2011

MICROBIAL COMMUNITY ASSEMBLY AND DIVERSIFICATION OF THE GENUS CHLOROFLEXUS ALONG AN ALKALINE HOT SPRING GRADIENT

Michael Louis Weltzer
The University of Montana

MICROBIAL COMMUNITY ASSEMBLY AND DIVERSIFICATION OF THE GENUS CHLOROFLEXUS ALONG AN ALKALINE HOT SPRING GRADIENT

By

Michael Louis Weltzer

Bachelor of Science, Colorado State University, Fort Collins, Colorado, 2006

Thesis

presented in partial fulfillment of the requirements
for the degree of

Master of Science
in Microbiology

The University of Montana
Missoula, MT

May 2011

Approved by:

Sandy Ross, Associate Dean of the Graduate School
Graduate School

Dr. Scott Miller, Chair
Division of Biological Sciences

Dr. James Gannon
Division of Biological Sciences

Dr. Winsor Lowe
Division of Biological Sciences

Weltzer, Michael, M.S., May 2011 Integrative Microbiology & Biochemistry

MICROBIAL COMMUNITY ASSEMBLY AND DIVERSIFICATION OF THE GENUS CHLOROFLEXUS ALONG AN ALKALINE HOT SPRING GRADIENT

Chairperson: Dr. Scott Miller

ABSTRACT

This research aims to achieve a greater understanding of the structure of bacterial communities present in alkaline hot springs in Yellowstone National Park. I focus specifically on White Creek and Rabbit Creek in the Lower Geyser Basin. I show that, overall, the bacterial communities of both creeks are non-randomly assembled. However, at finer taxonomic scales, bacterial groups differ in their community assembly patterns. Specifically, phototrophic groups show the strongest evidence for non-random assembly, most likely due to competition for light. A major exception to this pattern is the genus *Chloroflexus*, a major member of these communities. Members of this genus primarily grow phototrophically, yet they did not show evidence of non-random assembly, as only one major 16S ribosomal RNA (rRNA) sequence was detected. Therefore, I next explored whether this single 16S rRNA sequence represents a single, broadly-distributed generalist or several cryptic specialist lineages. I isolated eleven strains of *Chloroflexus* from White Creek and determined that these isolates are members of a group without previously cultured representatives. I show that strains isolated from different temperatures have recently diverged within White Creek, as they can be differentiated genetically by the propionyl Co-A synthase gene, as well as phenotypically by differences in thermotolerance.

ACKNOWLEDGMENTS

I am grateful for having the privilege of working in the laboratory my advisor, Dr. Scott Miller. I am very appreciative of the enormous amount of help and encouragement he provided throughout my studies at the University of Montana. I also thank my other committee members, Dr. Jim Gannon and Dr. Winsor Lowe, for their inspiration and feedback. Thanks to Dr. Scott Samuels, Dr. Mark Grimes, and Dr. Erick Greene for mentoring me during my teaching assistantships. I am grateful for the encouragement and feedback from other graduate students in the Division of Biological Sciences, including Chris Wall, Margaret Hendrick, Angela Stathos, Blair DeBuysscher, Ellen Lark, Franny Gilman, J.T. Van Leuven, Tom Brekke, Brett Addis, Matt Wilson, and Jay Pollock. I would also like to thank Dr. Frank Rosenzweig, Dr. Bill Holben, Dr. Philip Ramsey, Dr. Lila Fishman, Dr. Andrea Sweigart, and Dr. Camille Barr. Finally, I would like to thank all of my family and friends for their support over the last three years.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 2: Divergence of a Novel Group of *Chloroflexus* within White Creek

LIST OF TABLES

Chapter 1: Non-Random Assembly of Bacterial Communities from Alkaline-Silica Hot Springs in Yellowstone National Park

Chapter 2: Divergence of a Novel Group of *Chloroflexus* within White Creek

Chapter 1

NON-RANDOM ASSEMBLY OF BACTERIAL COMMUNITIES FROM ALKALINE-SILICA HOT SPRINGS IN YELLOWSTONE NATIONAL PARK

Abstract

Understanding how communities are assembled is central to the field of ecology. There has been much debate regarding whether communities are random assemblages, or alternatively, whether there are rules that determine which combinations of taxa can co-occur. For microbial systems, in particular, the question of whether taxonomic groups exhibit differences in community organization remains unresolved, but is critical for our understanding of community structure and function. Here, I used three co-occurrence indices to evaluate assembly patterns for eight bacterial divisions distributed along two hot spring outflow channels in the Lower Geyser Basin of Yellowstone National Park, USA. Four divisions (Cyanobacteria, Chloroflexi, Acidobacteria and Cytophaga-Flavobacter-Bacteroides) exhibited less co-occurrence than expected by chance, with phototrophic divisions showing the strongest evidence for non-random community structure. I propose that both temperature-driven, niche-based processes and competition among organisms, respectively, contribute to the observed non-random community assembly. Non-random assembly may not have been detected in some groups because the ca. 100 bp V3 sequence of the 16S rRNA gene used to identify taxa was insufficient to resolve existing structure for certain closely related groups.

Introduction

One of the central aims of ecology is to understand how communities are assembled. The issue has a history of controversy (reviewed by Gotelli 1999), however, particularly with respect to the statistical problem of how to identify non-random structure for community survey data sets. Using appropriate null models, a recent meta-analysis of 96 co-occurrence matrices collected for plant and animal communities indicated that approximately one half of the communities exhibited non-random structure (Gotelli and McCabe 2002). The study also noted taxon-specific differences in community assembly. Specifically, plants and homeothermic animal communities were assembled non-randomly, while poikilotherms (with the exception of ants) were not. It is unknown if bacterial community assembly is correlated with physiology, as appears to be the case for macroorganisms. Recently, it was determined that all three domains have similar community assembly patterns (Horner-Devine et al. 2007). However, it remains unclear if assembly of microorganisms is similar to macroorganisms when analyzed at finer taxonomic levels. It is possible that the metabolic diversity of a bacterial division may influence its assembly pattern. For example, divisions containing primarily photoautotrophs, such as the Cyanobacteria, may be more likely to be non-randomly structured because all of the members of that group are competing for the same resources: light and electron donors (as is speculated to be the case for plants by Gotelli and McCabe 2002). Divisions with greater metabolic diversity, such as the Proteobacteria, may be expected to exhibit less structure.

Inferring the mechanisms responsible for community structure poses an additional challenge. Following his study of the distribution of finch species in the Bismarck

Archipelago, Diamond (1975) proposed that competition creates a non-random "checkerboard" pattern of species distribution in which some species pairs never co-exist. Others, however, have pointed out possible alternative explanations for non-randomly structured taxon distributions, including taxon sorting due to differences in habitat preferences (i.e., niche-based processes; Gotelli and McCabe 2002), historical factors (Vuilleumier and Simberloff 1980, Cracraft 1988, Gotelli and McCabe 2002), and neutral processes (Hubbell 2001, Bell 2005).

Here, I used three indices of co-occurrence (C-score, Stone and Roberts 1990; number of checkerboards, Diamond 1975; number of species combinations, Pielou and Pielou 1968) to test the hypothesis that bacterial divisions have similar assembly patterns as macrobial taxa. This hypothesis predicts that phototrophic taxa will show more structure than non-phototrophic taxa. To test this hypothesis, I analyzed the distribution patterns of the eight bacterial divisions present in microbial communities along two strong environmental temperature gradients (the outflow channels of two alkaline hot springs in the Lower Geyser Basin of Yellowstone National Park). Taxa within three of these divisions (Cyanobacteria, Chloroflexi and Acidobacteria) consist of phototrophs in these systems, whereas the remaining divisions are either principally or exclusively not phototrophic. Phototrophic taxa and the Cytophaga-Flavobacter-Bacteroidetes (CFB) exhibited less co-occurrence than expected by chance. I propose that both niche-based processes (specifically, divergence in thermotolerance among lineages) and competition among taxa contribute to these patterns. I also consider the possible explanations for the failure to reject the null hypothesis of random assembly for groups that did not show evidence of non-random assembly.

