 (65AY4H20-AG, 65AY4H20-D, 65AY4H20-AAE, 65AY4H20-J, and 65AY3H20-A) (Figure 5.2, red) were cultivated from this water sample and Ti454-barcode data suggested that these strains were predominated by a single cyanobacterial strain. The hypothesized inability of variation in the *psaA* gene to discern ecologically distinct species within PE A1, that was mentioned in the introduction, resulted in a modification in our approach. Before obtaining the genome sequences of these five additional strains representative of PE A1, a PCR-based multi-locus sequence analysis will be conducted. This preliminary step will ensure that the strains representative of PE

A1 are properly classified before they are treated as representatives of strains within a single ecological species population.

Isolates Representative of PE B′24. As stated above, the mat sample collected on 26 July 2012 that was diluted to extinction and incubated under a high-light condition (∼450 μmol photons/m^2/sec), did not yield any isolates representative of the targeted PE B′9. Colonies only grew on low-dilution plates (diluted one order of magnitude from the original mat sample), and all of the strains derived from the colonies (YM60AY1HL-3C, YM60AY1HL-3D, YM60AY1HL-4E, Figure 5.2, black) were representative of a less abundant B′-like population, PE B′24. Although strains representative of this PE were not targeted, they were interesting for the following reason. In Table 5.2 it was shown that strain JA-2-3B′a (2-13) was not predominated by a single cyanobacterium, and PE B′24 was shown to comprise ∼15% of the batch culture. Moreover, an axenic strain that was derived from JA-2-3B′a (2-13) (Kilian et al., 2007), strain CIW-10, and obtained by the Ward lab via a colleague, was also shown to not be predominated by a single cyanobacterium. The most abundant *psaA* genotype in this culture was representative of PE B′24 (see Table 5.2). Strain CIW-10 was obtained by providing directional light as a selecting force to cause *Synechococcus* cells to glide away from the heterotrophic contaminants utilizing phototaxis. Surprisingly, Kilian et al. (2007) reported that strain CIW-10 was low-light adapted (it would grow at 200 but not at 400 μmol photons/m^2/sec), which

contradicts what the selection conditions infer about the isolates obtained here. Future studies to investigate the light adaptations of additional isolates representative of PE B′24 have been proposed.

Concluding Remarks. Collectively, the results from these cultivation attempts suggest that 52°C is an incubation temperature that is capable of selecting for a diverse array of *Synechococcus* populations, and that light intensity is an important selection criterion for obtaining isolates that are representative of the predominant PEs in the 60°C Mushroom Spring microbial mat. The cultivation of strains from high-temperature water samples (∼65°C) and incubation under low light (50 μmol photons/m^2/sec) appears to be an effective way to obtain representative isolates of PE A1. Additional isolates representative of A′-like PEs, which were also found in relatively high abundance in the high temperature water samples, were not obtained. This could be due to the fact that the A′-like PEs are hypothesized to be higher-temperature adapted (Ferris and Ward, 1997; Ward et al., 2006; Miller and Castenholz, 2000) and that 52°C was used as the incubation temperature. In fact, *Synechococcus* strains isolated from Hunter's Hot Springs, Oregon (Miller and Castenholz, 2000), that are closely related to the predominant 16S rRNA A′-like population found in Octopus Spring, were reported to be able to grow at 55°C but not at 50°C. It was hypothesized that this behavior was due to a thermotolerance

trade-off, i.e., in order to grow at higher temperatures, a possible evolutionary cost is the loss of the ability to grow at lower temperatures (Miller and Castenholz, 2000).

The results also suggest that, for the targeted isolates, a light intensity of 450 μmol photons/m^2/sec may be too high when supplied continuously during incubation. However, light provided in the 160-200 μmol photons/m^2/sec scalar intensity range, and sampling from 55-60°C water, may be ideal for selecting isolates representative of the PE B′9 population.

Finally, the success achieved with the Gelrite protocol must be reiterated. Eleven of the isolates discussed here were obtained by following the Gelrite protocol, including all of the strains representative of the hypothesized low-light adapted PEs (A4, A6, and A14). I hypothesize that either (i) the solid Gelrite medium possesses mat-like qualities that the populations inhabiting deeper mat layers may prefer or (ii) that the filtering process excludes the necessary nutrients, or co-inhabitants, that the deeper populations may require in order to grow on filters. Therefore, if additional isolates representative of PEs A4, A6, and A14 are to be obtained, the problems with the Gelrite protocol should be resolved.

Testing the Ecological Interchangeability of Strains within *psaA* PEs

Introduction

As discussed in the preamble of the chapter, PEs representative of ecologically distinct *Synechococcus* species populations were predicted by Ecotype Simulation

(Koeppel et al., 2008) from sequence variation at the *psaA* locus found in mat samples collected from Mushroom Spring. The underlying bacterial speciation theory that was simulated by the Ecotype Simulation algorithm (for the PEs predicted here) was the Stable Ecotype Model of species and speciation (Figure 5.4) (Cohan, 2002; Cohan, 2006; Cohan and Perry, 2007). This model is one of many models of

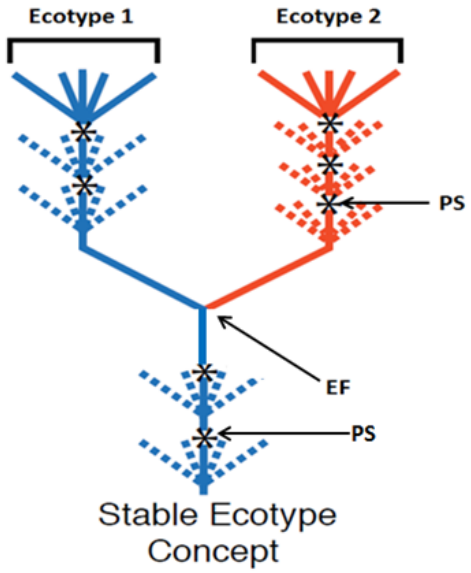


Figure 5.4. The Stable Ecotype Model of microbial speciation depicting the relationship between ecologically distinct populations and DNA sequence clusters. Different ecotypes (equivalent to ecologically distinct species) are represented by different colors. EF shows an ecotype formation event. PS (asterisks) shows examples of periodic selection events that purge population diversity within an ecotype except for one most-fit individual. Dashed lines represent extinct lineages (modified from Cohan and Perry, 2007).

bacterial evolution that have been proposed by Fred Cohan at Wesleyan University,

and has been shown to best match the observed distribution patterns of genetic diversity of *Synechococcus* populations inhabiting Mushroom Spring and Octopus Spring (Ward and Cohan, 2005). The basis of the Stable Ecotype Model is that mutation events create diversity within a given population. If a certain genetic variant acquires a beneficial trait, then that variant either (a) purges the diversity because the acquired trait made it more fit than other variants in the population (periodic selection) or (b) utilizes its newly acquired trait to inhabit a new environment (ecotype formation event), and in this case, future periodic selection events that affect the new population will not affect the parent population and vice versa (Figure 5.4). The members of populations whose periodic selection events affect one another are considered to be ecologically interchangeable, and the Stable Ecotype Model demarcates all such variants to be within a single PE. Note that this does not require individual members within a PE to be clonal. If the periodic selection events of one population do not affect another population then the two populations are considered to be two distinct PEs.

In this section the light adaptations of multiple strains within each of three *psaA* PEs are compared. The rationale for this analysis is to test the hypothesis that the *psaA* PEs predicted by Ecotype Simulation (based on the Stable Ecotype Model) are species populations whose members are ecologically interchangeable. If strains within a PE exhibit similar adaptations to light, this will provide support for the

above-stated hypothesis. If the strains within PEs exhibit different adaptations to light, alternative hypotheses must be developed.

Materials and Methods

Sample Collection, Cultivation, and Molecular Methods. See Chapter 4 methods and the preceding subsection of this chapter.

Growth Experiments. Growth experiments on cultivated isolates were conducted as described in the growth experiment protocol in Chapter 4. Two strains representative of PE A1, three strains representative of PE A14, and two strains representative of PE A6 were grown at different scalar light intensities (covering the range of intensities observed in nature), and bubbled with 6% CO_2 in air. The strains representative of PE A1 were incubated at 52°C, whereas the strains representative of PEs A14 and A6 were incubated at 60°C.

Results and Discussion

Cultivation and Sequencing. The strains that are discussed in this section are listed in Table 5.3. All of the strains were determined to be dominated by a single cyanobacterium based on the Ti454-barcode criterion established in Chapter 4 (see Table 5.2), and their genome sequences have either been or are in the process of being obtained. Included are two strains representative of PE A1 (65AY6Li

and JA-3-3Ab, the latter mentioned strain was obtained by Allewalt et al. (2006) from Octopus Spring, YNP), three strains representative of PE A14 (60AY4M2, 63AY4M1, 65AY640), and two strains representative of PE A6 (6-M65AY6, 65AY6-552-A6).

Table 5.3. Summary of light responses at 60°C and bubbled with 6% CO_2 in air during the experiment.

Strain name	PE	Sampling temperature	Dilution[a]	Incubation temperature	Upper-light limit	Optimal intensity[b]
65AY6Li	A1	65°C	10^{-7}	52°C	<2300	125-1600
JA-3-3Ab	A1	†	†	52°C	>1600	600
60AY4M2	A14	60°C	10^{-5}	60°C	1050	600-850
63AY4M1	A14	63°C	10^{-5}	60°C	>2500	125-600,2100
65AY640	A14	65°C	10^{-7}	60°C	2300	250
6-M65AY6	A6	65°C	10^{-7}	60°C	2400	1050
65AY6-552-A6	A6	65°C	10^{-7}	60°C	600	125

[a] Dilution from which isolate originated.
[b] Units are scalar μmol photons/m^2/sec.
†Cultivated by Allewalt et al. (2006).

Light Adaptations of Strains within and among PEs. At four of the six light conditions that were tested (25, 250, 1100 and 1600 μmol photons/m^2/sec), the two PE A1 strains exhibited similar light responses (at 52°C), which can be observed by the overlapping range bars shown in Figure 5.5A. At 125 μmol photons/m^2/sec the 65AY6Li strain (Figure 5.5A, solid line) grew noticeably faster than the JA-3-3Ab strain (Figure 5.5A, dashed line) (1.26 versus 0.88 doublings/day), whereas at 600 μmol photons/m^2/sec, strain JA-3-3Ab was the faster-growing strain (1.85 versus

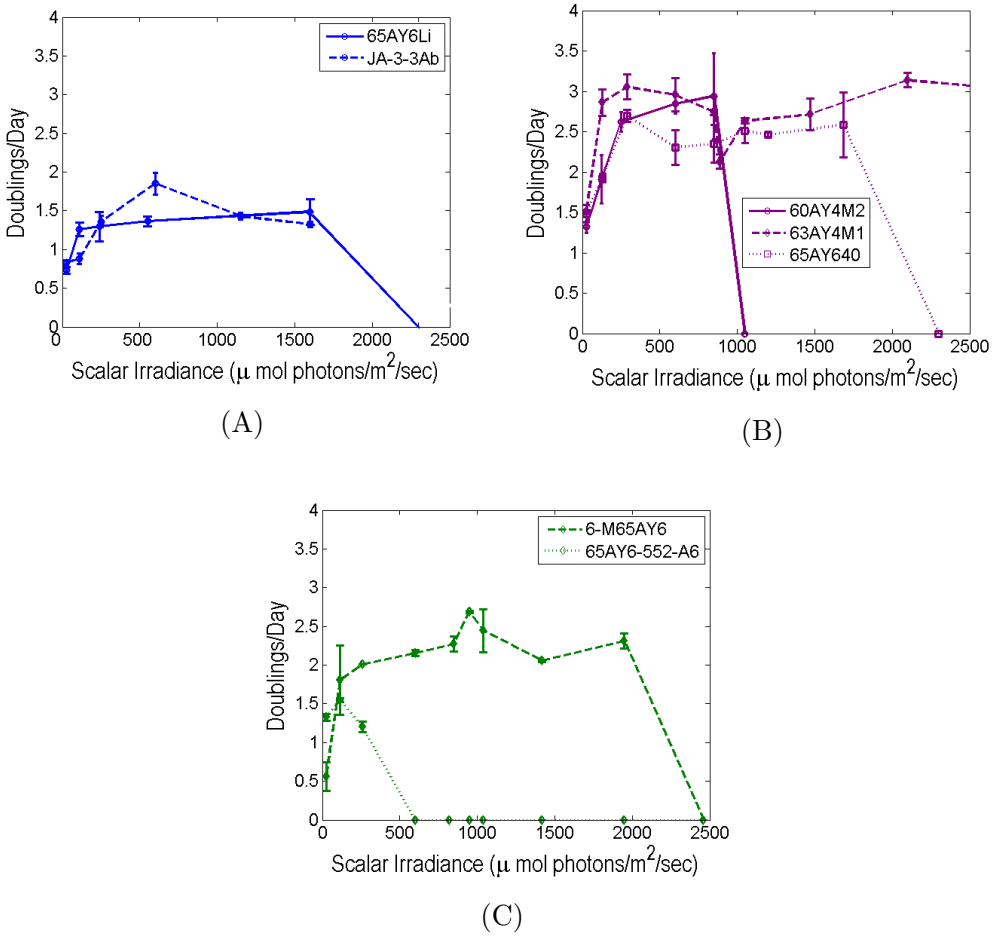


Figure 5.5. Light adaptations of multiple *Synechococcus* strains representative of three different *psaA*-defined PEs. (A) Two PE A1 strains grown at 52°C, (B) three PE A14 strains grown at 60°C, and (C) two PE A6 strains grown at 60°C. 6% CO_2 in air was bubbled in each experiment. The mean of the two replicates of each condition is denoted by the data point, and the range bars represent the growth rate of each replicate

1.37 doublings/day). The upper-light limit of the JA-3-3Ab strain has not yet been determined.

