Biodata of **Scott R. Miller**, author of the chapter ***"Diversity of the Cosmopolitan Thermophile Mastigocladus laminosus at Global, Regional and Local Scales"***

Scott R. Miller is currently an Assistant Professor in the Division of Biological Sciences at The University of Montana in Missoula, Montana, USA. He received his Ph.D. from the University of Oregon in 1999. His scientific interests include the evolution of environmental tolerance and the population genetics of microbial populations, with a principal focus on thermophilic cyanobacteria.

E-mail: **scott.miller@mso.umt.edu**

J. Seckbach (ed.), Algae and Cyanobacteria in Extreme Environments, 399–410.

DIVERSITY OF THE COSMOPOLITAN THERMOPHILE *MASTIGOCLADUS LAMINOSUS* AT GLOBAL, REGIONAL AND LOCAL SCALES

SCOTT R. MILLER
Division of Biological Sciences, The University of Montana, Missoula, MT 59812 USA

1. A Natural System for Investigating Microbial Biogeography

A recent theme in the study of microbial diversity has been the issue of whether and how the genetic and phenotypic variation of microorganisms is distributed along a geographic transect. Population genetic theory and the results of experimental evolution studies in the laboratory (e.g., Wright, 1931; Atwood et al., 1951; Bennett and Lenski, 1993) suggest that spatially structured microbial populations in nature should rapidly diverge from each other, provided that migratory gene flow among them is low, thereby creating geographic patterns of variation. Recent reports confirm that divergence of geographically isolated populations indeed outpaces the homogenization of genetic variation by migration, indicating the presence of dispersal barriers for microorganisms (e.g., Miller and Castenholz, 2000; Papke et al., 2003; Whitaker et al., 2003; Miller et al., 2006). These observations run counter to the longstanding idea that the abundance of a microorganism at a location is not limited by dispersal but is determined solely by environmental factors (Baas-Becking, 1934), a view that has been recently championed for eukaryotic microorganisms on the basis of morphological criteria (Finlay, 2002). Here, I will summarize our recent investigations of the biogeography of the moderately thermophilic, filamentous cyanobacterium, *Mastigocladus (Fischerella) laminosus*.

This bacterium provides an excellent system for investigating the organization of microbial diversity at a variety of geographic scales. *M. laminosus* is present virtually worldwide, though not necessarily in great abundance, in alkaline hot springs at temperatures below approximately 58ºC (Castenholz, 1996). For reasons that are not entirely clear, it tends to dominate fast-flowing thermal streams of appropriate temperature (Castenholz, 1978). Like other members of the Stigonematales, *M. laminosus* forms true branches as a result of changes in division plane during growth, typically resulting in a tuft of thick primary trichomes and narrow secondary trichomes. This complex morphology makes it easily recognizable during microscopic examination of field samples.

M. laminosus is also capable of differentiating specialized structures in response to changes in the environment. Under conditions of nitrogen limitation, it develops heterocysts for spatially separating the biochemically incompatible processes of nitrogen fixation and oxygenic photosynthesis. Other specialized

structures are hormogonia, motile dispersal structures released from the non-motile parental trichome by lysis of a necridial cell (Hernández-Muñiz and Stevens, 1987), and akinetes, which are freezing and desiccation-resistant resting cells that are produced in response to nutrient and/or light limitation.

The ability to produce akinetes may explain why *M. laminosus* is an excellent colonizer of new, often geographically distant hot spring habitat. A few examples serve as a testament to this bacterium's dispersal abilities. On the island of Surtsey, *M. laminosus* was observed around steam vents within several years of its formation by volcanic eruptions beginning in 1963 off the southern coast of Iceland (Castenholz, 1972). The nearest hot springs are 75–90 km away on the Icelandic mainland. It has also successfully colonized hot springs on Mount St. Helens formed following the 1980 eruption (Castenholz, 1996), as well as the former thermal effluent channels at Savannah River Nuclear Plant (Brock, 1978), where the closest site known to contain *M. laminosus* is Hot Springs, Arkansas. The above cases are notable for the absence of *Synechococcus cf. lividus*, a common resident of alkaline hot springs in western North America, which does not produce resting cells.