Methods

Data set

I analyzed the data set of Miller et al. (2009b), which used barcoded pyrosequencing of the V3 variable region of the bacterial 16S rRNA gene to determine the community composition of 39 samples collected from ten sites each along White Creek and Rabbit Creek, two alkaline hot springs located in the Lower Geyser Basin of Yellowstone National Park, USA. For the sample of 33,140 filtered sequences, 391 operational taxonomic units (OTUs) were identified based on unique V3 sequence signatures, resulting in a 391 × 39 abundance matrix for the full data set. Approximately 70% of the total sample was composed of either Cyanobacteria or Chloroflexi (i.e., green nonsulfur bacteria). Other abundant divisions included Cytophaga-Flavobacter-Bacteroidetes (CFB), Acidobacteria (specifically, *Chloracidobacterium*-like bacteria; Bryant et al. 2007), and candidate division OP10. Sequences from members of *Thermus*, Proteobacteria, Thermotogales, and Aquificales were also recovered. For each division, we organized the raw abundance data from Miller et al. (2009b) into presence-absence matrices for different taxonomic levels.

Analysis

C-score (Stone and Roberts 1990), number of checkerboard pairs (Diamond 1975), and number of species combinations (Pielou and Pielou 1968) were calculated for each matrix using Ecosim version 3.0.3.37 (Gotelli and Entsminger 2001). For each analysis, 5000 iterations were performed. Row and column sums were fixed, and

simulated matrices were created using the sequential swap algorithm. These settings were chosen because they have been shown to be less likely to result in Type I statistical error (Gotelli 2000, Gotelli and Entsminger 2001). The C-score is defined as the average number of checkerboard units between all pairs of species. A checkerboard unit consists of submatrix of two sites and two taxa, where each taxon is present at only one of the two sites, i.e., (10,01) or (01,10). The number of checkerboard units for each pair of taxa is calculated by the formula: $CU=(r_i-S)(r_j-S)$, where r_i and r_j represent the row totals for each taxon and S is the number of sites shared by the two taxa. If the C-score is higher than expected by chance, this indicates less co-occurrence than expected by chance. A smaller number of species combinations also indicates less co-occurrence, as does a greater number of checkerboards.

The standardized effect size (SES) was calculated for all analyses. The SES is the difference between the observed index and the simulated index normalized by the standard deviation of the simulated indices. It allows for the direct comparison of matrices with different numbers of taxa, since the value of the observed index is dependant on the number of taxa included in the analysis. Values close to zero indicate that the data are similar to the null model.

DNA isolation, amplification and sequence analysis

Environmental DNA was extracted with guanidium thhiocyanate following a procedure modified from Pitcher et al. (1989) and amplified with a forward primer specific for the 16S rRNA gene sequence of Candidate Division OP-10 (5'TCACGGCGGTATGGCTGACC3') and reverse primer 1492r (Lane 1991).

Amplification was performed under the following conditions: 1 minute initial denaturing at 94 °C followed by 40 cycles of 1 minute at 94°C , 1 minute at 54° C, 1.5 minutes at 72° C, and a final extension of 10 minutes at 72° C. Sequencing was performed at the University of Washington, Seattle, WA.

Results & Discussion

Full Data Set

The result of the analysis of the entire data set (analyzing all *Bacteria* together) was consistent with the results of previous analyses (Horner-Devine et al. 2007). The C-score for the full 391 OTU by 39 site incidence matrix was highly significant and exceeded the null model value (Table 1). This indicates that these communities are non-randomly structured with less taxon co-occurrence than predicted by the null model. Of the three co-occurrence analyses performed in this study, the C-score has been shown to have the most desirable statistical properties with respect to power and vulnerability to false positives (Gotelli 2000). The number of checkerboard pairs was greater than expected by chance (Table 1), but the number of taxon combinations was not fewer than the simulated null model. Taken together, these results indicated that the non-random assembly can be partially attributed to pairs of taxa that never co-occur (significant checkerboard number), but not to certain combinations of taxa that cannot co-exist (non-significant number of species combinations).

Evidence for the Non-random Structure of many Bacterial Divisions

Of the eight divisions evaluated, the Cyanobacteria, Chloroflexi, CFB, and the

Acidobacteria (i.e., *Chloracidobacterium*-like bacteria; Bryant et al. 2007) each had a significantly larger C-score (i.e., less co-occurrence) than predicted by the null model (Table 1). Of the divisions with statistically significant C-score results, only the Cyanobacteria matrix produced significant results for all three indices of co-occurrence. The three indices used have been shown to be only weakly correlated and therefore appear to measure different aspects of taxon co-occurrence (Gotelli 2000). Checkerboard scores and the number of taxon combinations are particularly sensitive to measurement errors in the presence-absence matrix that result in taxon rearrangement patterns, and are thereby more prone to Type II error.

As is typical for surveys of microbial diversity, the study of Miller et al. (2009b) undersampled the community, with only one of the thirty-nine samples beginning to plateau in a rarefaction analysis. One expected consequence of undersampling would be the difficulty of identifying the presence of rare taxa in a sample, with possible impact on the estimation of checkerboards and taxon combinations. In addition, Horner-Devine et al. (2007) demonstrated that undersampling decreases the statistical power of co-occurrence indices. Therefore, I would expect that further sampling of the community would only strengthen the statistical significance of our results, and I may be confident in the inferred non-random structure for the analyses for which the null hypothesis could be rejected.

Among the divisions with a non-significant C-score, Proteobacteria and candidate division OP-10 had a greater number of checkerboard pairs than predicted by the null model (Table 1). This result together with the non-significant C-score may be due to the rarity of many of the OTUs belonging to these divisions. Although 37 OTUs of

Proteobacteria were detected, 22 were present at only one or two sites, resulting in many checkerboard combinations. However, incidence matrices in which most taxa are absent at most sites result in low row totals, producing a C-score result that is not statistically significant. The matrix for OP-10 had a similar structure, with one OTU that was abundant at many sites, and several rare OTUs that were detected at few sites.

The results suggest differences in community assembly patterns among bacterial divisions. Specifically, phototrophic taxa were generally more structured than other taxa. This suggests that competition for light may be a major structuring force on phototrophic taxa across all domains. I expect light competition to have a greater effect on organisms within the same division than on organisms in different divisions, since members of different divisions contain different pigments that utilize different wavelengths of solar radiation. Divisions with high metabolic diversity, such as the Proteobacteria, were less structured. This may be due to less niche overlap in these divisions.

Are "randomly-assembled" divisions randomly assembled?

Though the divisions for which the null hypothesis was accepted for most or all indices (Proteobacteria, OP-10, Thermotogales, Thermus-Deinococcus) do not appear to exhibit the same degree of structure as the other groups, the question remains whether these divisions are truly random assemblages. One possible alternative explanation for the apparent random assembly of these bacteria is that the ca. 100 bp V3 sequence used by Miller et al. (2009b) was insufficient to resolve existing structure for certain closely-related taxa.

I investigated this possibility for two cases. The incidence matrix for candidate

division OP-10 consists of a predominant taxon (N = 2,241), which was present at all sample sites, and seven rare taxa. We amplified and sequenced an approximately 500 bp region of the 16S rRNA gene (surrounding the V3 region) for all OP-10 sequences from nine different field sites at White Creek and Rabbit Creek. My results revealed 10 polymorphic sites outside of the V3 region with relative frequencies that were associated with sample temperature (data not shown). This demonstrates that, in some cases, it is necessary to sequence a larger fragment of the gene in order to sufficiently capture diversity.