Light adaptation comparisons were also made for three PE A14 strains, but at 60°C (Figure 5.5B). At the lowest scalar light intensity tested (25 μmol photons/m^2/sec)

the three strains had very similar growth rates. At 125 and 250 μmol photons/m^2/sec the 60AY4M2 (Figure 5.5B, solid) and 65AY640 (Figure 5.5B, dotted) strains grew at a nearly identical rate, while the 63AY4M1 strain (Figure 5.5B, dashed) grew at a noticeably higher rate than the other two PE A14 strains at 125 μmol photons/m^2/sec (doubling 2.9 times per/day compared to 1.9 doublings/day), but only marginally faster at 250 μmol photons/m^2/sec (3.05 versus 2.62 and 2.69 doublings/day). In the 600-850 μmol photons/m^2/sec range all three PE A14 strains doubled between two and three times per day; the 60AY4M2 and 63AY4M1 strains grew at a comparable rate to one another and slightly faster than the 65AY640 strain. At light intensities between 850 and 1700 μmol photons/m^2/sec the 63AY4M1 and 65AY640 strains exhibited comparable growth rates. In contrast, the 60AY4M2 strain was unable to tolerate light intensities of 1400 μmol photons/m^2/sec. The 65AY640 strain could tolerate 1700 μmol photons/m^2/sec but not 2300 μmol photons/m^2/sec, and the upper-light limit of the 63AY4M1 strain was at least 2500 μmol photons/m^2/sec. It should be noted that collaborators from the Joint Genome Institute have informed us that the 63AY4M1 strain had several heterozygous single-nucleotide polymorphisms, suggesting that this strain may not be unicyanobacterial.

The light adaptations at 60°C of two strains representative of PE A6 (6-M65AY6 (Figure 5.5C, dashed) and 65AY6-552-A6 (Figure 5.5C, dotted)) were also determined. A third PE A6 strain (63AY4M2) was initially included in this analysis but the sequences of post-experiment samples revealed that its genotype had switched

under the experimental conditions (see discussion in Chapter 4 and Figure C.3), thus this strain has been excluded from the comparisons. Post-experiment genotype validation efforts are currently in progress for all of the other strains discussed in this section (see section on post-experiment genotype validation in Chapter 4 for methods). The two PE A6 strains exhibited very different responses to light (Figure 5.5C). The 6-M65AY6 strain (Figure 5.5C, dashed line) was able to tolerate light intensities up to 2000 but not 2500 μmol photons/m^2/sec, whereas strain 65AY6-552-A6 (Figure 5.5C, dotted line) was unable to sustain growth at or above 600 μmol photons/m^2/sec. The 65AY6-552-A6 strain had a doubling rate that was approximately twice that of the 6-M65AY6 strain at the lowest light intensity investigated (25 μmol photons/m^2/sec).

The results suggest that PE A1 fits the hypothesis that strains within predicted *psaA* PEs are ecologically interchangeable. In contrast, the strains representative of PE A14 and PE A6 do not appear to fit this hypothesis. The most notable difference in light adaptations of strains within PE A14 was the variability in growth rate behavior at high-light intensities. An obvious explanation for the lack of observed ecological interchangeability is that other, rare *Synechococcus* populations, possibly with higher-light adaptations, were undetected in the Ti454-barcode analyses that were used to test culture purity (Table 5.2). Several post-experiment samples from high-light conditions are currently being sequenced to determine if this may be the case.

If the post-experiment sequencing efforts reveal that the genotypes did not shift under the imposed light conditions, alternative hypotheses must be considered. One such hypothesis, regarding strains representative of PE A14, has already been developed. Observe that PE A14 has a deep distribution in the mat at 60°C (see Figure 4.1B), and Figure 4.1C suggests that this PE may be exposed to less than 10% of the incident irradiance in nature, which implies that strains representative of this PE may be low-light adapted. Next, observe that for lower-light conditions, between 0-850 μmol photons/m^2/sec, i.e., within the range of light intensities that PE A14 is hypothesized to experience in nature, the three PE A14 strains exhibited similar light responses (see Figure 5.6). Moreover, the light responses of the three

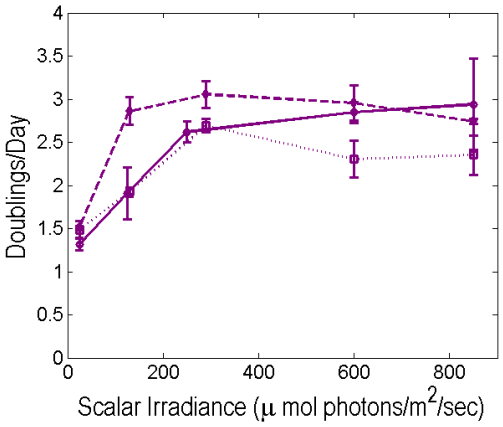


Figure 5.6. Lower-light adaptations of multiple strains within PE A14 grown at 60°C and bubbled with 6% CO_2 in air.

.

PE A14 strains were nearly identical at the lowest light intensity tested (25 μmol photons/m^2/sec). In contrast, when the strains representative of PE A14 were exposed to light conditions that they do not experience within their niche, extreme variations in growth rate and light tolerance were observed. Therefore, an alternative to the hypothesis that strains within *psaA*-defined PEs have similar growth responses with respect to light intensity, is that strains within PE A14 have similar growth responses to light over the range of intensities to which they are exposed to in nature. The growth rate behavior at high light may be inconsequential with respect to selection. Note that this alternative hypothesis would also fit the growth rate behavior of the PE A1 strains. The distribution data from Figure 4.1B suggest that the strains representative of PE A1 may be able to tolerate higher light because PE A1 is found in higher relative abundance near the mat surface than PEs A6 and A14, which could possibly explain why the PE A1 strains exhibited similar light responses at higher light intensities (1600 μmol photons/m^2/sec). If this alternative hypothesis is indeed true, this would suggest that strains within species may be able to maintain variation with respect to upper-light limit. Another way to interpret differences in upper light limit is that replicate strains belong to different temperature-adapted PEs that are embedded within the same *psaA* PE. The results presented in Figure 4.6 support this hypothesis, as it was shown that temperature had a considerable effect on the observed light responses.

Concluding Remarks

In an attempt to explain the results presented above, several alternative hypotheses have been developed. The first hypothesis is that some of the cultures are contaminated. DNA extracted from samples that were collected at the end of the growth experiments is currently being sequenced to determine if the *psaA* genotypes were the same at the end of the experiment as they had been at the beginning of the experiment. Post-experiment sequence data have already been analyzed for one of the strains discussed here (60AY4M2 (PE A14)); see asterisks in Figure 4.5), and have validated that its genotype did not switch under the imposed experimental conditions. But, it should also be noted that one of the PE A6 strains discussed in Chapter 4 did exhibit a switch in its genotype (63AY4M2), and therefore was excluded from those and the above analyses. If it is found that the genotypes of other strains also had undergone a genotype switching under experimental conditions, this would suggest that the initial cultures were not unicyanobacterial, and that the current cultivation protocol and/or purity criterion must be improved. If the post-experiment sequencing data confirm that the genotypes of the strains remained stable over the course of the experiment, then the hypothesis that variation at the *psaA* locus can be used to demarcate ecologically distinct populations should be revisited.

The genome sequences corresponding to the strains discussed here are currently

being investigated by Millie Olsen, a graduate student in the Ward Lab, in an attempt to identify a set of genes that can be used to apply a multi-locus sequence approach to demarcate species. The multi-locus approach will provide greater molecular resolution and buffer against the possibility that single loci that have recombined caused a variant to be classified in the wrong PE. This will explore the hypothesis that strains assigned to *psaA* PEs were misclassified either due to inadequate molecular resolution or to the impact of recombination. If the post-experiment samples do not show evidence of genotype switching, and also if the multi-locus analysis reveals that strains within a *psaA* PE are within the same multi-locus PE, yet another hypothesis is that the *Synechococcus* spp. inhabiting the microbial mats of Mushroom Spring may have the ability to change their phenotypic properties in response to environmental change (phenotypic plasticity), by maintaining ecological diversity (e.g., differences in upper light limits) within the species population.

CHAPTER 6

SUMMARY OF MAJOR RESULTS AND CONCLUSIONS

While some of the results that have been presented here are more mathematically relevant and some are more biologically relevant, the interaction between researchers from the two disciplines has been a driving force for many of the interesting discoveries that have been made. This statement is supported by the conclusions that follow, and summarize what my collaborators and I have learned about niche structure in a temporally varying environment through my thesis project.

A large part of this thesis has been devoted to theoretically determining the optimal behavior of an organism that is inhabiting a temporally varying environment. As was described in the implementation section of the introduction section, the optimization idea itself arose from communication with a microbial speciation expert (Fred Cohan, Wesleyan University), who expressed concern over assuming a Gaussian form to model the niche. This interaction motivated the idea of optimizing the niche function without making any assumptions about a fixed form. Implementing this strategy, first, led to results that agreed with previous studies – that some frequencies of temporal environmental variation may be important in determining niche structure (low and intermediate) and others may not (high) (Hutchinson, 1961; Levins, 1968; Gabriel and Lynch, 1987; Gilchrist, 1995). And second, this strategy

led to the novel discovery that the order in which the environmental conditions occur are important in the intermediate frequency regime, but not in the limiting high and low frequency regimes (which were determined to depend exclusively on the environmental density function (Chapter 2)). David Ward, a microbial ecologist, suggested that this result gave the appearance of a biological clocking mechanism. It should be noted that circadian clock genes are found in the genomes of many cyanobacterial species, including the *Synechococcus* spp. that inhabit the microbial mats found in the effluent channels of Mushroom Spring (Baca et al., 2010; M.Olsen, unpublished). Also note that *Synechococcus* species are oxygenic phototrophs that utilize light to fix CO_2, and that the light cycle is a daily cycle. It has been observed that the predominant *Synechococcus* species from Mushroom Spring have a doubling rate on the order of 1-3 doublings per day in laboratory culture (Kilian et al., 2007, Chapter 4 and Chapter 5). Combining these observations provides support for the following conclusion: biological clocking mechanisms may have evolved from temporal environmental fluctuations that occur with intermediate frequency, with respect to the growth rate of the organism, and that it is these frequencies where interaction between an organism and its temporally varying environment is most likely.

While the first conclusion provides an example of the contribution that the biologists have provided in this collaboration, i.e., interpreting theoretical results from a biological perspective, the information has flowed both ways. For example, another major result from the optimization study was that the environmental density

function plays a vital role in determining niche width (Chapter 2). After light data were collected from Mushroom Spring over a two-year period and the environmental density function was plotted, an important fact was revealed: over the course of a year, low-light intensities occur with much more regularity than high-light intensities (Chapter 2, Figure 2.5B). The model results, using light as the temporally varying environmental parameter, would then suggest that winter may actually be a better growing season than summer. Our empiricist collaborators that have studied these systems for many years were initially hesitant to accept this implication, but other observations that were made through this collaborative effort support this hypothesis. For example, the isolates whose light responses were measured under CO_2-limiting conditions all exhibited low-light adaptations (65AY6Li, 63AY4M2, 60AY4M2, and 63AY4M1, Figure 1.3), and these conditions are thought to be representative of *in situ* conditions due to the observed rise in pH at midday (Revsbech and Ward, 1984). In fact, none of these strains were able to withstand continuous light representative of the maximum light intensity that has been measured on a typical summer day. Specifically, both the model results and the growth rate behavior of a strain representative of PE A1 (grown under CO_2-limiting conditions) suggested that the upper-light limit for growth was roughly equivalent to the peak light intensity in December, the time of the year when light is most limited (Appendix A, Figure A.6). This provides an example of how simple theoretical models

can be used, not to exactly predict nature, but to assist in identifying the most important features in a given system.

As alluded to in the introduction, one of the key steps in applying any optimization approach is the choice of the objective function. In the appendices of Chapter 2 and in Chapter 3 the competitive ability of the optimal species was investigated in order to validate the choice of minimizing the geometric mean of substrate. In Chapter 2, results from numerical competition experiments over three frequency regimes were presented (Appendix A.2), and suggested that the optimal species was able to outcompete any other sub-optimal species in a two-species chemostat. Analytical evidence was sought in Chapter 3 via a non-autonomous stability analysis in order to confirm the numerical results. It was shown that in the high frequency regime the optimal species was able to outcompete any other species in the two-species model, but, in the low frequency regime, only certain invasibility results were obtained. Specifically, it was shown that the optimal species was the only species that could invade any other species, but the test to show that the optimal species was unable to be invaded, was inconclusive (due to one of the Floquet exponents being identically zero). Further investigation is currently underway (Appendix B.2). It should be noted that in order to conclude that minimizing the geometric mean of substrate does indeed result in a species that can outcompete any other, global stability properties must be determined. The stability analysis discussed in Chapter 3 was a local

stability analysis, and I have not yet studied the global stability properties of this system.

As a final note with respect to the theoretical work, further investigation into the role of the optimal species on observed community structure is also currently being investigated. Specifically, the interaction of the optimal species when other sub-optimal species are added to the chemostat is of interest. Initial investigations using a three species chemostat are underway, and at this time I have been unable to identify two sub-optimal species that can cooperatively exclude the optimal species in numerical simulations, although I hypothesize that such a situation does exist. The long-term behavior of the two and three species systems with the optimal species and other suboptimal species has the potential to reveal information regarding the effects of community dynamics on observed community structure.

My work in the laboratory regarding attempts to cultivate representative strains of the predominant *Synechococcus* organisms inhabiting the microbial mats of Mushroom Spring has resulted in some successes and some failures. Like others (Allewalt et al., 2006), I have been unable to obtain microscopic cells counts of *Synechococcus* spp. from an original mat sample that match the number of colony forming units that form on dilution plates. For example, a standard mat sample contains approximately 10^{10} cells/mL, yet when one mL of the sample is serially diluted ten-fold, colonies do not form on the 10^{-10} dilution plate. One success is that I have developed a protocol (Chapter 4) that results in colony growth out to the 10^{-8} dilution

plate sometimes, and out to 10^{-6} dilution plate with regularity. This is considered a success because the most successful attempt to date, before me, resulted in growth out to the 10^{-5} dilution plate sometimes, and only out to 10^{-3} with regularity (Allewalt et al., 2006). Many theories have been proposed to explain the discrepancy in the number of colonies that form on plates compared to the number of cells in the inoculum, most notably a nutrition deficiency theory (Allewalt, 2004). But, if that were the case, it should be possible to adjust the growth medium in a way that would result in obtaining colony growth on the 10^{10} dilution plate. One of my failures is that I have been unable to identify these missing nutrients. Obtaining isolates from highly-diluted mat samples is essential to ensure that the cultivated organisms represent the predominant organisms found *in situ*. It should be noted that I have obtained isolates representative of four of the five predominant *Synechococcus* species populations found in Mushroom Spring at 60°C, as so demarcated by sequence variation in the *psaA* gene, but have been trying for years, without success, to obtain the fifth one. Recent cultivation efforts suggest that progress in obtaining a representative of this uncultivated PE has been made (Chapter 5), and the sequence results that are currently in progress will either confirm or deny this statement. One of the important discoveries made through the cultivation efforts was that by exploiting the environmental data that have been collected in the past, I was able to achieve moderate success in predicting the sampling sites and incubation

conditions that would yield the targeted isolates. Further improvements in cultivation methods must be made, and a representative of the uncultivated organism must be isolated, in order to fully understand the community structure and function within the *Synechococcus* genus at Mushroom Spring.