2. Global Diversity: A Cosmopolitan Bacterium with Local Flavor

To investigate the genetic diversity of *M. laminosus* from thermal areas around the world (including Yellowstone National Park, Iceland and New Zealand), we have characterized 51 of the laboratory isolates of this cyanobacterium that have been deposited in the University of Oregon's Culture Collection of Microorganisms from Extreme Environments. Each of the isolates belonged to one of seven closely related groups based on the complete identity of a partial sequence (835 bp) of the 16S ribosomal RNA (rRNA) gene (Fig. 1). The level of dissimilarity at this locus for the two most divergent lineages was under 3%, similar to that observed for A and B clades of *Synechococcus* (Miller and Castenholz, 2000). That is, global *M. laminosus* diversity is comparable to the degree of genetic differentiation between *Synechococcus* lineages that have diverged in thermal ecology and may be separated by only millimeters along a hot spring channel. Whereas a few of these 16S rRNA groups were endemic to a single location, many were found at multiple, geographically disparate sites (Fig. 2), highlighting this bacterium's dispersal capabilities.

A closer look at the genetic diversity of these isolates tells a more complicated story, however. We have also obtained sequence data from these strains for three functional genes involved in nitrogen metabolism: *nifH*, encoding the iron protein of nitrogenase; *narB*, the gene encoding assimilatory nitrate reductase; and *devH*, which codes for a DNA-binding protein required for the development of a functional heterocyst (Ramírez et al., 2005). These additional data revealed 23 distinct multi-locus haplotypes (Fig. 2). Although a few of these haplotypes were observed in geographically distant populations (e.g., haplotype 16 in Oman,

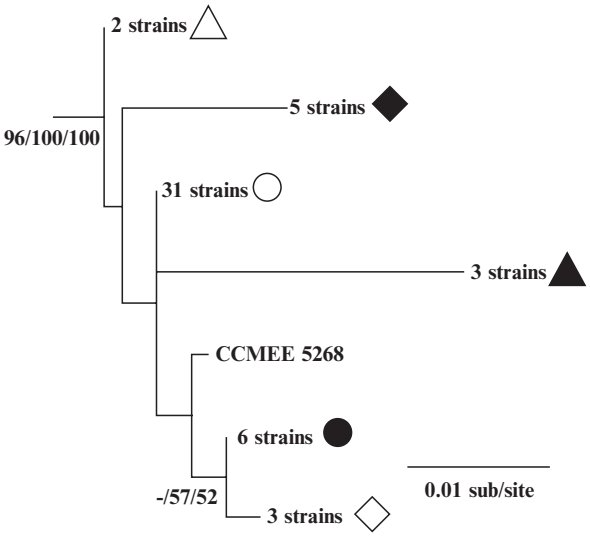

Figure 1. Maximum likelihood phylogeny of 51 *Mastigocladus* isolates reconstructed from 835 nucleotides of the 16S rRNA gene, rooted with outgroups *Chlorogloeopsis* PCC6718 and *Chroococcidiopsis* PCC7203 (not shown). The model of DNA substitution used (HKY + G + I) was selected by a hierarchical likelihood ratio analysis implemented in ModelTest 3.06. Symbols for each lineage appear in Fig. 2. Bootstrap values are indicated for likelihood, neighbor joining, and parsimony analyses for 1,000/10,000/10,000 pseudoreplicates.