In the second case, I analyzed population genetic data available for the cyanobacterium *Fischerella* (*Mastigocladus*) *laminosus* at the White Creek site (Miller et al. 2009a). The population exhibits sequence identity not only for the V3 region (N = 1330 sequences sampled between ~39-55 °C) but also for the entire 16S rRNA gene (Miller et al. 2007, Miller et al. 2009a). However, the population does harbor low amounts of sequence variation at more rapidly evolving loci (Miller et al. 2009a), and recently-evolved variation in thermal performance exists among population members that is strongly associated with environmental temperature. Therefore, there is good evidence that the population is not assembled randomly along the White Creek channel. Using an incidence matrix based on sequences for four nitrogen metabolism genes (Miller et al. 2009a), I confirmed that *M. laminosus* genotypes are assembled non-randomly at White Creek (C-score SES 1.9; P = 0.056). Similar to *Mastigocladus*, other abundant and broadly distributed taxa also showed no evidence of non-random assembly, including the phototrophic genus *Chloroflexus* (*Chloroflexus* results presented above are for the division Chloroflexi). An examination of other loci is necessary to determine if non-

random structure is present in these groups as well (see Chapter 2).

This pattern can be contrasted with that for the cyanobacterial genus *Synechococcus*, a more ancient group that has also diverged in thermotolerance (Miller and Castenholz 2000). In this case, non-random assembly is clearly distinguished by the V3 sequence (C-score SES = 12.5; $P = 0$; Miller et al. 2009b). Together, these results emphasize that there are taxon-specific differences among community members in the evolutionary timescale of diversification. Detection of non-random assembly that is the product of more recent diversification may require analysis of loci other than the rRNA genes traditionally used as phylogenetic markers in surveys of microbial diversity.

Mechanisms of community assembly

Horner-Devine et al. (2007) identified two possible explanations for why microbial and macrobial communities both tend to exhibit less co-occurrence than expected by chance. One is that both types of communities are shaped by similar processes, the other that these communities share similar structure due to their interdependencies. Because the communities in the present study are composed entirely of microorganisms, my analyses suggest that microbial communities are structured by similar segregating mechanisms as macrobial communities. These could include competition (Diamond 1975), allopatric speciation (Vuilleumier and Simberloff 1980, Cracraft 1988, Gotelli and McCabe 2002), and differences in environmental tolerance (Gotelli and McCabe 2002).

It is unlikely that any single factor is entirely responsible for taxon segregation in these hot spring systems; however, differences among taxa in physiological tolerance

(i.e., niche-based processes) likely make an important contribution to community assembly through its influence on potential taxon combinations. In particular, environmental temperature has been shown to exert strong control on both community similarity and taxon richness at both White Creek and Rabbit Creek (Miller et al. 2009b). Thus, although microbes are typically capable of surviving over a 30 to 40 degree temperature range, the thermal gradients of these hot springs (~37 to 72 °C) presumably create habitat gradients that filters taxa based on relative fitness at different temperatures.

The non-random structure of these matrices, however, cannot be solely due to differences among taxa in temperature tolerance. Matrices constructed for all taxa at sites of similar temperature from the two hot springs revealed non-random structure in all cases (Table 2). Therefore, other mechanisms must also contribute to the structure of these communities. For instance, differences in water chemistry among sites may contribute to non-random co-occurrence patterns. Dissolved magnesium is below detection level in Rabbit Creek, but ranges from 3.7-8.6 μM in White Creek. Combined nitrogen and phosphorus is substantially higher in Rabbit Creek than in White Creek (Miller et al. 2009b).

Finally, competition likely plays a key role in structuring these communities. One group for which there is good evidence that this is the case is the cyanobacterial genus *Synechococcus*. The breadth of the realized thermal niches of the *Synechococcus* OTUs observed by Miller et al. (2009b) spanned discrete ~10 °C regions along the gradients of each hot spring. However, laboratory data indicate that *Synechocoocus* strains are typically capable of growing over a nearly 30°C range (Miller and Castenholz 2000, Allewalt et al. 2006). I interpret this disparity between potential and realized niche

breadths to be the result of competition at range boundaries.

Concluding Remarks

I have shown that some bacterial divisions in White Creek and Rabbit Creek are assembled non-randomly. However, my ability to detect non-random community structure in other groups may have been limited by the size of the gene segment used for taxon identification. Communities of phototrophic bacteria showed the highest degree of non-random structure. Non-random structure resulted not only from differences in environmental tolerance but, in some cases, also from competition for a limiting resource.

References

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

Bell, G. 2005. The co-distribution of species in relation to the neutral theory of community ecology. *Ecology* 86:1757-1770.

Bryant, D.A., Garcia Costas, A.M., Maresca, J.A., Chew, A.G.M., Klatt, C.G., Bateson, M.M., Tallon, L.J., Hostetler, J., Nelson, W.C., Heidelberg, J.F., Ward, D.M. 2007.

Candidatus Chloracidobacterium thermophilum: an aerobic phototrophic Acidobacterium. *Science* 317:523-526.

Cracraft, J. 1988. Deep-history biogeography: retrieving the historical pattern of evolving continental biotas. Systematic Zoology 37:221–236.

Diamond, J. M. 1975. Assembly of species communities. Pages 342-444 in M. L. Cody and J. M. Diamond, editors. Ecology and evolution of communities. Harvard University Press, Cambridge, Massachusetts, USA.

Gotelli, N. J. 1999. How do communities come together? *Science* 286:1684–1685.

Gotelli, N.J. 2000. Null model analysis of species co-occurrence patterns. *Ecology* 81:2606-2621.

Gotelli, N.J. and Entsminger, G.L. 2001. EcoSim: Null models software for ecology. Version 7.0. Acquired Intelligence Inc. & Kesey-Bear. http://homepages.together.net/~gentsmin/ecosim.htm.

Gotelli, N.J., McCabe, D.J. 2002. Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. *Ecology* 83:2091-2096.

Horner-Devine, M.C., Silver, J.M., Leibold, M.A., Bohannan, J.B., Colwell, R.K.,

Fuhrman, J.A., Green, J.L., Kuske, C.R., Martiny, J.B.H., Muyzer, G., Oveas, L., Reysenbach, A., Smilth, V.H. 2007. A comparison of taxon co-occurrence patterns for macro- and microorganisms. *Ecology* 88:1345-1353.

Hubbell, S.P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, NJ.

Lane, D.J. (1991) 16S/23S rRNA sequencing. In: Nucleic Acid Techniques in Bacterial Systematics (Stackebrandt, E. and Goodfellow, M., Eds.), pp. 115–147. John Wiley and Sons, Chichester.

Miller, S.R. and Castenholz, R.W. 2000. Evolution of thermotolerance in hot springs Cyanobacteria of the genus *Synechococcus*. *Applied and Environmental Microbiology* 66:4222-4229.

Miller, S. R., Castenholz, R.W. and Pedersen, D. 2007. Phylogeography of the thermophilic cyanobacterium *Mastigocladus laminosus*. *Applied and Environmental Microbiology* 73:4751–4759.

Miller, S.R., Williams, C., Strong, A.L., Carvey, D. 2009a. Ecological specialization in a spatially structured population of the thermophilic Cyanobacterium *Mastigocladus laminosus. Applied and Environmental Microbiology* 75:729-734.

Miller, S.R., Strong, A.L., Jones, K.L., Ungerer, M.C. 2009b. Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Applied and Environmental Microbiology* **75:** 4565-4572.

Pielou, D.P., and Pielou, E.C. 1968. Association among species of infrequent occurrence: the insect and spider fauna of *Polyporus betulinus* (Bulliard) Fries. *Journal of Theoretical Biology* 21:202-216.

Pitcher, D.G., Saudnders, N.A., Owen, R.J. 1989. Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Letters in Applied Microbiology* 8:151-156.

Stone, L., and Roberts, A. 1990. The checkerboard score and species distributions. *Oecologia* 85:74-79.

Vuilleumier, F., and Simberloff, D. 1980. Ecology versus history as determinants of patchy and insular distribution in high andean birds. Evolutionary Biology 12:235–379.