With the *Synechococcus* isolates that I have obtained, distinct ecological adaptations with respect to light, temperature, and CO_2 have been observed among these closely-related organisms. In Chapter 4 light responses at 52 and 60°C were presented for isolated *Synechococcus* strains that are representative of three of the predominant PEs found *in situ* (Figures 4.5 and 4.6). The closely-related strains not only exhibited distinct light adaptations, but their light adaptations also corresponded to vertical distributions of the represented PEs in the 60°C Mushroom Spring microbial mat (Figure 4.1B). That is, the strain exhibiting the highest light adaptation corresponded to the PE that was most abundant near the surface. The other three strains were found to have lower light adaptations, and the PEs that these strains represent were found to be most abundant in the subsurface of the mat. Furthermore, the genome sequence data have revealed that the isolates representative of PEs that were found in relatively high abundance in the subsurface of the mat (Figure 4.1B) all possess certain light-harvesting antennae genes that the other isolates do not possess (M. Olsen, unpublished). This suggests that light quality may also be driving adaptation. Future experiments have been planned to test the this hypothesis by using filters that mimic the *in situ* light environments of

organisms inhabiting different depth intervals. The isolates that were obtained to study the growth rates of the predominant organisms with respect to light intensity, can now be used to study the adaptations of these organisms with respect to certain wavelengths of visible light.

In Chapter 4 it was also shown that the form of dissolved inorganic carbon available, and the temperature at which the experiments were conducted, played an important role in determining the light responses of the isolated strains (Figures 4.6 and 4.8). First, much higher growth rates were observed when 6% CO_2 in air was bubbled to the cultures versus when CO_2 was provided only by diffusion in air. And second, all of the strains exhibited a higher growth rate at 60°C, at all light intensities, than they did at 52°C. Along with the hypothesized importance of wavelength to light, based on the presence or absence of putative light-harvesting antennae genes, the effects of CO_2 and temperature on light responses highlight the need to consider the interconnectivity of various environmental parameters when attempting to quantify the fundamental niche of an organism.

The species demarcation method applied here, namely the use of sequence variation in the *psaA* gene, is not necessarily, and probably not even likely, to be the best possible species demarcation method. In fact, the results presented in Chapter 5 suggested that isolated strains within PEs A1 and A14 had similar growth rates over the range of light intensities that the vertical distributions suggest that these PEs are exposed to in nature. But, at higher light intensities, the three strains within

PE A14 exhibited different growth rates with respect to light. It was also observed that, at all light intensities, the two strains representative of PE A6 exhibited very different growth rates. Although the *psaA* gene encodes a protein and, as such, is more highly diverged than the 16S rRNA gene that typically has been used (97% nucleotide identity; Stackebrandt and Goebel, 1994) to demarcate microbial species populations, it is still only a single gene. The observed variability in light responses within *psaA* PEs further supports the need to address the microbial species issue. For example, if the 16S rRNA molecular cut-offs mentioned above are applied to *Synechococcus* species in the Mushroom Spring community, the definition of a predominant organism changes because several ecologically distinct species are lumped together. This causes a problem if the characteristics of an isolate that is representative of a relatively rare organism, whose 16S rRNA sequence is identical or nearly identical to a more predominant organism, are used to make inferences about niche structure. Precisely this has happened on at least two occasions (Allewalt et al., 2006; Kilian et al., 2007), and results from experiments conducted in these studies have shown a mismatch between what is found in nature compared to what was observed in the laboratory (regarding light adaptations). A multi-locus approach is likely to be more effective than a single-locus approach in identifying the ecologically distinct populations, and investigations of this nature have been initiated (Melendrez et al., in preparation; M. Olsen and D.M. Ward, unpublished). One advantage of having obtained the isolates and their corresponding genome sequences is that it

may now be possible to identify a set of genes that can be used to more accurately discern the ecologically distinct populations. A rather lofty, but possibly attainable goal of the comparative genomic analyses, is that this effort may contribute to a universally-accepted definition of microbial species.

A final conclusion of importance is: I have adapted. I started working in the Ward lab approximately five years ago as a mathematics graduate student with little knowledge in any field outside of mathematics. Not only did my thesis project afford me the opportunity to collaborate with researchers in other fields, but my project also required me to submerse myself in the field of microbial ecology. I have spent a significant amount of time at the field sites in Yellowstone National Park, and in the laboratory doing benchwork, and now, five years later, I can confidently say that my experience has given me a unique perspective to apply to mathematically modeling ecological systems.

In summary, mathematical and microbiological methods have been combined to investigate niche character in a temporally varying environment. The collaborative approach that was taken has been beneficial to both the mathematicians and the biologists involved in this project far beyond what would have been possible otherwise. More importantly, the information that has been exchanged has advanced our overall understanding of temporal effects on niche structure.

REFERENCES CITED

[1] Allewalt, J. A. (2004). "Temperature and light adaptations of *Synechococcus* isolates from a hot spring microbial community." M.S. thesis. Montana State University, Bozeman, MT.

[2] Allewalt, J. P., M. M. Bateson, N. P. Revsbech, K. Slack and D. M. Ward (2006). "Effect of temperature and light on growth of and photosynthesis by *Synechococcus* isolates typical of those predominating in the Octopus Spring microbial mat community of Yellowstone National Park." *Applied and Environmental Microbiology* **72**:544-550.

[3] Alzer, H. (1996). "A proof of the arithmetic mean-geometric mean inequality." *The American Mathematical Monthly* **103**:585.

[4] Angilletta, M. J. and M. W. Sears (2011). "Coordinating theoretical and empirical efforts to understand the linkages between organisms and environments." *Integrative and Comparative Biology* **51**:653-661.

[5] Baca, I., D. Sprockett and V. Dvornyk (2010). "Circadian input kinases and their homologs in cyanobacteria: evolutionary constraints versus architectural diversification." *Journal of Molecular Evolution* **70**:453-465.

[6] Barros, S. P. and S. Offenbacher (2009). "Epigenetics: connecting environment and genotype to phenotype and disease." *Journal of Dental Research* **88**:400-408.

[7] Bauld, J. and T. D. Brock (1974). "Algal excretion and bacterial assimilation in hot spring algal mats." *Journal of Phycology* **10**:101-106.

[8] Becraft, E. D. (2014). "The ecology and evolution of thermophilic *Synechococcus*." Ph.D. thesis. Montana State University, Bozeman, MT.

[9] Becraft, E. D., F. M. Cohan, M. Kühl, S. I. Jensen and D. M. Ward (2011). "Fine-scale distribution patterns of *Synechococcus* ecological diversity in microbial mats of Mushroom Spring, Yellowstone National Park." *Applied and Environmental Microbiology* **77**:7689-7697.

[10] Becraft, E. D., F. M. Cohan, M. Kühl, S. I. Jensen, D. B. Rusch, J. M. Wood and D. M. Ward. "Pyrosequencing analyses of *Synechococcus* ecological species inhabiting the microbial mat of Mushroom Spring, Yellowstone National Park." Submitted.

[11] Bernstein, H. C., J. P. Beam, M. A. Kozubal, R. P. Carlson and W. P. Inskeep (2013). "*In situ* analysis of oxygen consumption and diffusive transport in high-temperature acidic iron-oxide microbial mats." *Environmental Microbiology* **15**:2360-2370.

[12] Bhaya, D., A. R. Grossman, A. S. Steunou, N. Khuri, F. M. Cohan, N. Hamamura, M. C. Melendrez, M. M. Bateson, D. M. Ward and J. F. Heidelberg (2007). "Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analysis." *The ISME Journal* **1**:703-713.

[13] Brock, T. D. (1978). Thermophilic Microorganisms and Life at High Temperatures. Springer Verlag, New York.

[14] Broyden, C. G. (1965). "A class of methods for solving nonlinear simultaneous equations." *Mathematics of Computation* **19**:577-593.

[15] Buckley, L. B., M. C. Urban, M. J. Angilletta, L. G. Crozier, L. J. Rissler and M. W. Sears (2010). "Can mechanism inform species' distribution models?." *Ecology Letters* **13**:1041-1054.

[16] Castenholz, R. W. (1969). "Thermophilic blue-green algae and the thermal environment." *Bacteriological Reviews* **33**:476-504.

[17] Chase, J. M. and M. A. Leibold (2003). Ecological Niches: Linking Classical and Contemporary Approaches. University of Chicago Press.

[18] Chesson, P. (2000). "Mechanisms of maintenance of species diversity." *Annual Review of Ecology and Systematics* **31**:343-366.

[19] Chesson, P. (2008). "Quantifying and testing species coexistence mechanisms." *In* Unity in Diversity: Reflections on Ecology after the Legacy of Ramon Margalef. F. Valladares, A. Camacho, A.Elosegui, C. Gracia, M. Estrada, J. C. Senar, and J. M. Gili (eds.). Fundacion BBVA: Bilbao, Spain, pp 119-164.

[20] Chesson, P. and N. Huntly (1997). "The roles of harsh and fluctuating conditions in the dynamics of ecological communities." *The American Naturalist* **150**: 519-553.

[21] Chevin L-M., R. Lande and G. M. Mace (2010). "Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory." *PLoS Biology* **8**:e1000357.

[22] Chicone, C. (2006). Ordinary Differential Equations with Applications. Springer.

[23] Cohan, F. M. (2002). "What are bacterial species?." *Annual Reviews in Microbiology* **56**:457-487.

[24] Cohan, F. M. (2006). "Towards a conceptual and operational union of bacterial systematics, ecology and evolution." *Philosophical Transactions of the Royal Society B* **361**:1985-1996.

[25] Cohan, F. M. and E. B. Perry (2007). "A systematics for discovering the fundamental units of bacterial diversity." *Current Biology* **17**:R373-R386.

[26] Descamps-Julien, B. and A. Gonzalez (2005). "Stable coexistence in a fluctuating environment: an experimental demonstration." *Ecology* **86**:2815-2824.

[27] Dobzhansky, T. (1956). "What is an adaptive trait?." *The American Naturalist* **90**:337-347.

[28] Doebeli, M. and U. Dieckmann (2003). "Speciation along environmental gradients." *Nature* **421**:259-264.

[29] Drake, J. W. (2009). "Avoiding dangerous missense: thermophiles display especially low mutation rates." *PLoS Genetics* **5**: e1000520.

[30] Ferris, M. J., A. L. Ruff-Roberts, E. D. Kopczynski, M. M. Bateson and D. M. Ward (1996). "Enrichment culture and microscopy conceal diverse thermophilic *Synechococcus* populations in a single hot spring microbial mat habitat." *Applied and Environmental Microbiology* **62**:1045-1050.

[31] Ferris, M. J. and D. M. Ward (1997). "Seasonal distributions of dominant 16S rRNA-defined populations in a hot spring microbial mat examined by denaturing gradient gel electrophoresis." *Applied and Environmental Microbiology* **63**:1375-1381.

[32] Ferris, M. J., S. C. Nold, N. P. Revsbech and D. M. Ward (1997). "Population structure and physiological changes within a hot spring microbial mat community following disturbance." *Applied and Environmental Microbiology* **63**:1367-1374.

[33] Ferris, M. J., M. Kühl, A. Wieland and D. M. Ward (2003). "Different light-adapted ecotypes in a 68°C *Synechococcus* mat community revealed by analysis

of 16S-23S intervening transcribed spacer variation." *Applied and Environmental Microbiology* **69**:2893-2898.

[34] Gause, G. F. (1934). The Struggle for Existence. Williams and Wilkins. Baltimore, Maryland.

[35] Gilchrist, G. W. (1995). "Specialists and generalists in changing environments. I. Fitness landscapes of thermal sensitivity." *American Naturalist*, **146**:252-270.

[36] Gilchrist, G. W. (2000). "The evolution of thermal sensitivity in changing environments." *In* Cell and Molecular Responses to Stress. J. M Storey and K. B Storey (eds.). Elsevier, pp 55-70.

[37] Giovannoni, S. J. and U. Stingl (2005). "Molecular diversity and ecology of microbial plankton." *Nature* **437**:343-348.

[38] Gomulkiewicz, R. and D. Houle (2009). "Demographic and genetic constraints on evolution." *The American Naturalist* **174**:E218-E229.

[39] Gyllenberg, M. and G. Meszéna (2005). "On the impossibility of coexistence of infinitely many strategies." *Journal of Mathematical Biology* **50**:133-160.

[40] Haldane, J. B. S. and S. D. Jayakar (1963). "Polymorphism due to selection of varying direction." *Journal of Genetics* **58**:237-242.

[41] Holmes, M. H. (1995). Introduction to Perturbation Methods. Springer.

[42] Huey, R. B. and J. G. Kingsolver (1993). "Evolution of resistence to high temperature in ectotherms." *American Naturalist* **142**:S21-S46.

[43] Hutchinson, G. E. (1957). "Concluding remarks." *In* Cold Spring Harbor Symposia on Quantitative Biology. Cold Spring Harbor Laboratory Press, pp 415-427.

[44] Hutchinson, G. E. (1961). "The paradox of the plankton." *The American Naturalist* **882**:137-145.

[45] Jackson, B., S. Nowack, E. D. Becraft, I. Klapper and D. M. Ward. "Relationship between microorganisms inhabiting benthic communities in alkaline siliceous hot springs and overflowing water." In preparation.