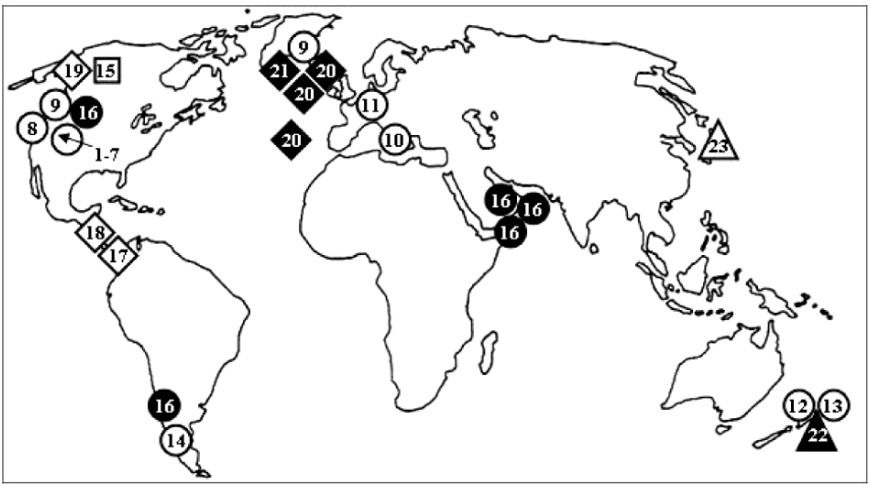

Figure 2. Genetic diversity of *M. laminosus* on a global scale. The symbols represent the seven 16S rRNA groups depicted in Fig. 1. Numerals refer to individual haplotypes observed in the survey based on approximately 3 kbp of sequence data obtained for three functional genes involved in nitrogen metabolism (see text for details).

Montana, and Chile), most haplotypes were only observed at a single location. For example, within Yellowstone National Park, haplotypes 1–4 are unique to White Creek in the Lower Geyser Basin, haplotype 5 to Boiling River, haplotype 6 to Chocolate Pots and haplotype 7 to Obsidian Pool. This suggests that genetic differentiation between populations generally occurs on very local geographic scales, despite the greater dispersal capability of *M. laminosus* compared with other thermophilic cyanobacteria such as *Synechococcus*.

Why this local differentiation in an apparently strong disperser? Factors in addition to dispersal capability appear to be required to explain the existence of both the wide geographic distribution of some haplotypes and the genetic differentiation among local populations of *M. laminosus*. While dispersal probability is the primary factor determining colonization of new suitable habitat, the probability of a migrant successfully establishing itself at a populated site is expected to depend not only on its ability to disperse to the location, but also on other factors, including its relative fitness compared with the extant population. For example, an established population could act as a preemptive barrier to migration simply by occupying substrate.

Because geographic barriers to migration appear to be important in this system, we would expect there to be a positive association between the amount of genetic divergence between a pair of strains and the geographic distance that separates their sites of origin. We investigated whether there is indeed evidence for isolation by distance in *M. laminosus*. To do so, we used correlation analysis to examine the relationship between nucleotide diversity (the probability of two sequences having a different nucleotide at a site) and the distance separating each pair of sequences for members of the most abundant 16S rRNA group in the dataset (open circles in Fig. 1). There was a strong and statistically significant positive relationship between these variables (unpublished data), providing further evidence for the importance of dispersal barriers in shaping the observed geographic patterns of genetic diversity in this group.

The major phylogenetic groups identified in the CCMEE survey exhibit interesting differences in thermal performance, particularly with respect to the maximum temperature permissive of growth. Whereas the thermal limit of strains from certain 16S rRNA lineages (open triangle and open circle in Fig. 1) was observed to be 55ºC in batch culture experiments, others were capable of growth at 57ºC (strain CCMEE 5268, closed circle and open diamond in Fig. 1). There was no obvious relationship between the temperature of the collection from which the strain had been isolated and its thermal maximum. The ecological importance of these physiological differences therefore requires further investigation. Because the phylogeny of these 16S rRNA lineages is currently unresolved (Fig. 1), it is also not possible at the present time to infer with confidence the direction that phenotypic evolution has taken during *M. laminosus* diversification (e.g., whether strains with a higher thermal limit were derived from less thermotolerant ancestors).