Group	N	C-score				Number of Checkerboards				Number of Taxon Combinations			
		Obs.	Mean (var)	P	SES	Obs.	Mean (var)	P	SES	Obs.	Mean (var.)	P	SES
Bacteria (full data set)	32,099	21.6	20.0 (.003)	0	28.7	13332	11595.74 (16921.28)	0	13.4	20	20 (0)	1	0
Acidobacteria	1796	78.4	65.5 (.723)	0	15.2	1	0.133 (.120)	0.13	2.5	13	18.6 (1.4)	0	-4.7
Cyanobacteria	10,877	19	16.5 (.038)	0	13	393	336.0 (175.6)	0	4.3	36	39.0 (.03)	0	-16.0
CFB	2,656	13.3	12.6 (.027)	0	4.3	242	183.1 (80)	0	6.5	31	30.9 (.96)	0.67	0.1
Chloroflexi	13,108	23.9	22.7 (.022)	0	7.8	303	276.2 (163.6)	0.03	2.1	39	39 (0)	1	0
Thermus-Deinococcus	611	32.2	32.3 (.583)	0.5	0.1	5	5.0 (1.6)	0.65	-0.02	18	20.2 (1.6)	0.07	-1.8
Thermotogales	231	7.0	7.5 (.204)	0.4	-1.1	0	0 (0)	1	0	7	6.7 (.2)	0.71	0.6
OP-10	2,322	28.4	27.8 (.567)	0.2	0.8	4	0.69 (0.42)	0	5.1	22	23.0 (1.9)	0.35	-0.7
Proteobacteria	498	7.2	7.0 (.018)	0.1	1.4	465	423.46 (203.7)	0	2.9	31	30.9 (.71)	0.7	0.1

Table 1 Results for C-score, number of checkerboards, and number of species combinations for the entire bacterial data set and each bacterial division detected in White Creek and Rabbit Creek in Yellowstone National Park. The observed (obs.) index is reported as well as the mean and variance (var.) of simulated indices.

Average Temperature (°C)	Field Sites	C-score	Mean	Variance	P	SES
38.5	WC1 (39 C), RC1 (38C)	0.5410	0.5110	0.0002	0	6.5
47	WC3 (47C), RC2 (47 C)	0.6294	0.5977	0.0003	0	6.0
54.75	WC 5 (54 C),WC 6 (57 C); RC 4 (53 C),WC 5 (55 C)	1.6185	1.5540	0.0008	0	7.2
61	WC7 (61 C), RC6 (61 C)	0.8278	0.7400	0.0002	0	19.1
67.67	WC9 (67 C); RC 9 (67 C),RC10 (69 C)	0.9929	0.9385	0.0001	0	4.7

Table 2 C-score results for every OTU detected at field sites in White Creek (WC) and Rabbit Creek (RC) with approximately the same temperature.

Chapter 2

DIVERGENCE OF A NOVEL GROUP OF *CHLOROFLEXUS* ALONG A HOT SPRING THERMAL GRADIENT

Abstract

This study explores diversity in the genus *Chloroflexus* along the White Creek thermal gradient in the Lower Geyser Basin of Yellowstone National Park. Unlike many other phototrophs, *Chloroflexus* in White Creek have a single dominant 16S rRNA sequence. This suggests either the presence of one generalist or cryptic diversification on a recent evolutionary time scale. Here, I use a combination of genetic and physiological approaches to show that novel *Chloroflexi* isolates have recently diverged in White Creek. Isolates from cooler temperatures showed thermal specialization, while an isolate from the warmest temperature had an increased thermal range but decreased maximal performance. Therefore, my data support the traditional view of a trade-off between ecological breadth and maximal performance. Due to the recent origin of *Chloroflexus* lineages, I reject the long-standing hypothesis of ancient co-adaptation between *Chloroflexi* and the cyanobacterium *Synechococcus* in alkaline hot spring systems.

Introduction

Understanding the factors which contribute to the origins and maintenance of ecological variation is a central goal of the investigation of microbial diversity. Spatially structured environments are predicted both to enhance the rate of diversification (Dobeli and Dieckmann 2003) and to maintain greater diversity through the co-existence of

multiple ecological specialists that are locally adapted to different niches (Kassen and Rainey 2004). The process of ecological specialization entails a fitness cost, or trade-off, in alternative environments. Most evolutionary theory assumes the existence of such trade-offs (Lynch and Gabriel 1987, Huey and Kingsolver 1993); evidence for trade-offs is often, but not always, found during adaptation of experimentally evolved laboratory microorganisms (Elena and Lenski 2003, Kassen and Rainey 2004, Bennett and Lenski 2007).

The contribution of trade-offs to niche diversification in natural communities of microorganisms is less clear. The environmental gradients of Yellowstone National Park alkaline hot springs are excellent systems for investigating this issue. In particular, trade-offs in thermotolerance appear to be an important mechanism of structuring diversity in populations of phototrophs in these systems. Members of phototrophic clades often exhibit substantial sequence variation in the slowly evolving 16S rRNA gene, suggesting that divergence of these groups is ancient. For example, divergent laboratory strains of the *Synechococcus* A/B group characteristic of these systems can vary in 16S rRNA gene sequence by over 5% (Miller and Castenholz 2000). Given the estimated 50-100 million years required for the 16S rRNA gene to diverge by 1% (Ochman and Wilson 1987), this suggests that the *Synechococcus* A/B clade began to diverge more than 100 million years ago. *In situ*, these genetically distinct lineages are restricted to different regions along thermal gradients (Ferris and Ward 1997, Miller et al. 2009b). Laboratory studies have determined that they represent specialists with divergent thermal performance curves (Miller and Castenholz 2000, Allewalt et al. 2006). Thermal niche specialization has also occurred in a group of phototrophic Acidobacteria (*Candidatus* Chloracidobacterium

thermophilum). In White Creek (Lower Geyser Basin), the four major 16S rRNA-defined Chloracidobacterium lineages present each occupy a relatively narrow realized thermal niche and peak in abundance at a different location along the gradient (Miller et al. 2009b).

Members of the genus *Chloroflexus* (division *Chloroflexi*) represent an apparent exception to this pattern of ancient specialization. This metabolically versatile group of anoxygenic phototrophs can account for as much as 50% of environmental sequences retrieved from these communities (Miller et al. 2009b). At White Creek, individual 16S rRNA-defined lineages of *Chloroflexus* are present in high abundance across a much broader range of temperatures than that observed for other phototrophic groups. Most notably, one lineage (*Chloroflexus* OTU10) was the most abundant *Chloroflexus* OTU across the entire gradient (39-68 °C). Its broad distribution raises the question of whether it represents a single generalist or, rather, a group of recently diverged, cryptic specialists that are indistinguishable by sequencing the 16S rRNA gene.

This study addresses whether *Chloroflexus* OTU10 in White Creek has diverged in thermotolerance. I first cultivated eleven strains of *Chloroflexus* OTU10. These are the first cultured members of this group of *Chloroflexi*, which is divergent from the previously cultured members of the genus, including laboratory model strains *C. aurantiacus* J-10-fl (Pierson and Castenholz 1974) and *C. aggregans* strain DSM9486 (Hanada et al. 1995). To test the hypothesis that recent divergence in thermotolerance has produced several specialists in this group, I have taken a combination of phylogenetic and physiological approaches. Genetic differentiation of strains was assessed with partial sequences for the 16S rRNA gene as well as for two loci diagnostic for the 3-

hydroxypropionate pathway of autotrophy used by these bacteria: malonyl Co-A reductase (*mcr*) and propionyl Co-A synthase (*pcs*). All White Creek strains had identical sequences for the 16S rRNA and *mcr* genes. However, strains could be differentiated by *pcs* gene sequence. Phenotypic differentiation in thermotolerance was assessed by measuring growth rates at seven temperatures. Together, these results show that strains of this novel group of *Chloroflexi* have diverged both genetically and phenotypically within White Creek.