[46] Jacquet, S., F. Partensky, D. Marie, R. Casotti, and D. Vaulot (2001). "Cell cycle regulation by light in *Prochlorococcus* strains." *Applied and Environmental Microbiology* **67**:782-790.

[47] Jensen, S. I., A. S. Steunou, D. Bhaya, M. Kühl and A. R. Grossman (2010). "*In situ* dynamics of O_2, pH and cyanobacterial transcripts associated with CCM, photosynthesis and detoxification of ROS." *The ISME Journal* **5**:317-328.

[48] Kilian, O., A. S. Steunou, F. Fazeli, S. Bailey, D. Bhaya and A. R. Grossman (2007). "Responses of a thermophilic *Synechococcus* isolate from the microbial mat of Octopus Spring to light." *Applied and Environmental Microbiology* **73**:4268-4278.

[49] Klapper, I. and J. Dockery (2010). "Mathematical description of microbial biofilms." *SIAM Review* **52**:221-265.

[50] Koeppel, A., E. B. Perry, J. Sikorski, D. Krizanc, A. Warner, D. M. Ward, A. P. Rooney, E. Brambilla, N. Connor, R. M. Ratcliff, E. Nevo and F. M. Cohan (2008). "Identifying the fundamental units of bacterial diversity: A paradigm shift to incorporate ecology into bacterial systematics." *Proceedings of the National Academy of Sciences* **105**:2504-2509.

[51] Kondo, T. and M. Ishiura (2000). "The circadian clock of cyanobacteria." *Bioessays* **22**:10-15.

[52] Krimbas, C. B. (2004). "On fitness." *Biology and Philosophy* **19**:185-203.

[53] Kühl, M. (2005). "Optical microsensors for analysis of microbial communities." *Methods in Enzymology* **397**:166-199.

[54] Kühl, M., C. Lassen and B. B. Jørgensen (1994). "Optical properties of microbial mats: light measurements with fiber-optic microprobes." *In* Microbial Mats: Structure, Development and Environmental Significance. L. J. Stal and P. Caumette (eds.). Springer, Berlin, pp 149-167.

[55] Leimar, O., M. Doebeli and U. Dieckmann (2008). "Evolution of phenotypic clusters through competition and local adaptation along an environmental gradient." *Evolution* **62**:807-822.

[56] Lenas, P. and S. Pavlou (1995). "Coexistence of three competing microbial populations in a chemostat with periodically varying dilution rate." *Mathematical Biosciences* **129**:111-142.

[57] Levins, R. (1968). Evolution in Changing Environments: Some Theoretical Explorations. Princeton University Press.

[58] Litchman, E. and C. A. Klausmeier (2001). "Competition of phytoplankton under fluctuating light." *The American Naturalist* **157**:170-187.

[59] Luo, C., S. T. Walk, D. M. Gordon, M. Feldgarden, J. M. Tiedje and K. T. Konstantinidis (2011). "Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species." *Proceedings of the National Academy of Sciences* **108**:7200-7205.

[60] Lynch, M. and W. Gabriel (1987). "Environmental tolerance." *American Naturalist*, **129**:283-303.

[61] Lynch, M. and R. Lande (1993). "Evolution and extinction in response to environmental change." *In* Biotic Interactions and Global Change. P. Kareiva, J. Kingsolver, and R. Huey (eds.). Sinauer, Sunderland, MA, pp 234-250.

[62] May, R. M. and R. H. MacArthur (1972). "Niche overlap as a function of environmental variability." *Proceedings of the National Academy of Sciences* **69**:1109-1113.

[63] Melendrez, M. C., R. K. Lange, F. M. Cohan and D. M. Ward (2011). "Influence of molecular resolution on sequence-based discovery of ecological diversity among *Synechococcus* populations in an alkaline siliceous hot spring microbial mat." *Applied and Environmental Microbiology* **77**:1359-1367.

[64] Melendrez, M. C., J. M. Wood, E. D. Becraft, F. M. Cohan, J. F. Heidelberg, D. Rusch and D. M. Ward. "The relative importance of recombination and selection on the evolution of *Synechococcus* species inhabiting a hot spring cyanobacterial mat." In preparation.

[65] Miller, M. B. and B. L. Bassler (2001). "Quorum sensing in bacteria." *Annual Review of Microbiology* **55**:165-199.

[66] Miller, S. R. and R. W. Castenholz (2000). "Evolution of thermotolerance in hot spring cyanobacteria of the genus *Synechococcus*." *Applied and Environmental Microbiology* **66**:4222-4229.

[67] Moore, L. R., G. Rocap and S. W. Chisholm (1998). "Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes." *Nature* **393**:464-467.

[68] Nagylaki, T. (1975). "Polymorphisms in cyclically varying environments." *Heredity* **35**:67-74.

[69] Norberg, J., D. P. Swaney, J. Dushoff, J. Lin, R. Casagrandi, and S. A. Levin (2001). Phenotypic diversity and ecosystem functioning in changing environments: a theoretical framework. *Proceedings of the National Academy of Sciences* **98**: 11376-11381.

[70] Nocedal, J. and S. J. Wright (2006). Penalty and Augmented Lagrangian Methods. Springer, New York.

[71] Nowack, S., I. Klapper and D. M. Ward. "Consequences of temporal frequency regime on optimal behavior." Submitted.

[72] Papke, R. T., N. B. Ramsing, M. M. Bateson and D.M. Ward (2003). "Geographical isolation in hot spring cyanobacteria." *Environmental Microbiology* **5**:650-659.

[73] Peary, J. A. and R. W. Castenholz (1964). "Temperature strains of a thermophilic blue-green alga." *Nature* **202**:720-721.

[74] Pfeiffer, T. and S. Bonhoeffer (2004). "Evolution of cross-feeding in microbial populations." *American Naturalist* **163**:E126-E135.

[75] Pigolotti, S., C. Lopez and E. Hernandez-Garcia (2007). "Species clustering in competitive Lotka-Volterra models." *Physical Review Letters* **98**:258101.

[76] Pittendrigh, C. S. (1993). "Temporal organizations: reflections of a Darwinian clock-watcher." *Annual Review of Physiology* **55**:17-54.

[77] Pittman, J. K., A. P. Dean and O. Osundeko (2011). "The potential of sustainable algal biofuel production using wastewater resources." *Bioresource Technology* **102**:17-25.

[78] Powell, M. J. D. (1978). "A fast algorithm for nonlinearly constrained optimization calculations." *In* Numerical Analysis. Springer, Berlin, pp 144-157.

[79] Radakovits, R, R. E. Jinkerson, A. Darzins and M. C. Posewitz (2010). "Genetic engineering of algae for enhanced biofuel production." *Eukaryotic Cell* **9**:486-501.

[80] Ramsing, N. B., M. J. Ferris and D. M. Ward (2000). "Highly ordered vertical structure of *Synechococcus* populations within the one-millimeter thick photic zone of a hot spring cyanobacterial mat." *Applied and Environmental Microbiology* **66**:1038-1049.

[81] Rasmussen, M. (2007). Attractivity and Bifurcation for Nonautonomous Dynamical Systems. Springer.

[82] Reese, M. O., S. A. Gevorgyan, M. Jørgensen, E. Bundgaard, S. R. Kurtz, D. S. Ginley, D. C. Olson et al. (2011). "Consensus stability testing protocols for organic photovoltaic materials and devices." *Solar Energy Materials and Solar Cells* **95**:1253-1267.

[83] Retchless, A. C. and J. G. Lawrence (2007). "Temporal fragmentation of speciation in bacteria." *Science* **317**:1093-1096

[84] Revsbech, N. P. and D. M. Ward (1984). "Microelectrode studies of interstitial water chemistry and photosynthetic activity in a hot spring microbial mat." *Applied and Environmental Microbiology* **48**:270-275.

[85] Rinnan, R., A. Michelsen, E. Baath and S. Jonasson (2007). "Fifteen years of climate change manipulations alter soil microbial communities in a subarctic heath ecosystem." *Global Change Biology* **13**:28-39.

[86] Roukos, D. H., C. Katsios and T. Liakakos (2010). "Genotype-phenotype map and molecular networks: a promising solution in overcoming colorectal cancer resistance to targeted treatment." *Expert Review of Molecular Diagnostics* **10**:541-545.

[87] Sakamoto, T., V. B. Delgaizo and D. A. Bryant (1998). "Growth on urea can trigger death and peroxidation of the cyanobacterium *Synechococcus* sp. strain PCC 7002." *Applied and Environmental Microbiology* **64**:2361-2366.

[88] Santegoeds, C. M., S. C. Nold and D. M. Ward (1996). "Denaturing gradient gel electrophoresis used to monitor the enrichment culture of aerobic chemoorganotrophic bacteria from a hot spring cyanobacterial mat." *Applied and Environmental Microbiology* **62**:3922-3928.

[89] Schwenk, K., D. K. Padilla, G. S. Bakken and R. J. Full (2009). "Grand challenges in organismal biology." *Integrative and Comparative Biology* **49**:7-14.

[90] Siepielski, A. M. and M. A. McPeek (2010). "On the evidence for species coexistence: a critique of the coexistence program." *Ecology* **91**:3153-3164.

[91] Slatkin, M. and R. Lande (1976). "Niche width in a fluctuating environment-density independent model." *American Naturalist* **110**:31-55.

[92] Smith, H. L. and P. Waltman (1995). The Theory of the Chemostat. Cambridge University Press, Cambridge, UK.

[93] Stackebrandt, E and B. M. Goebel (1994). "Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology." *International Journal of Systematic Bacteriology* **44**:846-849.

[94] Steunou, A. S., D. Bhaya, M. M. Bateson, M. C. Melendrez, D. M. Ward, E. Brecht, J. W. Peters, M. Kühl and A. R. Grossman (2006). "*In situ* analysis of nitrogen fixation and metabolic switching in unicellular thermophilic cyanobacteria inhabiting hot spring microbial mats." *Proceedings of the National Academy of Sciences of the United States of America* **103**:2398-2403.

[95] Stewart, F. M. and B. R. Levin (1973). "Partitioning of resources and the outcome of interspecific competition: a model and some general considerations." *American Naturalist* **103**:171-198.

[96] Stumm, W. and J. J. Morgan (2012). Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters. John Wiley and Sons.

[97] Swenson, W., J. Arendt and D. S. Wilson (2000). "Artificial selection of microbial ecosystems for 3-chloroaniline biodegradation." *Environmental Microbiology* **2**:564-571.

[98] Ward D. M. (1998). "A natural species concept for prokaryotes." *Current Opinions in Microbiology* **1**:271-277.

[99] Ward, D. M., M. J. Ferris, S. C. Nold and M. M. Bateson (1998). "A natural view of microbial biodiversity within hot spring cyanobacterial mat communities." *Microbiology and Molecular Biology Reviews* **62**:1353-1370.

[100] Ward, D. M. and F. M. Cohan (2005). "Microbial diversity in hot spring cyanobacterial mats: pattern and prediction." *In* Geothermal Biology and Geochemistry in Yellowstone National Park. W. P. Inskeep and T. McDermott (eds.). Thermal Biology Institute: Bozeman, MT, pp 185-201.

[101] Ward, D. M., M. M. Bateson, M. J. Ferris, M. Kühl, A. Wieland, A. Koeppel and F. M. Cohan (2006). "Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function." *Philosophical Transactions of the Royal Society B: Biological Sciences* **361**:1997-2008.

[102] Ward, D. M., R. W. Castenholz and S. R. Miller (2012). "Cyanobacteria in geothermal habitats." *In* Ecology of Cyanobacteria II. B. A. Whitton (ed.). Springer, Netherlands, pp 39-63.

[103] Wayne, L. G., D. J. Brenner, R. R. Colwell, P. A. D. Grimont, O. Kandler, M. I. Krichevsky, L. H. Moore, W. E. C. Moore, R. G. E. Murray et al.(1987). "Report of the ad hoc committee on reconciliation of approaches to bacterial systematics." *International Journal of Systematic and Evolutionary Microbiology* **37**:463-464.

[104] Weltzer, M. L. and S. R. Miller (2013). "Ecological divergence of a novel group of *Chloroflexus* strains along a geothermal gradient." *Applied and Environmental Microbiology* **79**:1353-1358.

[105] West, N. J. and D. J. Scanlan (1999). "Niche-partitioning of *Prochlorococcus* populations in a stratified water column in the eastern North Atlantic Ocean." *Applied and Environmental Microbiology* **65**:2585-2591.

[106] Wolin, E. A., R. S. Wolfe and M. J. Wolin (1964). "Viologen dye inhibition of methane formation by *Methanobacillus omelianskii*." *Journal of Bacteriology* **87**:993-998.

[107] Wolkowicz, G. S. K. and X. Q. Zhao (1998). "N-species competition in a periodic chemostat." *Differential and Integral Equations* **11**:465-491.

[108] Xia, D. F., S. L. Xu, and F. Qi (1999). "A proof of the arithmetic mean-geometric mean-harmonic mean inequalities." *RGMIA Research Report Collection* **2**:99-102.

[109] Yamauchi, A. and T. Miki (2009). "Intraspecific niche flexibility facilitates species coexistence in a competitive community with a fluctuating environment." *Oikos* **118**:55-66.

APPENDICES

APPENDIX A

CHAPTER 2 APPENDICES

Appendix A.1. Solution Details of Model ODE

Multiplying the second equation in (2.2) by 1/Y and adding it to the first equation results in

$$\frac{d}{dt}\left(S+\frac{x}{Y}\right)=D\left(S^0-\left(S+\frac{x}{Y}\right)\right). \tag{A.1}$$

Letting $U=S+\dfrac{x}{Y}$ allows (A.1) to be rewritten as

$$\frac{dU}{dt}=D(S^0-U)$$

which has solution

$$U=S^0+C_1e^{-Dt},$$

where C_1 is an integration constant. Thus after $O(D^{-1})$ time,

$$x=Y\left(S^0-S\right), \tag{A.2}$$

to an exponentially small approximation. The ODE in (A.1) can then be converted into a standard Riccati equation (recall that T is a function of t)

$$\frac{dx}{dt}=\left(\frac{rS^0}{K_s}f(T)-D\right)x-\frac{rf(T)}{KY}x^2,$$

which is used to obtain the exact solution of (2.2)

$$S(t)=S^0-\frac{K_s}{r}\frac{e^{\int_0^t(\frac{rS^0}{K_s}f(T(B))-D)dB}}{\int_0^t e^{\int_0^B(\frac{rS^0}{K_s}f(T(\gamma))-D)d\gamma}f(T(B))dB+C} \tag{A.3}$$

$$x(t)=\frac{K_sY}{r}\frac{e^{\int_0^t(\frac{rS^0}{K_s}f(T(B))-D)dB}}{\int_0^t e^{\int_0^B(\frac{rS^0}{K_s}f(T(\gamma))-D)d\gamma}f(T(B))dB+C},$$

where C is an integration constant determined by initial conditions.