3. Regional Diversity: Life Under Different Nutrient Regimes

As noted above, populations of *M. laminosus* within geographic regions such as the greater Yellowstone area are genetically differentiated from each other. This raises the question of whether populations are locally adapted to the prevailing chemical and physical conditions of their environment, which is a fundamental question in the study of microbial diversity. We have investigated this issue for two populations of *M. laminosus* within Yellowstone National Park that exhibit dramatic phenotypic differences in situ as a result of environmental differences in nitrogen availability. White Creek is a thermal discharge fed stream in the Lower Geyser Basin that is dominated in biomass by *M. laminosus* where mean annual temperature is between 39 and 54ºC (unpublished data). Combined nitrogen is undetectable in this system, and heterocyst-forming *M. laminosus* fixes nitrogen at appreciable rates as assayed by acetylene reduction (~150 nmol ethylene produced per μg Chl *a* per hour; Miller et al., 2006). In contrast, Boiling River, a short (~150 m) channel of Mammoth Hot Springs outflow located approximately 50 km from White Creek, is nitrate-replete (130 μg/L). *M. laminosus* from this site does not develop heterocysts, and ethylene production in acetylene reduction assays did not differ from background levels (Miller et al., 2006).

We have recently investigated the degree of genetic differentiation between these populations in considerable detail. Despite their close proximity, *M. laminosus* from White Creek and Boiling River are genetically distinct. No genotype was shared between populations, based on the distribution pattern of 25 polymorphic nucleotide sites from approximately 8 kbp of sequence data collected for six nitrogen metabolism genes (Miller et al., 2006). The average amount of divergence observed between a randomly chosen pair of sequences from the total sample was comparable to that observed for the human global population at noncoding autosomal loci (Yu et al., 2004), indicating that identity of these Yellowstone strains of *M. laminosus* at the 16S rRNA gene and the internal transcribed spacer of the *rrn* operon belie substantial genetic diversity. This diversity is simply at a different scale of evolutionary differentiation than is typically investigated in microbial ecology. Another way to quantify the amount of genetic differentiation between populations is with F_{ST}. This parameter takes on values between 0 (when different populations harbor the same alleles at a locus in the same proportions) and 1 (when the populations are fixed for different alleles). Observed values of F_{ST} at polymorphic loci were considerable, and ranged between 0.22 and 0.94 for the six nitrogen metabolism genes. For comparison, F_{ST} for two Yellowstone populations of the hyperthermophilic archaeon *Sulfolobus* was estimated to be 0.37 (Whitaker et al., 2003).

But are these genetically differentiated populations adapted for the utilization of the different nitrogen sources available in their respective environments? Given sufficient evolutionary time, we would expect that the relative fitness of each population will have increased on its available nitrogen source in response to

selection, and, conversely, for the ability to assimilate an unutilized nitrogen source (either because it is not available in the environment or because it is not preferred) to have declined due to relaxed selective constraints on loci involved in its metabolism. As a first step toward testing whether local adaptation has occurred in the Boiling River and White Creek populations, we assayed growth with either nitrate or dinitrogen as sole nitrogen source for ten randomly selected strains from each population (Miller et al., 2006). No differences in performance were observed between the populations when grown on either kind of medium: Growth rate constants μ (h^{-1}) for each population were approximately 0.02 when grown in the presence of nitrate, and about 0.017 under nitrogen-fixing conditions (Miller et al., 2006).

Why haven't these populations apparently adapted to the prevailing nutrient conditions of their respective environments? After all, they appear to spend generation after generation expressing idiosyncratic suites of nitrogen metabolism genes, while repressing others. Population genetic models suggest that local adaptation is mutation-limited, but also that the populations are expected to continue to diverge due to low migratory gene flow (Miller et al., 2006). This suggests that locally adaptive mutations may ultimately become fixed in their respective environments given sufficient time.