Methods

Isolation

Microbial mat samples were collected from White Creek (WC) sites WC3 (Universal Transverse Mercator (UTM) coordinates: 516002E 4931137N, approximate temperature 47°), WC5 (UTM 516362E 4930907N, approximate temperature 54°) and WC7 (516471E 4930848N, approximate temperature 61° C) between May 2009 and May 2010. Environmental samples were spread on plates of D medium (Castenholz 1988) with additions of 0.00082g/L acetate, 0.0027g/L succinate, 0.0011g/L butyrate, 1.87 μL/L lactate, 0.00053g/L ammonium Cl at pH 8.2. Phytagel (0.7% wt/vol) was used as the gelling agent, rather than Bacto-agar, as this has been a successful strategy for improving the cultivation of novel microorganisms from soil (Janssen 2008). Spread plates of the samples collected from WC3 were incubated at 50° C, and spread plates of samples collected from White Creek sites WC5 and WC7 were incubated at 55° C. Plates were examined under a dissecting microscope several times per week. After one to two weeks of incubation, colonies that appeared to be *Chloroflexi* were transferred to

fresh plates. *Chloroflexi* are capable of gliding motility and will glide away from non-motile cells with which they are in co-culture (Hanada and Pierson 2006). To obtain axenic cultures, I used forceps to transfer the "leading edge" of the culture to a fresh plate. This process was repeated until an axenic culture was obtained. Cultures were maintained in liquid Castenholz D medium (pH 8.2) with additions 0.8g/L glycylglycine, 2.0 g/L yeast extract, 0.2 g/L ammonium chloride at 50° C for WC3 isolates, or 55°C for WC5 and WC7 isolates. All cultures were maintained under 24 hour irradiance at 150 μmol photons m-2 s-1 of tungsten radiation.

Phylogenetic analyses

I sequenced nearly the full-length of the 16S rRNA gene, using primers 23f and 1492r (Lane 1991) under the following cycling conditions: 1 minute initial denaturing at 94 °C followed by 40 cycles of 1 minute at 94°C , 1 minute at 54° C, 1.5 minutes at 72° C, and a final extension of 10 minutes at 72° C. Amplified products were directly sequenced with primer 23f at the University of Washington, Seattle, WA. *Chloroflexi* 16S rRNA gene sequences from the metagenomes of White Creek, Chocolate Pots, Fairy Spring, and Bath Lake Vista (at 2 time points) were obtained from the DOE-JGI Yellowstone Metagenome Project (Klatt et al. submitted, Inskeep et al. unpublished data). 16S rRNA sequence data for *C. aurantiacus* (strain DSM637, accession number J308501.1; strain J-10-fl, accession number CP000909.1), *C. aggregans* (strain DSM 9486, accession number AJ308499.1; sequence from Hanada, et al. 1995, accession number D32255.2), and *Chloroflexus* sp. 396-1 (accession number AJ308498.1) were obtained from the NCBI database. These sequences were aligned with 16S rRNA

sequences from White Creek isolates using CLUSTAL W (Thompson et al. 1994). Ambiguous base calls were excluded from the analysis. *Roseiflexus* sequences from the Yellowstone metagenome project were also included in the alignment as an outgroup. Maximum likelihood trees were generated by RaxML version 7.0.3 (Stamatakis 2006). Modeltest version 3.7 (Posada and Crandall 1998) selected the TIM + gamma + I model of DNA evolution, but the more general GTR + gamma + I model was used for this analysis, since the former model is not implemented by RaxML. Starting trees were generated randomly and 100 bootstrap replicates were performed using the rapid bootstrapping algorithm. MrBayes (Huelsenbeck 2001) version 3.1.2 was used to generate phylogenies by Bayesian inference. As with the likelihood analysis, the GTR + gamma + I model was used. Two replicates were run simultaneously for 10,000,000 generations until convergence was attained, as assessed by a convergence diagnostic below 0.01. Chain heating was adjusted from the program default of 0.2 to 0.025 to ensure sufficient chain swapping. The chains were sampled every 100 generations and the first 20% of sampled trees were discarded as burn in.

The *mcr* gene was amplified using forward primer 5'CATCTTTCCCGGCCCGATTG3' and reverse primer 5'CACAGGCAAATTCTAACCCTTC3'. The *pcs* gene from each strain was amplified using forward primer 5'AGAAGCGTAYACCGATCARG3' and reverse primer 5'CACCRACCACAATACAATTACC3'. Both sets of primers were designed from the Yellowstone National Park metagenome sequences. Both genes were amplified under the following cycling conditions: 1 minute initial denaturing at 94 °C followed by 40 cycles of 1 minute at 94°C , 1 minute at 54° C, 2 minutes at 72° C, and a final extension of 10

minutes at 72° C. Genes were sequenced using the forward primers at the University of Washington and the Murdock Lab, University of Montana, Missoula, MT. Sequences of both genes were aligned with CLUSTALW (Thompson et al. 1994). A sequence of each gene from the Bath Lake Vista metagenome (Klatt et al. submitted, Inskeep et al. unpublished data) was also included in the alignments. Identical sequence haplotypes were identified with DNAsp (Librado and Rozas 2009). A minimum spanning tree was inferred in Arlequin (Excoffier et al. 2005) and manually adjusted to produce a genealogical network for which the distance between any two sequences was identical to the number of observed nucleotide differences between them.

Physiological Tests

For growth rate experiments, strains were grown in liquid Castenholz D medium (pH 8.2) with additions of 0.8g/L glycylglycine, 2.0 g/L yeast extract, 0.2 g/L ammonium chloride. Test tubes were inoculated with enough stationary phase culture (generally 2mL) to obtain a starting OD of 0.01 to 0.02. Optical densities (OD) of 2mL sub-samples were determined by spectrophotometry at 660 nm every 24 hours until the cultures reached stationary phase (3-5 days). Growth rates were measured at 41°C, 45°C, 50°C, 55°C, 58°C, 60°C, 65°C, and 67°C (67°C for strain WC7-3 only) at a light intensity of 150 μmol photons m^{-2} s^{-1} of tungsten radiation. Triplicates of each strain were performed for each temperature treatment. However, only two replicates were used in the analysis of strains W3L at 50° and 55°, and WC5-1 at 55°, due to the failure of one of the replicates to begin growth. Growth rate was determined by calculating the slope of log of optical density at 660nm vs. time. This value was transformed to doublings per day by

dividing by the natural log of 2.

Results

Isolation

Spread plates of environmental samples resulted in the isolation of eleven axenic *Chloroflexus* strains (Table 1). All strains grew axenically on 0.01 D+ phytagel plates, but only six were capable of growth in liquid medium. Light microscopy was used to confirm that isolates had characteristic *Chloroflexus* morphology. An absorption spectrum showed that the isolates had typical pigmentation of *Chloroflexus*, including large amounts of Bacteriochlorophyll *c* (not shown).

Phylogenetic and Population Genetic Analysis

All strains were identical at the 680 nucleotides of the 16S rRNA gene sequenced. Phylogenies of the 16S rRNA gene were reconstructed to determine the evolutionary relationship of these isolates to *Chloroflexus* sequences in the metagenomes of other springs in Yellowstone National Park, as well as to cultured representatives of the genus. This analysis confirmed that all Yellowstone sequences are divergent from well-studied strains of *C. aurantiacus* and *C. aggregans* and belong to a group with no previously cultured representatives (Fig. 1). The White Creek sequences were most similar and nearly identical to (>99.5% nucleotide identity) to sequences from the Bath Lake Vista, Chocolate Pots, and Fairy Springs metagenomes.

To determine if isolates could be distinguished genetically, the malonyl Co-A reductase (*mcr*) and propionyl Co-A synthase (*pcs*) genes of the 3-hydroxypropionate

carbon fixation pathway were sequenced. All White Creek sequences were identical over the 700 nucleotides sequenced of the *mcr* gene, and differed from a sequence detected in Bath Lake Vista by less than 1% (1/151 nucleotides in the region of overlap). However, White Creek strains could be distinguished by *pcs* gene sequence. Three *pcs* alleles were present in White Creek. All were distinct from the allele detected in Bath Lake Vista. Of the 657 nucleotides sequenced, there were 40 polymorphic sites (6%), all of which were two-variant sites (each allele contained one of two different nucleotides). Eight of these sites were singletons, while the remaining 34 sites were shared by two strains. All WC3 and WC5 strains shared the same allele, two of the three WC7 strains (WC7-1, WC7-2) had a second allele and strain WC7-3 had a unique allele.