Appendix A.2. Selection of Objective Function - Numerical Competition

The purpose of this appendix is to provide support for the choice of the objective function chosen in (2.4). We consider a two-species chemostat model and through numerical competition experiments we provide additional support for choosing to minimize the geometric mean of the nutrient concentration in the optimization calculations. The two species model is

$$\begin{aligned} \frac{dS}{dt} &= D(S^0 - S) - \frac{r}{K_s Y} S\left(x_1 f_1(T(t)) + x_2 f_2(T(t))\right) \\ \frac{dx_1}{dt} &= x_1\left(\frac{rS}{K_s} f_1(T(t)) - D\right) \\ \frac{dx_2}{dt} &= x_2\left(\frac{rS}{K_s} f_2(T(t)) - D\right). \end{aligned} \tag{A.4}$$

Note that all parameters are assumed to be non-negative with the parameters in both the one-species and two-species systems assumed to be equal. Also note $S, x_1, x_2, f_1, f_2 \geq 0 \ \forall \ t$.

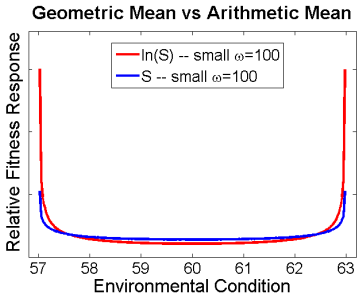
In silico competition experiments were run by solving (A.4) with a fourth-order Runge-Kutta method using the numerical formulation of the optimization problem defined in Appendix A.4. For each competition experiment an environment was chosen (e.g., low ($\omega = 100$), intermediate ($\omega = 1$) or high frequency ($\omega = .01$)), and two species competed for the available resource. Using the optimization method described in Appendix A.4, f_1 and f_2 were either replaced with one of the optimal fitness responses that was calculated numerically (and shown in Figure 2.3), or optimal Gaussian curves. All competing species were constrained according to (2.5).

Optimum from Geometric Mean versus Optimum from Arithmetic Mean

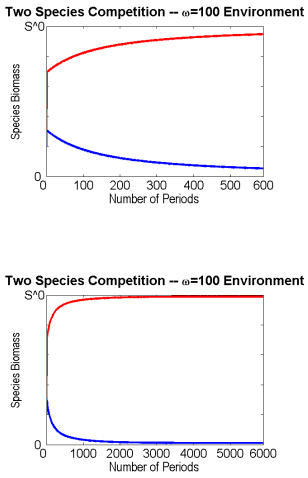
The first competition experiment competes species x_1 with fitness response f_1, corresponding to the optimal fitness response obtained from minimizing the geometric mean (in the low frequency regime) versus species x_2 with fitness response f_2, corresponding to the the optimal fitness response from minimizing the arithmetic mean (also in the low frequency regime). While both species were optimal in this particular environment (with respect to the objective function used in their respective calculations), and subjected to the same constraints, species x_1 appears to be more fit. The numerical results shown in Figure A.1, at least in the low frequency regime, suggest that the geometric mean (red) is a better choice than the arithmetic mean (blue) as an objective function for this optimization method. Note that competitive exclusion was rather slow, in the sense that decay of the excluded species takes place over thousands of cycles.

Competition of Optimal Species from Different Environments

In the next set of competition experiments the optimal species (with respect to the geometric mean) from one frequency regime was competed against the optimal species from another frequency regime (Figure A.2 and Figure A.3). In each experiment, with one exception (Figure A.3D, top), the optimal species for the selected environment competitively excluded any other species that was considered. The reason for apparent coexistence in the exceptional instance is unclear at this time and is possibly a result of numerical approximation. We have also included another instance of coexistence (Figure A.2B, middle). In this particular experiment both species were suboptimal, that is, when the low-frequency optimal species



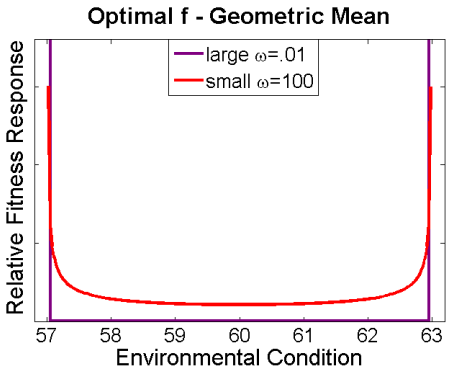
(A)



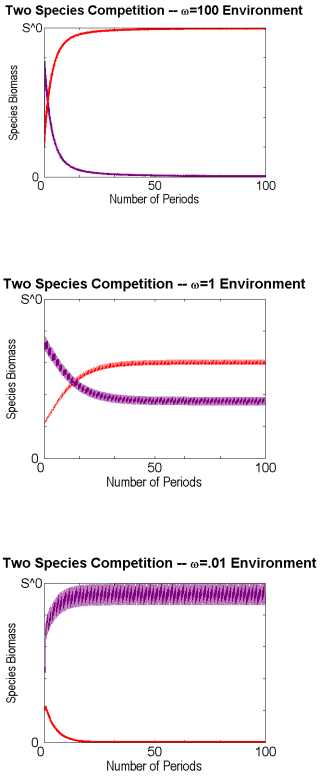
(B)

Figure A.1. Numerical competition between a species with fitness determined by minimizing the arithmetic mean versus a species with fitness determined by minimizing the geometric mean.(A) Plots of optimal fitness responses, f_1 and f_2, in the low frequency regime ($\omega = 100$). f_1 was computed by numerically minimizing the geometric mean (red) and f_2 was computed by numerically minimizing the arithmetic mean (blue). (B) Competition results of species x_1 with fitness f_1 versus species x_2 with fitness f_2 after 600 and then 6000 periods.

competes against the high-frequency optimal species in an environment undergoing intermediate fluctuations, they appear to coexist.



(A)

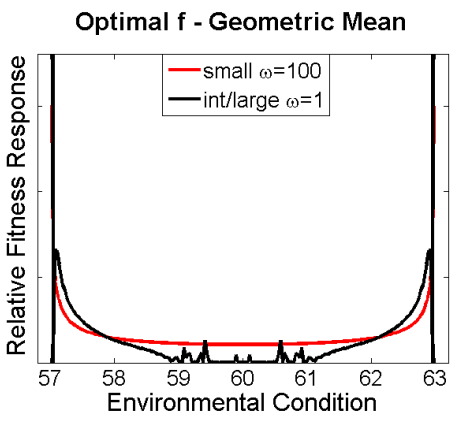


(B)

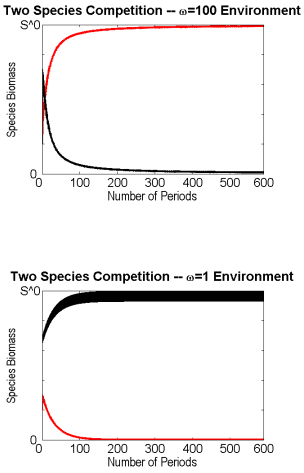
Figure A.2. Numerical competition experiments between optimal and sub-optimal species. Panel (A) shows a plot of the optimal fitness responses (with respect to the geometric mean), f_1 and f_2, for the competitors shown in the right panel. Colors of fitness responses in left panel correspond to colors of species in right panel. Panel (B) shows competition results of species x_1 with fitness f_1 versus species x_2 with fitness f_2 after 100 periods in environments that fluctuate under different frequency regimes. The optimal species in each environment competitively excluded the suboptimal species. Numerical coexistence is shown for a case (middle) when both species were suboptimal with respect to the environment.

Optimum without Specified Form Versus Gaussian Optimum - Geometric Mean

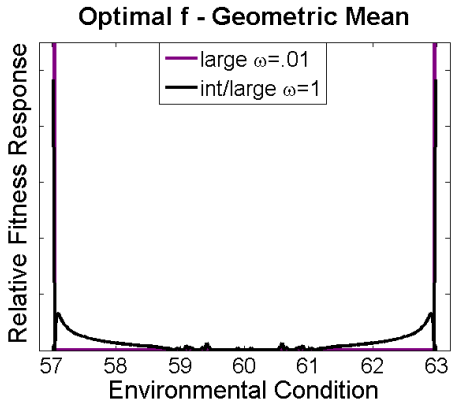
In this last set of competition experiments a presupposed Gaussian form was



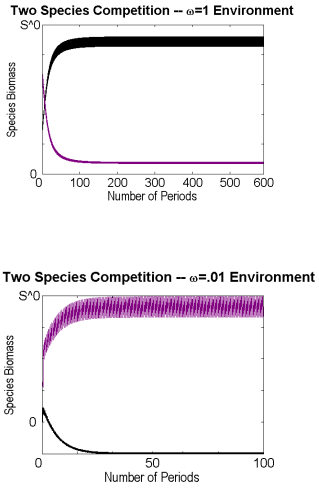
(A)



(B)



(C)



(D)

Figure A.3. Additional numerical competition experiments. Panels (A) and (C) show plots of optimal fitness responses (with respect to the geometric mean), f_1 and f_2, for the competitors shown in the right panel. Colors of fitness responses in the left panel correspond to colors of species in the right panel. Panels (B) and (D) show competition results of species x_1 with fitness f_1 versus species x_2 with fitness f_2 after either 100 or 600 periods (depending on rate of convergence) under environments that fluctuate under different frequency regimes. The optimal species won the competition battle in all but one of the depicted experiments (Fig. B3.D, $\omega = 1$ environment).

assumed and then subsequently, the mean and variance were optimized by minimizing the geometric mean of the nutrient concentration. In each case the species with optimal fitness resulting from the unrestricted form excluded the species with optimal Gaussian fitness (Figure A.4).

Appendix A.3. High and Low Frequency Asymptotics

High Frequency Asymptotic Analysis

Let $\mathbf{x} = (S, x)$. System (2.2) can then be written as

$$\dot{\mathbf{x}} = \mathbf{z}(t, \mathbf{x}(t)). \tag{A.5}$$

Suppose the chemostat operates under high frequency temperature oscillations, that is $\omega \gg \frac{rS^0}{K_s}\bar{f} - D$. Introducing two time scales (Holmes, 1995),

$$\begin{aligned} t_1 &= t \\ t_2 &= \epsilon^{-1} t, \end{aligned}$$

where $\epsilon = \omega^{-1}$, allows (A.5) to be rewritten as

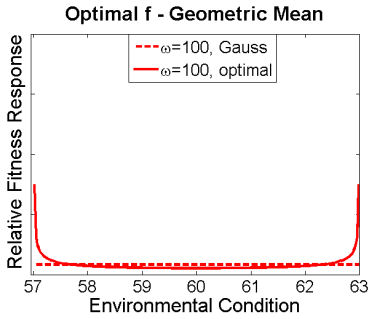
$$\dot{\mathbf{x}} = \mathbf{z}(t_1, t_2, \mathbf{x}(t_1, t_2)). \tag{A.6}$$

After expanding $\mathbf{x}$ in powers of ϵ as

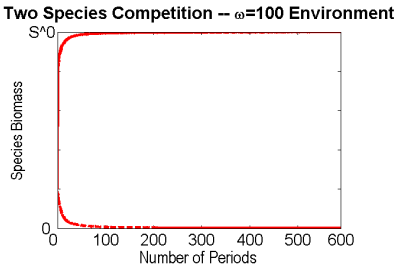
$$\mathbf{x} = \mathbf{x}_0(t_1, t_2) + \epsilon \mathbf{x}_1(t_1, t_2) + \epsilon^2 \mathbf{x}_2(t_1, t_2) + O(\epsilon^3), \tag{A.7}$$

the right-hand side of equation (A.6) becomes

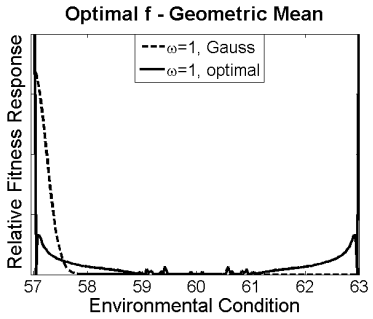
$$\mathbf{z}(t_1, t_2, \mathbf{x}(t_1, t_2)) \approx \mathbf{z}(t_1, t_2, \mathbf{x}_0 + \epsilon \mathbf{x}_1 + O(\epsilon^2)),$$



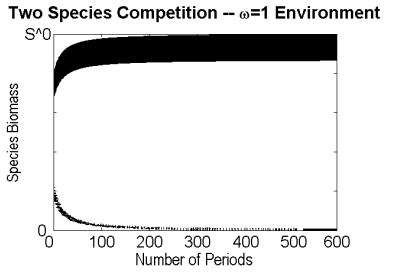
(A)



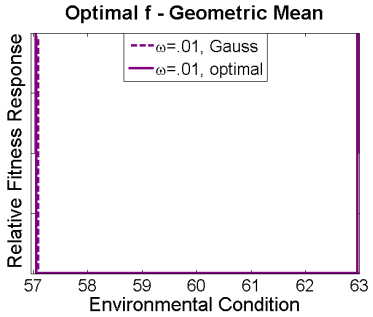
(B)



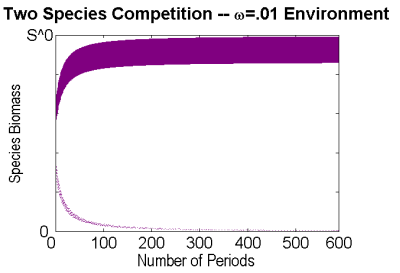
(C)



(D)



(E)



(F)