Highly similar gene sequences often violate the assumptions of traditional phylogenetics that bifurcating descendants are related by an extinct ancestor and that there has been no recombination between sequences (Posada and Crandall 2002). Therefore, to infer the relationships among the *pcs* alleles, a minimum-spanning genealogical network was reconstructed (Fig. 2a). This network is characterized by a loop connecting four alleles, three of which were not observed in my sample. Possible explanations for the loop include recombination or recurrent mutation. Since there were no three- or four-variant nucleotide sites in the alignment, the data conform to the infinite sites mutation model and thereby strongly implicate recombination. The presence of a single loop should be interpreted as a single recombination event in the sample. This was corroborated by Hudson and Kaplan's (1985) estimator of the minimum number of recombination events (Rm), which was equal to one. By comparing nucleotide sequences of polymorphic sites, the WC7-3 allele appears to be a recombinant of the other two

White Creek alleles (Figure 2b). This suggests recent or on-going gene flow among *Chloroflexus* lineages along the White Creek gradient.

Physiological Tests

To test for ecological diversification among strains, I experimentally determined the growth rate for six strains at seven different temperatures, ranging from 41 to 67° C. For these experiments I used only strains that grew in liquid medium. Strains isolated from different sites along the thermal gradient exhibited considerable phenotypic variation in both optimum and maximum growth temperatures as well as in maximal performance. Optimal temperature closely matched environmental conditions from which the strains were isolated. WC3 strains maximized growth rates at 50° C, while WC5 strains had maximal growth rates at 55 or 58° C (Fig. 3). The potential thermal niches of strains isolated from sites 3 and 5 had comparable breadths of thermotolerance. Two of the WC3 strains and one of the WC5 strains were capable of growth between 45° and 58° C. The remaining strain from each site had a slightly more limited potential niche and were capable of growth at temperatures between 45° and 55° C. Strain WC7-3 had a broader potential niche ranging from 45° to 65° C. However, it was never the fastest growing strain at any temperature, except for 60 and 65° C, temperatures at which no other strains were capable of growth.

Discussion

Chloroflexus strains have recently diverged

Chloroflexus isolates from White Creek have identical 16S rRNA sequences and

closely matched those from other Yellowstone hot springs, indicating that this group has diverged more recently than other phototrophs in these systems. Since the 16S rRNA gene sequences closely matched those from other Yellowstone hot springs and unique *pcs* alleles were observed at White Creek, this suggests that diversification of these *Chloroflexus* occurred within White Creek itself. Therefore, divergence of *Chloroflexus* is on a vastly different time scale than *Synechococcus* or Chloracidobacterium, which diverged prior to the origin of White Creek.

In several respects, the *Chloroflexus* community is similar to that of the cyanobacterium *Fischerella* (*Mastigocladus*) *laminosus*, which co-occurs with *Chloroflexus* in White Creek mats between temperatures of 38 and ~55° C (Miller et al. 2009b). Both populations are monomorphic at the 16S rRNA sequence across a broad range of temperatures but are genetically variable at more rapidly evolving loci. In both organisms, different genotypes are restricted to different locations on the thermal gradient, but there is also evidence for gene flow (Miller et al. 2009a). Recombination has played an important role in producing genetic variation in *Mastigocladus* (Miller et al. 2007), and future studies can address more comprehensively whether this is also the case for *Chloroflexus*.

Implications for the co-adaptation hypothesis

A long-standing hypothesis is that *Synechococcus* and *Chloroflexus* have co-evolved along a thermal gradient due to a tight producer-consumer relationship (Ward et al. 1987, Ward et al. 1998). However, this co-adaptation hypothesis has recently been questioned by pyrosequencing data showing that the abundance of the two groups are

negatively correlated in two Yellowstone hot springs (Miller et al. 2009b), which suggests a primarily competitive interaction. Furthermore, my data show that the two groups have diverged on different evolutionary timescales and, therefore, argue against their coordinated co-evolution along the thermal gradient. Any co-evolved ecological interactions between these groups would have to have developed relatively recently, following diversification of *Chloroflexus* along the thermal gradient.

Ecological variation and the nature of trade-offs in thermotolerance

Theories of the evolution of environmental tolerance (e.g., Lynch and Gabriel 1987) commonly assume that the area under organismal fitness curves remains constant during diversification. Implicit in this assumption is a "jack-of-all-trades is a master-of-none" trade-off, in which an extension in niche breadth comes at the cost of a reduction in maximal performance. Such a trade-off is also predicted by the Principle of Allocation (Levins 1968), which posits that energy limitations prevent an organism from excelling in all environments. However, there is often no correlation between maximal performance and thermal breadth (Carriere and Bovin 1997, Palaima and Spitze 2004). In fact, it is possible for generalists to maintain high maximal performance across a wide range of temperatures (that is, "a jack-of-all-trades" can be "master of all"; Huey and Hertz 1984). *Chloroflexus* isolates, however, do show evidence for a trade-off between increased maximal performance and decreased thermal range. The strain tested from site WC7 had a potential niche spanning 20° C, but never grew faster than 0.8 doublings per day. In contrast, strains isolated from lower temperatures were limited to a 10° to 13° potential niche, but were able to grow at rates of up to 1.2 doublings per day.

The "hotter is better" rule (Kingsolver and Huey 2008) has been proposed to explain why organisms with relatively high optimal temperatures often have relatively high maximal performance. The rule is based on the thermodynamic argument that reaction rates are higher at increased temperatures. Supporting evidence for this rule is found in diverse taxa including phages, bacteria, plants, and animals (Angilletta et al. 2010). To date, the only exception to the rule is hot springs strains of *Synechococcus*. Contrary to the predictions of the hotter is better rule, *Synechococcus* strains adapted to the highest temperature had the lowest maximal performance (Miller and Castenholz 2000). This group defines the upper temperature limit for photosynthesis for any organism. This suggests that performance does not monotonically increase with temperature and that exceptions to the hotter is better rule are likely to be found as taxa approach a fundamental evolutionary constraint. Here, I show that *Chloroflexus* strains are a second exception to this rule. The WC7-3 strain did not achieve a maximal growth rate comparable to that of strains isolated from lower temperatures. Further testing is needed to determine if other strains isolated from WC7 follow the same pattern.

Concluding remarks

I have isolated eleven axenic strains of a novel group of *Chloroflexus*. I have shown that, although the strains have identical 16S rRNA gene sequences, they can be distinguished genetically at the *pcs* locus and phenotypically by differences in thermal performance. These results suggest that future studies on the population genetics and thermal performance of these organisms is warranted. Finally, by showing that *Chloroflexus* and *Synechococcus* have diverged on different evolutionary timescales, I

have provided additional evidence against the longstanding co-adaptation hypothesis. Therefore, the ecological relationships between *Chloroflexus* and Cyanobacteria should be reassessed.

References

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

Angilletta, M.J., Huey, R.B., Frazier, M.R. 2010. Thermodynamic effects on organismal performance: is hotter better? *Physiological and Biochemical Zoology* 82:197-206.

Bennett, A.F. and Lenski, R.E. 2007. An experimental test of evolutionary trade-offs during temperature adaptation. *PNAS* 104:8649-8654.

Carriere, Y., and Bovin, G. 1997. Evolution of thermal sensitivity of parasitization capacity in egg parasitoids. *Evolution* 51:2028-2032.

Castenholz, R.W. 1988. Culturing methods for cyanobacteria. *Methods in Enzymology* 167: 68-93.

Dobeli, M. and Dieckmann, U. 2003. Speciation along environmental gradients. *Nature*

421: 259-264.