Figure A.4. Competition between optimal species and Gaussian species. Panels (A), (C), and (E) show plots of fitness responses, f_1 and f_2, for the competitors. The Gaussian curves (dashed) are optimal in the sense that the mean and average were optimized by minimizing the geometric mean of nutrient concentration (as was done without an assumed fitness form (solid)). Line styles of fitness responses in left panel correspond to line styles of competing species in right panel. Panels (B), (D), and (F) show numerical competition results. The species with Gaussian fitness was competitively excluded over all frequency regimes.

which has the two-term Taylor expansion about $\mathbf{x_0}$

$$\mathbf{z}(t_1, t_2, \mathbf{x}(t_1, t_2)) \approx \mathbf{z}(t_1, t_2, \mathbf{x}_0(t_1, t_2)) + \epsilon \mathbf{x}_1(t_1, t_2)\mathbf{z}'(t_1, t_2, \mathbf{x}_0(t_1, t_2)). \tag{A.8}$$

The prime here denotes the Jacobian of $\mathbf{z}$ with respect to $\mathbf{x}$. Note that

$$\frac{d}{dt} = \frac{\partial}{\partial t_1} + \frac{1}{\epsilon}\frac{\partial}{\partial t_2},$$

so differentiating both sides of (A.7) with respect to time results in

$$\dot{\mathbf{x}} = \frac{\partial \mathbf{x}_0}{\partial t_1} + \frac{1}{\epsilon}\frac{\partial \mathbf{x}_0}{\partial t_2} + \epsilon\left(\frac{\partial \mathbf{x}_1}{\partial t_1} + \frac{1}{\epsilon}\frac{\partial \mathbf{x}_1}{\partial t_2}\right) + \epsilon^2\left(\frac{\partial \mathbf{x}_2}{\partial t_1} + \frac{1}{\epsilon}\frac{\partial \mathbf{x}_2}{\partial t_2}\right) + ... \tag{A.9}$$

Equating like terms of the right-hand side of equation (A.8) with terms from the right-hand side of (A.9), justified by the relation in (A.6), yields

$$\begin{aligned} \epsilon^{-1} : \frac{\partial \mathbf{x}_0}{\partial t_2} &= 0 \\ \epsilon^{0} : \frac{\partial \mathbf{x}_0}{\partial t_1} + \frac{\partial \mathbf{x}_1}{\partial t_2} &= \mathbf{z}(t_1, t_2, \mathbf{x}_0(t_1, t_2)). \end{aligned}$$

Integrating both sides of the ϵ^0-component equations, over one cycle in fast time (t_2), results in

$$\begin{aligned} \int_{t_L}^{t_L+R}\left(\frac{\partial S_0}{\partial t_1} + \frac{\partial S_1}{\partial t_2}\right)dt_2 &= \int_{t_L}^{t_L+R}\left(D(S^0 - S_0) - \frac{r}{YK_s}S_0 x_0 f(t_2)\right)dt_2 \\ \int_{t_L}^{t_L+R}\left(\frac{\partial x_0}{\partial t_1} + \frac{\partial x_1}{\partial t_2}\right)dt_2 &= \int_{t_L}^{t_L+R} x_0\left(\frac{r}{K_s}S_0 f(t_2) - D\right)dt_2. \end{aligned}$$

Since x and S are cycling in fast time

$$\int_{t_L}^{t_L+R}\frac{\partial S_1}{\partial t_2}dt_2 = \int_{t_L}^{t_L+R}\frac{\partial x_1}{\partial t_2}dt_2 = 0.$$

Note that the order ϵ^{-1}-equation implies $\mathbf{x}_0$ is not a function of t_2, which implies

$$R\left(\frac{\partial S_0}{\partial t_1}\right) = \int_{t_L}^{t_L+R}\left(D(S^0 - S_0) - \frac{r}{YK_s}S_0 x_0 f(t_2)\right)dt_2$$
$$R\left(\frac{\partial x_0}{\partial t_1}\right) = \int_{t_L}^{t_L+R} x_0\left(\frac{r}{K_s}S_0 f(t_2) - D\right)dt_2. \tag{A.10}$$

But, since solutions that are in steady state on the slow time scale are of interest, i.e., solutions that are independent of t_1, $\mathbf{x}_0$ should then be constant. Thus, the left-hand sides of the (A.10) can be set to zero. The second equation can be solved for S_0, which can then be substituted into the first equation in (A.10) to obtain

$$S_0 = \frac{DK_s}{r\bar{f}} \tag{A.11}$$
$$x_0 = Y(S^0 - S_0),$$

where

$$\bar{f} = \frac{1}{R}\int_{t_L}^{t_L+R} f(t_2)dt_2 = \frac{\omega}{R}\int_{t_L}^{t_L+R/\omega} f(t)dt. \tag{A.12}$$

Low Frequency Asymptotic Analysis

An analogous procedure to that used in the high frequency case is followed for the low frequency case, that is $\omega << \frac{rS^0}{K_s}\bar{f} - D$. This time the two time scales are

$$t_1 = t$$
$$t_2 = \epsilon t,$$

where t_2 is now slow time, with $t_2 = \omega t$ and $\omega = \epsilon$. The calculation here is identical to that of the high frequency case except $\frac{dt_1}{dt} = 1$, and $\frac{dt_2}{dt} = \epsilon$. Expanding

$$\dot{\mathbf{x}} = \frac{\partial \mathbf{x}_0}{\partial t_1} + \epsilon\frac{\partial \mathbf{x}_0}{\partial t_2} + \epsilon\left(\frac{\partial \mathbf{x}_1}{\partial t_1} + \epsilon\frac{\partial \mathbf{x}_1}{\partial t_2}\right) + \epsilon^2\left(\frac{\partial \mathbf{x}_2}{\partial t_1} + \epsilon\frac{\partial \mathbf{x}_2}{\partial t_2}\right) + ... \tag{A.13}$$

and equating like terms again, this time the right-hand side of (A.13) equated with

the right-hand side of (A.8), yields

$$\begin{aligned} \epsilon^0 : \frac{\partial \mathbf{x}_0}{\partial t_1} &= \mathbf{z}(t_1, t_2, \mathbf{x}_0(t_1, t_2)) \\ \epsilon : \frac{\partial \mathbf{x}_0}{\partial t_2} + \frac{\partial \mathbf{x}_1}{\partial t_1} &= \mathbf{z}'(\mathbf{x}_0(t_1, t_2))\mathbf{x}_1. \end{aligned} \tag{A.14}$$

Since now (fast time) solutions that are in steady state on the fast time scale are of interest, the left-hand side of the ϵ^0 equation (A.14) should be zero. Its component equations are

$$\begin{aligned} 0 &= D(S^0 - S_0) - \frac{r}{YK_s} S_0 x_0 f(t_2) \\ 0 &= x_0 \left(\frac{r}{K_s} S_0 f(t_2) - D \right). \end{aligned}$$

Solving the second equation for S_0 (assuming $x_0 \neq 0$) and then substituting it into the first equation to obtain x_0 results in

$$\begin{aligned} S_0(t_2) &= \frac{DK_s}{rf(t_2)} \\ x_0(t_2) &= Y \left(S^0 - \frac{DK_s}{rf(t_2)} \right). \end{aligned} \tag{A.15}$$

Appendix A.4. Details of Optimization Calculations

High Frequency Optimization Details

To minimize $\int_{t_L}^{t_L+R/\omega} \ln S(t) dt$ asymptotically in the high frequency case, $S(t)$ is approximated with $S_0(t)$ from (2.6). Note that minimizing $\int_{t_L}^{t_L+R/\omega} \ln S_0(t) dt$ amounts to maximizing $\bar{f}$. In order to obtain the optimal fitness response $f(T)$ as a function of T, the environmental history function (e.g., temperature or light), a

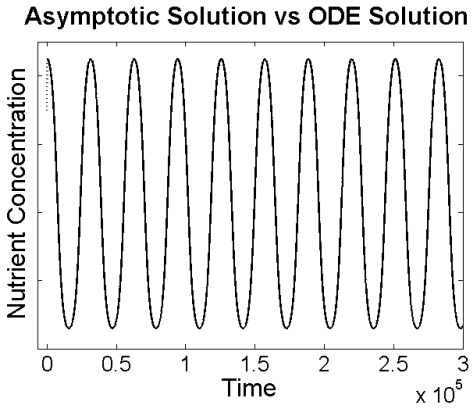


Figure A.5. Comparison of the numerical solution from an ODE solve in MATLAB (dotted) to the 0th-order asymptotic approximation (solid) for the low frequency case. The curves represent the species concentration x, and other than the region near $t = 0$, the curves are nearly indistinguishable.

change of variables is required. Let

$$F(\tau) := \int_0^{R/\omega} H\left(\tau - T(\alpha)\right) d\alpha,$$

where H is the Heaviside function and note that $F(\tau)$ measures the total time in a cycle that $T(\alpha) < \tau$. Now define

$$W(\tau)d\tau := \frac{dF}{d\tau}d\tau, \tag{A.16}$$

and note $W(\tau)d\tau$ measures the total time in a cycle that $\tau \leq T(\alpha) \leq \tau + d\tau$. The function W will be referred to as the environmental density function. The time integral in the $\bar{f}$ definition in equation (A.12) can now be changed to an integral of T,

$$\int_{t_L}^{t_L+\frac{R}{\omega}} f(T(t))dt = \int_{T_{\text{lo}}}^{T_{\text{hi}}} f(T)W(T)dT. \tag{A.17}$$

Note that

$$0 \leq \int_{T_{\text{lo}}}^{T_{\text{hi}}} f(T)W(T)dT \leq W(\tau) \int_{T_{\text{lo}}}^{T_{\text{hi}}} f(T)dT = W(\tau), \tag{A.18}$$

where $W(\tau) = \max_{T_{\text{lo}} \leq T \leq T_{\text{hi}}} W(T)$.

Let $M := \{\tau \in [T_{\text{lo}}, T_{\text{hi}}] : W(\tau) \geq W(T) \; \forall \; T \in [T_{\text{lo}}, T_{\text{hi}}]\}$ and let the number of elements in M be denoted by N. If we let $f(T) = \frac{1}{N} \sum_{\tau \in M} \delta(T - \tau)$ then $\int_{T_{\text{lo}}}^{T_{\text{hi}}} f(T)W(T)dT = W(\tau)$, where δ is the Dirac delta function. It can then be concluded that the right-hand side of (A.17) is maximized when $f(T) = \frac{1}{N} \sum_{\tau \in M} \delta(T-\tau)$. Another observation from (A.18) is that if $W(T)$ is constant with respect to the environment, which would be the case if say, the environment was a piecewise-linear function of time (over one environmental cycle), then any $f(T)$ that satisfies the constraint would be an optimizer.

Low Frequency Optimization Details

Approximating S with S_0 using (2.8) and applying the change of variables introduced in the subsection above, (2.4) and (2.5) take the form of an isoperimetric problem from the calculus of variations: minimize

$$L = \int_{T_{\text{lo}}}^{T_{\text{hi}}} \ln\left(\frac{DK_s}{rf(T)}\right) W(T)dT$$

subject to the constraint

$$G = \int_{T_{\text{lo}}}^{T_{\text{hi}}} f(T)dT = 1.$$

The Lagrangian, L^* is defined as

$$L^* = L + \lambda G = \ln\left(\frac{DK_s}{rf(T)}\right) W(T) + \lambda f(T),$$

where L is the integrand for the quantity that is to be minimized and G is the integrand in the constraint. Since the Lagrangian is not dependent on f' then

$L_f^* = 0$ at a minimum, so that,

$$L_f^* = -\frac{W(T)}{f(T)} + \lambda = 0$$

and thus

$$f(T) = \frac{W(T)}{\lambda}. \tag{A.19}$$

Integrating both sides of (A.19) from T_{lo} to T_{hi} and using the supplied constraint to solve for λ yields the result

$$\begin{aligned} \lambda &= \frac{\int_{T_{\text{lo}}}^{T_{\text{hi}}} W(T)dT}{\beta} \\ f(T) &= \beta \frac{W(T)}{\int_{T_{\text{lo}}}^{T_{\text{hi}}} W(T)dT}. \end{aligned} \tag{A.20}$$

Intermediate Frequency Optimization Details: A Numerical Approach

To perform the numerical calculations a fitness response function that is piecewise constant with respect to T is assumed, specifically, $f(T) = \alpha_i$, where $T_{i-1} \leq T \leq T_i$, with $T_i = T_0 + \frac{T_{\text{hi}} - T_{\text{lo}}}{n} i$, $1 \leq i \leq n$ and n is the number of subintervals over which T is equally partitioned. Next, define

$$G(\boldsymbol{\alpha}) := \int_{t_L}^{t_L + R/\omega} \ln S(t, \boldsymbol{\alpha}) dt \tag{A.21}$$

$$H(\boldsymbol{\alpha}) := \Delta T \left(\sum_{i=1}^{n} \alpha_i \right) - 1, \tag{A.22}$$

where G and H are the objective function and constraint from (2.4) and (2.5), respectively, rewritten using the piecewise constant assumption for f. In this context the optimization is with respect to $\boldsymbol{\alpha}$ and the resulting Kuhn-Karush-Tucker (KKT)

equations are

$$\nabla G = \lambda \nabla H - \boldsymbol{\mu} \tag{A.23}$$
$$H = 0$$
$$-\boldsymbol{\alpha} \leq \mathbf{0},$$

where λ and $\boldsymbol{\mu}$ are Lagrange multipliers and ∇ is the gradient with respect to $\boldsymbol{\alpha}$. The solutions of system (A.23) provide the necessary conditions required to minimize (A.21) subject to (A.22) (Nocedal and Wright, 2006), and these solutions were computed numerically using a sequential quadratic programming approach.

Appendix A.5. Derivation of $T_2(t)$

In the main body of Chapter 2 it was shown, both numerically and analytically, that in the low and high-frequency limiting cases that the optimal fitness response function was exclusively dependent on the environmental density function, $W(T)$. The complicated nature of the equations only allowed the intermediate cases to be studied from a numerical approach. To gain insight into the dependence of the optimal fitness response on $W(T)$ in the intermediate frequency regime, the optimal fitness responses for two environmental history functions, $T_1(t)$ and $T_2(t)$, that possessed the same density function, were determined. The optimal fitness responses with respect to the two different environmental history functions were compared to one another over all three frequency regimes.