Elena, S.F., Lenski, R.E. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics* 4:457-469.

Excoffier, L. Laval, G., and Schneider, S. (2005) Arlequin ver. 2.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47-50.

Ferris, M.J. and Ward, D.M. 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.

Hanada, S., Hiraishi, A., Shimada, K., Matsuura, K. 1995. *Chloroflexus aggregans* sp. nov., a filamentous phototrophic bacterium which forms dense cell aggregates by gliding movement. *International Journal of Systematic Bacteriology* 45:676-681.

Hanada, S. And Pierson, B.K. 2006. The family Chloroflexaceae. *Prokaryotes* 7:815-842.

Hudson, R.R. and Kaplan, N.L. 1985. Statistical Properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* 111:147-164.

Huelsenbeck, J. P. and Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17:754-755.

Huey, R.B., and Hertz, P.E. 1984. Is a jack-of-all-temperatures a master of none. *Evolution* 38:441-444.

Huey, R. B., and Kingsolver, J.G. 1993. Evolution of resistance to high temperature in ectotherms. *American Naturalist* 142:S21–S46.

Janssen, P.H. 2008. New Cultivation Strategies for Terrestrial Microorganisms. In:Accessing Uncultivated Microorganisms. (Zengler, K. Ed.), pp.173-192. ASM press, Washington, DC.

Kasseen, R. and Rainey, P.B. 2004. The ecology and genetics of microbial diversity. *Annual Review of Microbiology* 58:207-231.

Kingsolver, J.G. and Huey, R.B. 2008. Size, temperature, and fitness: three rules. *Evolutionary Ecology Research* 10:251-268.

Knies, J. L., Kingsolver, J.G., Burch, C.L. 2009. Hotter is better and broader: thermal sensitivity of fitness in a population of bacteriophages. *The American Naturalist* 173: 419-430.

Lane, D.J. (1991) 16S/23S rRNA sequencing. In: Nucleic Acid Techniques in Bacterial Systematics (Stackebrandt, E. and Goodfellow, M., Eds.), pp. 115–147. John Wiley and Sons, Chichester.

Levins, R. 1968. Evolution in changing environments. Princeton University Press, Princeton, New Jersey, USA.

Librado, P. and Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451-1452 | doi: 10.1093/bioinformatics/btp187.

Lynch, M., Gabriel, W. 1987. Environmental tolerance. *The American Naturalist* 129: 283-303.

Miller, S.R. and Castenholz, R.W. 2000. Evolution of thermotolerance in hot springs Cyanobacteria of the genus *Synechococcus*. *Applied and Environmental Microbiology* 66:4222-4229.

Miller, S. R., Castenholz, R.W. and Pedersen, D. 2007. Phylogeography of the thermophilic cyanobacterium *Mastigocladus laminosus*. *Applied and Environmental Microbiology* 73:4751–4759.

Miller, S.R., Williams, C., Strong, A.L., Carvey, D. 2009a. Ecological specialization in a spatially structured population of the thermophilic Cyanobacterium *Mastigocladus* laminosus. *Applied and Environmental Microbiology* 75:729-734.

Miller, S.R., Strong, A.L., Jones, K.L., Ungerer, M.C. 2009b. Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Applied and Environmental Microbiology* 75:4565-4572.

Ochman, H. & Wilson, A. C. 1987. Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *Journal of Molecular Evolution* 26:74–86.

Palaima, A., and Spitze, K. 2004. Is a jack-of-all-temperatures a master of none? An experimental test with *Daphnia pulicaria* (Crustacea: Cladocera). *Evolutionary Ecology Research* 6:215-225.

Pierson, B.K. and Castenholz, R.W. A phototrophic gliding filamentous bacterium of hot springs, *Chloroflexus aurantiacus*, gen. and sp. nov. *Archives of Microbiology* 100:5-24.

Posada, D. and Crandall, K.A. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14:817-818.

Posada, D. and Crandall, K.A. 2002. The effect of recombination on the accuracy of

phylogeny estimation. *Journal of Molecular Evolution* 54:396-402.

Stamatakis, A. 2006. "RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models." *Bioinformatics* 22:2688–2690, 2006.

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

Ward, D. M., T. A. Tayne, K. L. Anderson, and M. M. Bateson. 1987. Community structure and interactions among community members in hot spring cyanobacterial mats. Symp. *Soc. Gen. Microbiol*. 41:179–210.

Ward, D. M., M. J. Ferris, S. C. Nold, and M. M. Bateson. 1998. A natural view of microbial diversity within hot spring cyanobacterial mat communities. *Microbiol. Mol. Biol. Rev.* 62:1353–1370.

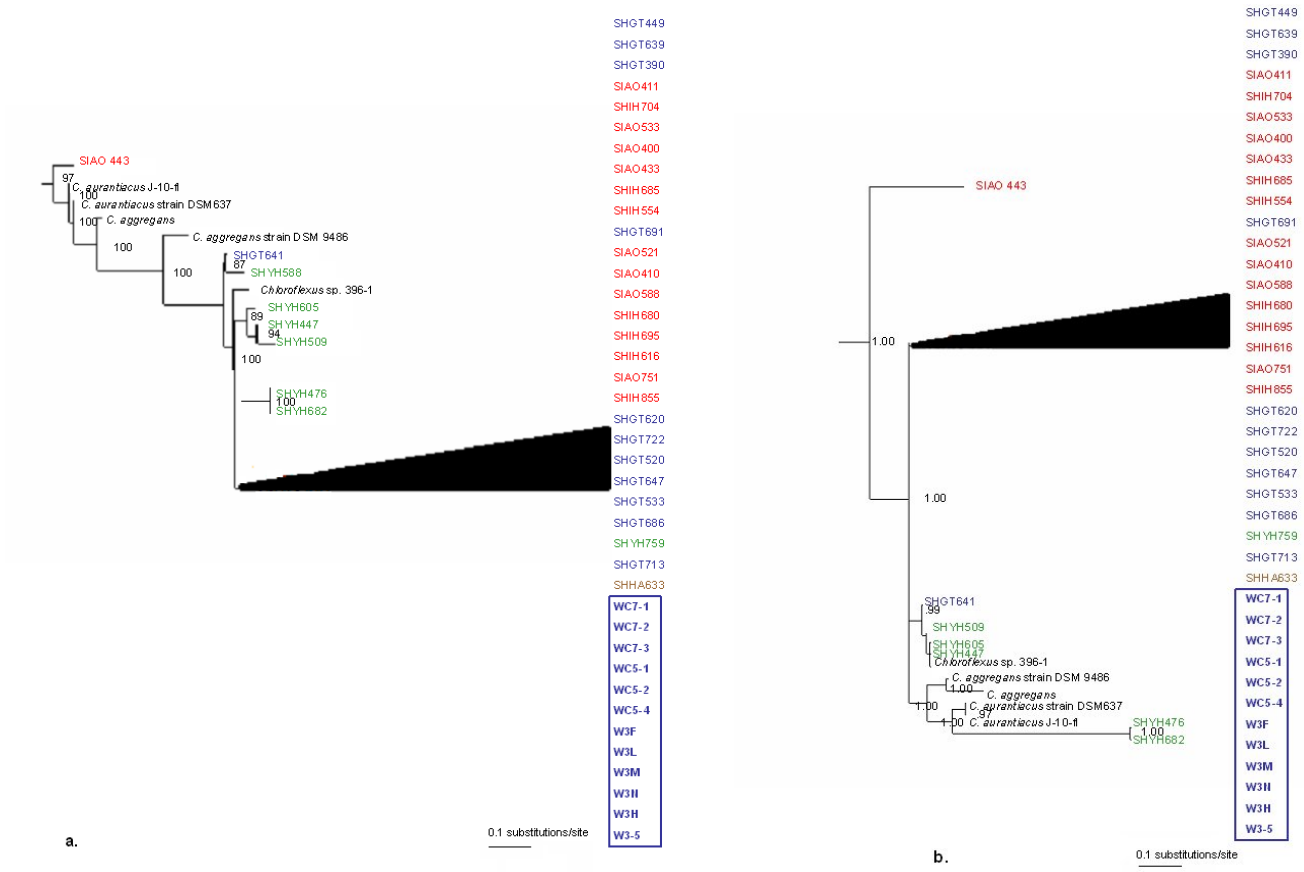


Figure 1. 16S rRNA gene phylogenies of *Chloroflexus* generated by (A) maximum likelihood analysis and (B) Bayseian inference. The phylogenies are outgroup-rooted with *Roseiflexus* sequences. Numbers at the nodes represent bootstrap values for 100 bootstrap replicates in (A) and posterior probabilities in (B). Sequences are color coded by origin: White Creek (blue), Bath Lake Vista Annex (red), Fairy Springs (green), and Chocolate Pots (brown). White Creek strains are in bold and boxed.