The intuitive idea was to rearrange the order in which the temperatures occurred. To do so, the environmental density function of $T_1(t)$ was numerically computed. Recall that the environmental density function is a proxy for the amount of

time spent at each environmental condition in the range. Therefore, with the environmental range partitioned into 200 equal subintervals (the same discretization applied to numerically compute the optimal fitness response), Δt was computed for each ΔT, using the inverse of $T_1(t)$ (where it is defined). To obtain $T_2(t)$, the order in which the environments occurred were permuted, which could be done in any number of ways, as long as the Δt:ΔT ratio remained constant. It was decided that a $T_2(t)$ that was qualitatively different from $T_1(t)$, in terms of the number and location of extrema, would be appropriate to answer the questions of interest. The permutation for the $T_2(t)$ shown in Figure 2.4A is now defined.

$$F(t) = \begin{cases} G_1(t) := T_1\left(8t + \frac{\pi}{\omega}\right) & : 0 \leq t \leq t_1 \\ G_2(t) := T_1\left(4(t - t_1) + (8t_1 + \frac{\pi}{\omega})\right) & : t_1 \leq t \leq t_2 \\ G_3(t) := T_1\left(8(t - t_2) + 4(t_2 - t_1) + (8t_1 + \frac{\pi}{\omega})\right) & : t_2 \leq t \leq t_3 \\ G_3(-t + 2t_3) & : t_3 \leq t \leq t_4 \\ G_3(t - 2(t_3 - t_2)) & : t_4 \leq t \leq t_5 \\ G_3(-t + (t_3 + t_5)) & : t_5 \leq t \leq t_6 \\ G_2(-t + (t_2 + t_6)) & : t_6 \leq t \leq t_7 \\ G_1(-t + (t_1 + t_7)) & : t_7 \leq t \leq t_8 \\ G_1(t - t_8) & : t_8 \leq t \leq t_9 \\ G_1(-t + (t_1 + t_9)) & : t_9 \leq t \leq \pi/w, \end{cases} \tag{A.24}$$

where

$$t_1 = 2\left(\frac{\frac{2\pi}{\omega} - \frac{1}{\omega}\cos^{-1}\left(\frac{(59-\Delta T)-57}{3}\right) - \frac{\pi}{\omega}}{8}\right)$$

$$t_2 = 2\left(t_1 + \frac{\frac{2\pi}{\omega} - \frac{1}{\omega}\cos^{-1}\left(\frac{(61+\Delta T)-(59-\Delta T)}{3}\right) - 8t_1}{4}\right)$$

$$t_3 = 2\left(t_2 + \frac{\frac{2\pi}{\omega} - \frac{1}{\omega}\cos^{-1}\left(\frac{63-(61-\Delta T)}{3}\right) - (8t_1 + 4t_2)}{8}\right)$$

$$t_4 = t_3 + (t_3 - t_2)$$

$$t_5 = t_4 + (t_3 - t_2)$$

$$t_6 = t_5 + (t_3 - t_2)$$

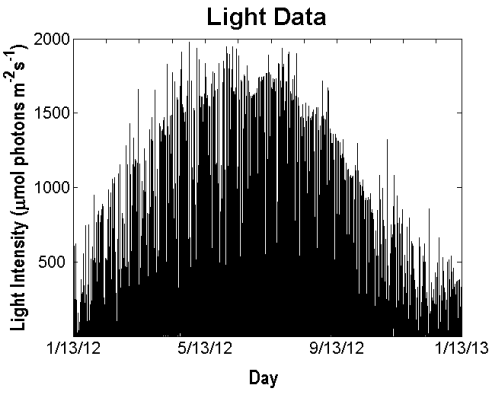
$$t_7 = t_6 + (t_2 - t_1)$$

$$t_8 = t_7 + t_1$$

$$t_9 = t_8 + t_1.$$

And then

$$T_2(t) = \begin{cases} F(t) & : 0 \leq t \leq \pi/\omega \\ F(t - \pi/w) & : \pi/\omega \leq t \leq 2\pi/\omega. \end{cases} \tag{A.25}$$



(A)

(B)

Figure A.6. (A) Light data collected from 13 January 2012 to 13 January 2013 at Mushroom Spring in Yellowstone National Park. (B) Light sensor (LI-COR) housed in a rock near an effluent channel at Mushroom Spring. Inset shows a close-up image of the light sensor.

APPENDIX B

CHAPTER 3 APPENDICES

Appendix B.1. Periodicity of One Species Solution

Recall from Chapter 2, equation 2.3, that the species solution was

$$x(t) = \frac{K_s Y}{r} \frac{e^{\int_0^t (\frac{rS^0}{K_s} f(T(B)) - D) dB}}{\int_0^t e^{\int_0^B (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} f(T(B)) dB + C}. \tag{B.1}$$

It is assumed here that the environmental history function $T(t)$ is periodic. Let

$$\begin{aligned} L(t) &= e^{\int_0^t (\frac{rS^0}{K_s} f(T(B)) - D) dB} \\ M(t) &= \int_0^t e^{\int_0^B (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} f(T(B)) dB, \end{aligned} \tag{B.2}$$

so that $x(t)$ can be rewritten as

$$x(t) = \frac{K_s Y}{r} \frac{L(t)}{M(t) + C}. \tag{B.3}$$

Note that the exponent of $L(t)$ can be rewritten as

$$\begin{aligned} \int_0^t \left(\frac{rS^0}{K_s} f(T(B)) - D \right) dB = &\int_0^Q \left(\frac{rS^0}{K_s} f(T(B)) - D \right) dB \\ &+ \int_Q^{2Q} \left(\frac{rS^0}{K_s} f(T(B)) - D \right) dB \\ &+ ... + \int_{Q(N-1)}^{NQ} \left(\frac{rS^0}{K_s} f(T(B)) - D \right) dB \\ &+ \int_{NQ}^t \left(\frac{rS^0}{K_s} f(T(B)) - D \right) dB \\ = N &\int_0^Q \left(\frac{rS^0}{K_s} f(T(B)) - D \right) dB \\ &+ \int_{NQ}^t \left(\frac{rS^0}{K_s} f(T(B)) - D \right) dB, \end{aligned} \tag{B.4}$$

where N is the number of periods elapsed before the last period, the one containing t, and Q is the duration of the period. The last equality holds because $f(T(t+Q)) = f(T(t))$ as the environmental history function $T(t)$ is assumed to be periodic.

Next, a similar procedure is applied to $M(t)$,

$$\begin{aligned}
M(t) &= \int_0^t e^{\int_0^B (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} f(T(B)) dB \\
&= \int_0^Q e^{\int_0^B (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} f(T(B)) dB \\
&\quad + \int_Q^{2Q} e^{\int_0^Q + \int_Q^B (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} f(T(B)) dB \\
&\quad + \int_{2Q}^{3Q} e^{2\int_0^Q + \int_{2Q}^B (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} f(T(B)) dB \\
&\quad + \qquad ... \\
&\quad + \int_{Q(N-1)}^{NQ} e^{(N-1)\int_0^Q + \int_{Q(N-1)}^B (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} f(T(B)) dB \\
&\quad + \int_{NQ}^{t} e^{N\int_0^Q + \int_{NQ}^B (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} f(T(B)) dB.
\end{aligned}$$

Note that the second term through the N-1 term can all be written as integrals from $[0, Q]$ and this yields

$$\begin{aligned}
M(t) = &\left[\sum_{i=1}^{N} e^{(i-1)\int_0^Q (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} \right] \int_0^Q e^{\int_0^B (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} f(T(B)) dB \\
&+ e^{N\int_0^Q (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} \int_{NQ}^{t} e^{\int_{NQ}^B (\frac{rS^0}{K_s} f(T(\gamma)) - D) d\gamma} f(T(B)) dB.
\end{aligned} \tag{B.5}$$

Next, let $C_p = \int_0^Q \left(\frac{rS^0}{K_s} f(T(\gamma)) - D \right) d\gamma$ and consider three cases.

Case 1: $C_p = 0$

When $C_p = 0$ the bracketed expression in (B.5) is equal to N and as $t \to \infty$, $N \to \infty$. This implies $M(t) \to \infty$ as $t \to \infty$ as well, since the multiplicative factor of the bracketed expression is positive ($f(T(t))$ is non-negative and not identically 0), and the second term in $M(t)$ is always non-negative. Also note that when $C_p = 0$, $L(t)$ will be bounded, so that it can be concluded that

$$\lim_{t \to \infty} x(t) = \lim_{t \to \infty} \frac{K_s Y}{r} \frac{L(t)}{M(t) + C} = 0. \tag{B.6}$$

To investigate the cases where $C_p < 0$ and $C_p > 0$, note that the bracketed expression in (B.5) is a finite geometric sum, that is

$$\begin{aligned}\sum_{i=1}^{N} e^{(i-1)C_p} &= \sum_{i=0}^{N-1} \left(e^{C_p}\right)^i \\ &= \frac{1 - e^{NC_p}}{1 - e^{C_p}}\end{aligned} \tag{B.7}$$

Case 2: $C_p < 0$

When $C_p < 0$ this implies $e^{C_p} < 1$, so in the limit as $N \to \infty$, $\sum_{i=1}^{N} e^{(i-1)C_p}$ will converge to a positive number. The factor multiplying the sum in $M(t)$ is always positive, and this time, the second term of $M(t)$ goes to 0, which in turn implies $M(t)$ converges to a positive, bounded number. Also note that $L(t) \to 0$ as $t \to 0$ in this case. Therefore

$$\lim_{t\to\infty} x(t) = \lim_{t\to\infty} \frac{K_s Y}{r} \frac{L(t)}{M(t) + C} = 0. \tag{B.8}$$

Case 3: $C_p > 0$

When $C_p > 0$ this implies $e^{C_p} > 1$, which in turn implies the geometric series diverges to ∞ as $N \to \infty$. In this case note that $L(t)$ also diverges to ∞, so L'Hospital's rule can be applied

$$\begin{aligned}\lim_{t\to\infty} x(t) &= \lim_{t\to\infty} \frac{K_s Y}{r} \frac{L(t)}{M(t) + C} = 0. \\ &= \lim_{t\to\infty} \frac{K_s Y}{r} \frac{\frac{rS^0}{K_s} f(T(t)) - D}{f(T(t))} \\ &= \lim_{t\to\infty} Y \left(S^0 - \frac{DK_s}{rf(T(t))} \right)\end{aligned} \tag{B.9}$$

Since $T(t)$ is assumed to be periodic, in all three cases $x(t)$ converges to the same thing to which a periodic function converges. q.e.d.

Appendix B.2. Asymptotics to Study Zero Eigenvalue

The rationale for including this appendix is to provide an alternative approach to the stability analysis investigated in the main text of this chapter. An alternative approach for the low frequency regime is required because of the inconclusive nature of that analysis – due to one of the Floquet exponents always being identically zero. Here, to study the long term behavior of the perturbed ODE, the following change of variables is introduced,

$$\tilde{x}_1 = \epsilon z_1$$

$$\tilde{x}_2 = \epsilon z_2.$$

Since the non-linear terms of (3.8) are quadratic, applying this change of variables to (3.8) and dividing both sides of the system by ϵ results in

$$\begin{pmatrix} \dot{z_1} \\ \dot{z_2} \end{pmatrix} = \mathbf{A}(t) \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \epsilon \mathbf{p} \left(t, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right). \tag{B.10}$$

This form suggests studying this system as a perturbation in ϵ. Let

$$\mathbf{z} = \mathbf{z_0} + \epsilon \mathbf{z_1} + \epsilon^2 \mathbf{z_2} + O(\epsilon^3), \tag{B.11}$$

where $\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ and $z_{i,j}$ represents the jth component of the ith-order solution. Substituting the 2nd-order approximation into (B.10) results in

$$\begin{aligned} \dot{\mathbf{z_0}} + \epsilon \dot{\mathbf{z_1}} + \epsilon^2 \dot{\mathbf{z_2}} =& A(t) \left(\mathbf{z_0} + \epsilon \mathbf{z_1} + \epsilon^2 \mathbf{z_2} \right) + \epsilon \left[-\frac{r}{K_s Y} \left((z_{0,1} + z_{0,2}) + \epsilon \left(z_{1,1} + z_{1,2} \right) \right. \right. \\ & \left. \left. + \epsilon^2 \left(z_{2,1} + z_{2,2} \right) \right) \begin{pmatrix} f_1 \left(z_{0,1} + \epsilon z_{1,1} + \epsilon^2 z_{2,1} \right) \\ f_2 \left(z_{0,2} + \epsilon z_{1,2} + \epsilon^2 z_{2,2} \right) \end{pmatrix} \right]. \end{aligned} \tag{B.12}$$

Equating like terms yields

$$\begin{aligned}
\epsilon^0 &: \dot{\mathbf{z_0}} = A(t)\mathbf{z_0} \\
\epsilon^1 &: \dot{\mathbf{z_1}} = A(t)\mathbf{z_1} - \frac{r}{K_s Y}\begin{pmatrix} f_1 z_{0,1}\left(z_{0,1}+z_{0,2}\right) \\ f_2 z_{0,2}\left(z_{0,1}+z_{0,2}\right)\end{pmatrix} \\
\epsilon^2 &: \dot{\mathbf{z_2}} = A(t)\mathbf{z_2} - \frac{r}{K_s Y}\begin{pmatrix} f_1\left(z_{1,1}\left(z_{0,1}+z_{0,2}\right)+z_{0,1}\left(z_{1,1}+z_{1,2}\right)\right) \\ f_2\left(z_{1,2}\left(z_{0,1}+z_{0,2}\right)+z_{0,2}\left(z_{1,1}+z_{1,2}\right)\right)\end{pmatrix}
\end{aligned} \tag{B.13}$$

Each equation is now linear and separable with respect to the state variables. Solving the 0th-order equation with $(\bar{x}_1, \bar{x}_2) = (y(t), 0)$, and initial conditions of $\begin{pmatrix} z_1(t_0) \\ z_2(t_0)\end{pmatrix} = \begin{pmatrix} C_1 \\ C_2\end{pmatrix} = \begin{pmatrix} C_{0,1} \\ C_{0,2}\end{pmatrix}$, yields

$$\begin{aligned}
z_{0,1} &= e^{\int_0^t\left[-D+\frac{r}{K_s}f_1\left(S^0-\frac{2y(B)}{Y}\right)\right]dB} \\
&\quad \cdot\left[-\frac{rC_{0,2}}{K_sY}\int_0^t\left(e^{-\frac{r}{K_s}\int_0^B\left[f_1\left(S^0-\frac{2y(\alpha)}{Y}\right)-f_2\left(S^0-\frac{y(\alpha)}{Y}\right)\right]d\alpha}f_1 y(B)\right)dB + C_{0,1}\right] \\
z_{0,2} &= C_{0,2}e^{\int_0^t\left[\frac{r}{K_s}f_2\left(S^0-\frac{y(B)}{Y}\right)-D\right]dB}
\end{aligned} \tag{B.14}$$

The solution to the 1st-order equation is

$$\begin{aligned}
z_{1,1} &= e^{\int_0^t\left[-D+\frac{r}{K_s}f_1\left(S^0-\frac{2y(B)}{Y}\right)\right]dB} \\
&\quad \cdot\left[\int_0^t e^{\int_0^B\left[D-\frac{r}{K_s}f_1\left(S^0-\frac{2y(B)}{Y}\right)\right]d\alpha}\right. \\
&\quad \left.\cdot\left[\left(-\frac{r}{K_sY}f_1 y(B) z_{1,2}\right)\left(-\frac{r}{K_sY}f_1\left(z_{0,1}\left(z_{0,1}+z_{0,2}\right)\right)\right)\right]dB\right] \\
z_{1,2} &= z_{0,2}\left[-\frac{r}{K_sY}\int_0^t\left(z_{0,1}+z_{0,2}\right)f_2 dB\right]
\end{aligned} \tag{B.15}$$

Thus far I have been able to show that $z_{0,1} \to -z_{0,2}$ as $t \to \infty$. However, higher order solutions are complicated and it has therefore been difficult to obtain any further information using this approach.