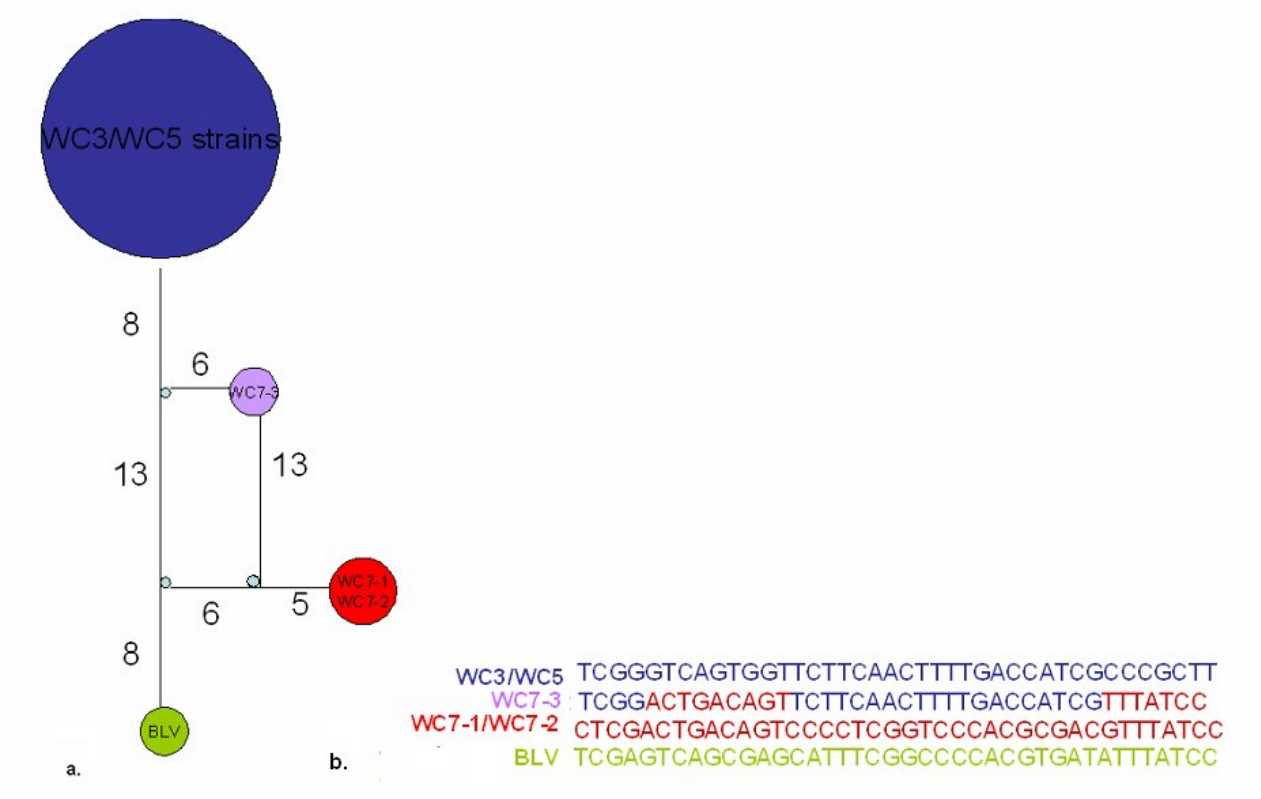
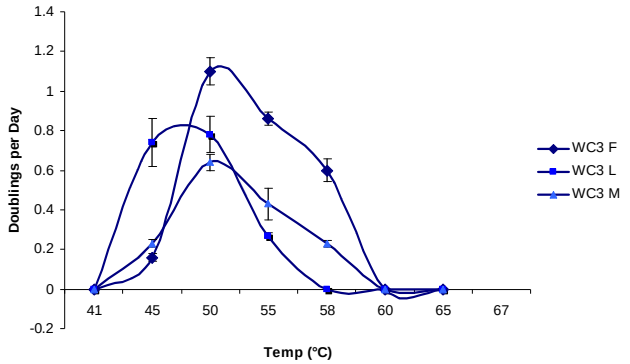
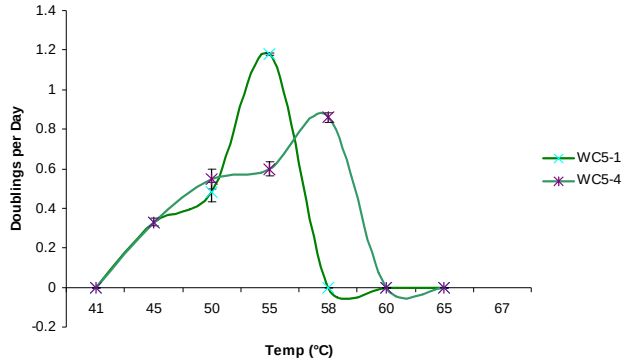


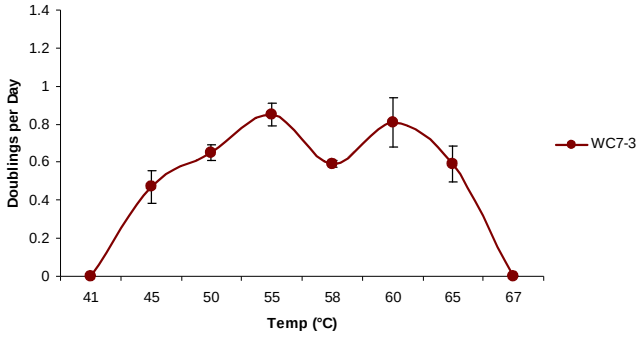
Figure 2 a. Genealogical network of *pcs* alleles. Allele 1 was detected in strains W3F, W3H, W3L, W3M, W3N, WC3-5, WC5-1, WC5-4; allele 2 was detected in strain WC7-3; allele 3 was detected in strains WC7-1 and WC7-2, allele 4 was detected in the Bath Lake Vista (BLV) metagenome. Numbers on branches indicate the number of mutations separating alleles. **b.** Sequence alignment of the 40 polymorphic nucleotide sites in the *pcs* sample. Allele 2 (strain WC7-3) appears to be a recombinant of alleles 1 and 3.



a.



b.



c.

Figure 3. Growth rates of *Chloroflexus* strains isolated from WC3 (**a**), WC5 (**b**), and WC7 (**c**). Growth rates were measured by optical density at 660 nm. Error bars represent standard errors for triplicate samples.

Stain Name	Collection Site	Collection Date
WC3F	White Creek 3	May 2009
WC3H	White Creek 3	May2009
WC3L	White Creek 3	July 2009
WC3M	White Creek 3	July 2009
WC3N	White Creek 3	July 2009
WC5-1	White Creek 5	May 2010
WC5-2	White Creek 5	May 2010
WC5-4	White Creek 5	May 2010
WC7-1	White Creek 7	May 2010
WC7-2	White Creek 7	May 2010
WC7-3	White Creek 7	May 2010

Table 1. Summary of collections from which *Chloroflexus* strains were isolated