APPENDIX C

CHAPTER 4 APPENDIX

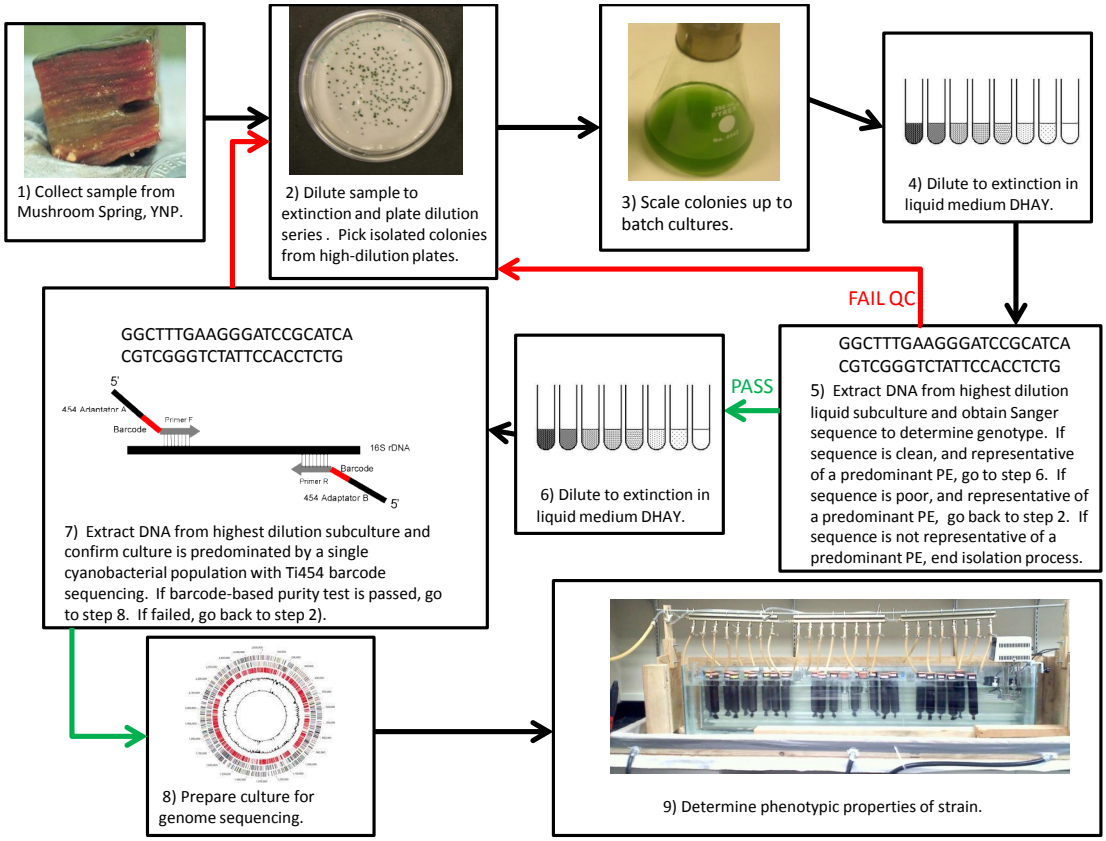


Figure C.1. Summary of protocol for obtaining *Synechococcus* isolates for phenotypic and genomic analyses.

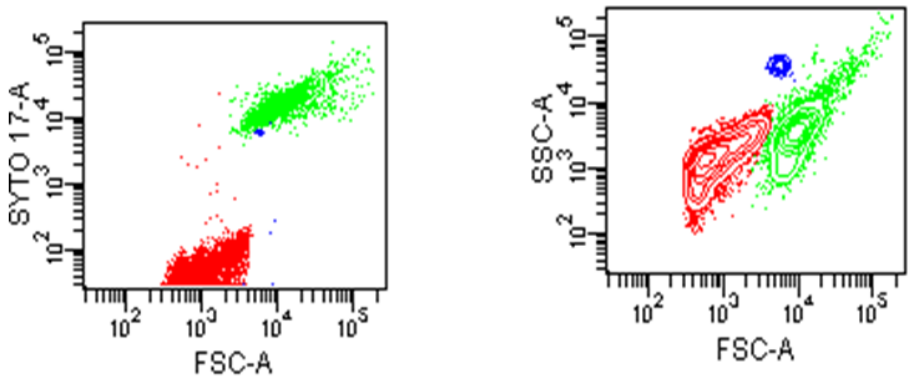
DNA extraction protocol for Ti454-barcode sequencing. Extraction was performed by first adding lysozyme at 200 μg/mL (final concentration) to the sample and incubating for 45 minutes at 37°C, mixing the sample every 15 minutes in a vortex mixer. Next, 110 μL SDS at (1% concentration) and 200 μg/ml proteinase K (final concentration) were added to the sample and this mixture was incubated for 50 minutes at 50°C, mixing the contents every 10 minutes. DNA was extracted by adding 950 μL of Tris-EDTA (pH 7.5) buffered phenol, mixing gently for 3 minutes, and then centrifuging for 5 minutes at 6600 x g. The aqueous layer was removed, then 450 μL phenol and 450 μL chloroform/isoamyl alcohol (24:1) were added to

the sample, which was mixed gently for 3 minutes, then centrifuged for 5 minutes at 6600 x g. The aqueous layer was removed and 900 μL chloroform/isoamyl alcohol (24:1) was added and the solution was mixed gently for 3 minutes, and then centrifuged for 5 minutes at 6600 x g. The aqueous layer was removed and 40 μL 3 M NaAcetate (pH 5.2) and 1 mL 70 % ethanol were added. The DNA was then allowed to precipitate overnight at -20°C. The sample was centrifuged at 4°C (6600 x g) for 30 minutes, washed in 0.9 mL 70% ethanol twice, and then dried using a SpeedVac (Thermo Scientific) for 10 minutes, at ambient temperature. The sample was resuspended with 200 μL Tris-EDTA buffer (pH 7.5) and the DNA was allowed to hydrate with no mixing for two hours.

Table C.1. Percentage of 16S rRNA Ti454-barcode variant closest relatives comprising the four *Synechococcus* isolates in this study and two of the 16S rRNA isolates from Allewalt et al. (2006).

Closest relative[a]	Percent in culture					
	65AY6Li (PE A1)	65AY6A5 (PE A4)	63AY4M2 (PE A6)	60AY4M2 (PE A14)	JA-3-3Ab (PE A1)	JA-2-3B′a (2-13) (PE B′11/B′24)
Synechococcus spp.	48.73	80.44	76.69	76.66	39.07	60.78
Meiothermus spp.	49.64	10.69	17.99	18.25	58.74	35.26
Chloroflexus spp.	0	1.97	0	0	0	0
Thermocrinis spp.	0	2.1	0	0	0	0
Caldilinea aerophila	1.22	2.62	2.6	3.01	0	0
Other	.41	2.2	2.72	2.09	2.19	3.96

[a] Results provided by The Research and Testing Laboratory (Lubbock, TX), who demarcate sequences based on identity scores, to well- characterized 16S rRNA sequences at the "species" level as those with greater than 97% identity (<3% divergence), and at the "genus" level as those with between 95% and 97% identity.



(A) (B)

Figure C.2. Flow cytometer (BD FACSAria II) output of 65AY6Li (PE A1) strain. (A) Scatter plot (left) showing relative fluorescence signal (SYTO 17-A) and size (FSC-A) of cells grown at 52°C and an irradiance of 25 μmol photons/m^2/sec in medium DHAY (without addition of any dissolved inorganic carbon) ∼36 hours after inoculation. *Synechococcus* cells (green) are typically five times longer than the heterotrophic contaminants that are able to pass through the 70 μm filter. *Synechococcus* cells also contain chlorophyll *a*, a pigment that is excited by the SYTO 17-A laser, that the heterotrophic cells do not possess. The blue population represents the fluorescent counting beads. (B) Contour plot showing cell size (FSC-A) versus cell complexity (SSC-A). A combination of these two plots was used to differentiate heterotrophic contaminants from *Synechococcus* cells.

Table C.2. Incubation times used to determine exponential growth phase under low-light conditions (scalar intensities of 25 and 125 μmol photons/m^2/sec). Experiment was done in duplicate and the times of exponential growth phase are provided for each replicate.

Light condition		65AY6Li (PE A1)	65AY6A5 (PE A4)	63AY4M2 (PE A6)	60AY4M2 (PE A14)
25 μmol photons/m^2/sec	replicate 1	36-96 h	72-144 h	60-108 h	84-144 h
	replicate 2	24-120 h	72-144 h	60-108 h	84-144 h
125 μmol photons m^2/sec	replicate 1	24-84 h	36-84 h	36-96 h	24-60 h
	replicate 2	24-84 h	24-72 h	24-84 h	24-60 h

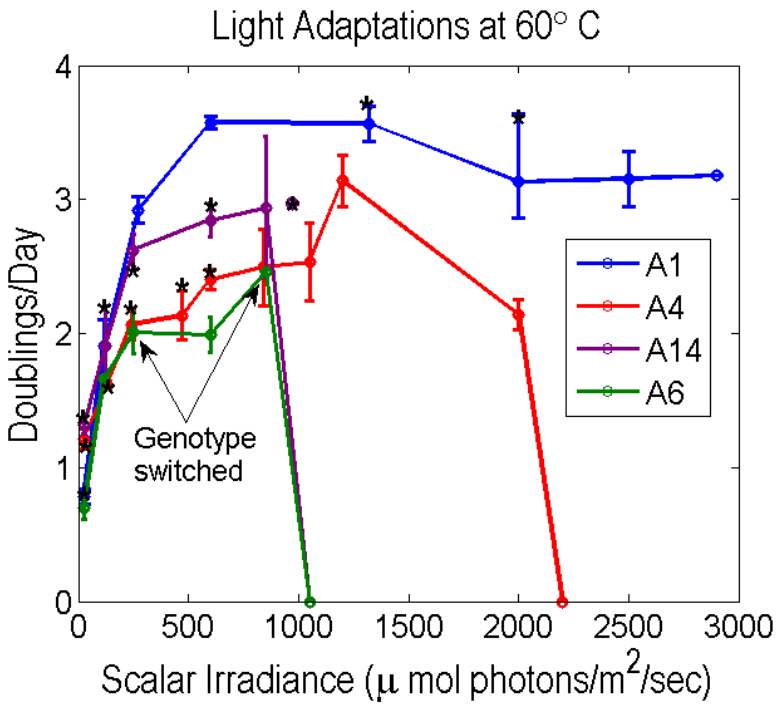


Figure C.3. Light adaptations of the PE A6 strain grown at 60°C and bubbled with 6% CO_2 in air. All growth curves are as in Figure 4.5A, except that the contaminated PE A6 strain is included here as well. The arrows point to the two light conditions where the post-experiment sequencing analysis revealed that the PE A6 genotype had switched to another predominant PE. Error bars are range bars and asterisks represent the post-experiment samples for which *psaA* genotypes were confirmed.

Table C.3. Summary of light responses of the contaminated PE A6 strain grown at 60°C and bubbled with 6% CO_2 in air.

Strain name	*psaA* PE/ 16S rRNA genotype	Temperature/ dilution[a]	Upper light limit[b]	Optimal light intensity[b]
63AY4M2	A6/A	$63°/10^{-5}$	1050	800

[a] Temperature of mat sample and dilution from which isolate originated.
[b] Units are scalar μmol photons/m^2/ sec.

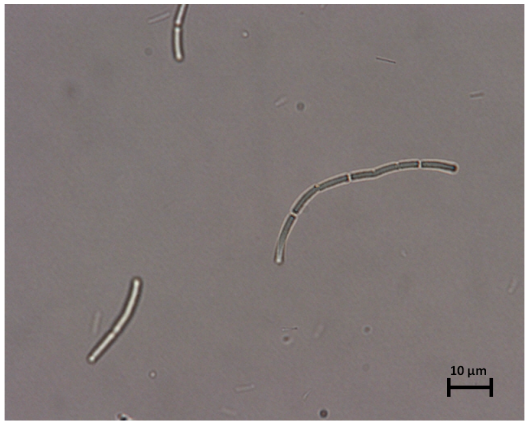


Figure C.4. Differential interference contrast photomicrograph of the PE A14 strain grown at a scalar irradiance of 600 μmol photons/m^2/sec. The image depicts the inability of daughter cells to separate from parent cells after division.