4. Local Diversity: Population Structure Along a Thermal Gradient

Whereas the changes in microbial community composition along certain environmental gradients are comparatively well understood (e.g., see Ward et al. (1998) for a review on community shifts along hot spring gradients), how *population*-level microbial diversity is partitioned along environmental gradients is still unknown, in large part due to the predominant use of slowly evolving markers such as the 16S rRNA gene. Yet there are several compelling reasons for exploring this largely hidden component of microbial diversity. Of central importance is the recognition that genetic differences between individuals within populations provide the raw material for *all* evolutionary diversification, whether that variation is generated by new mutations or by the acquisition of novel DNA from a donor. Consequently, it is at the population scale that one can investigate the *process* of diversification that produces the resultant patterns that we describe when we investigate diversity at the macroevolutionary (phylogenetic) scale. Second, genetic differences within populations may prove to be important for community function, particularly if different population members specialize on different regions along an environmental gradient. A population-genetic perspective on microbial diversity would also enable the empirical evaluation of predictions derived from experimental evolution studies of microbial populations in the laboratory.

One of the intrinsic challenges of investigating this issue with microorganisms is the difficulty of unambiguously identifying a distinct population that is distributed along a clear environmental gradient. *M. laminosus* from White Creek

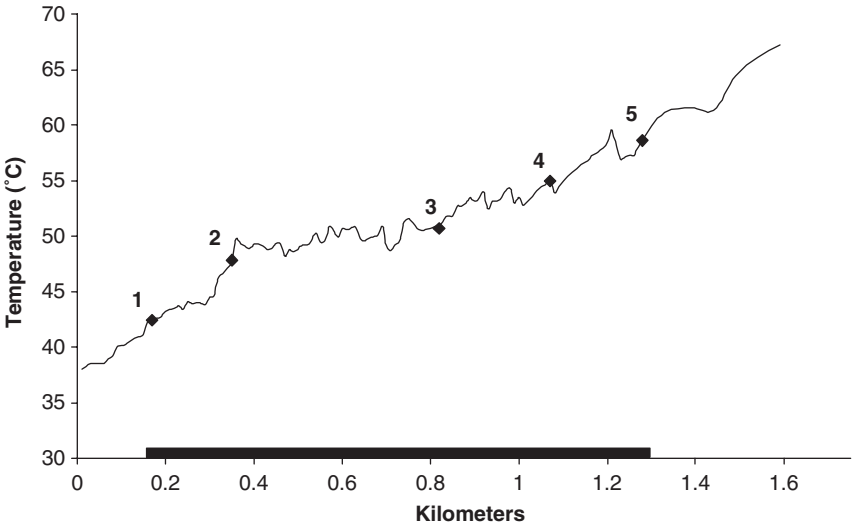

Figure 3. White Creek thermal gradient on September 11, 2004, with the region occupied by *M. laminosus* indicated by the shaded bar. Collection sites for population structure analysis are also shown (diamonds).

in the Lower Geyser Basin of Yellowstone National Park provides just such a system. Benthic streamers of this bacterium dominate the biomass where the mean annual temperature is between approximately 39 and 54ºC, a region of the channel that stretches for greater than a kilometer (Fig. 3). For perspective, this 15ºC difference in mean temperature is comparable to that observed for a 15º latitudinal gradient in eastern US Atlantic coastal waters.

To test whether analogous differences in allele frequencies and thermal performance exist among *M. laminosus* that occupy different regions along the White Creek thermal gradient, we first made multiple live collections from five sites spanning the range of *M. laminosus* habitat (Fig. 3). Individual trichomes were directly isolated in laboratory culture from these collections to provide a representative sampling of in situ diversity that is not biased by selective enrichment. Clonal isolates were next characterized both genetically and phenotypically. Though these bacteria share identical nucleotide sequences at slowly evolving loci including the SSU rRNA gene and the downstream internal transcribed spacer, *M. laminosus* from White Creek had been found in a previous study to be variable at several nitrogen metabolism loci (Miller et al., 2006). We genotyped approximately 150 isolates at four of these loci (*narB*, *devH*, *nifH*, and *nirA* (nitrite reductase). A total of 14 multi-locus haplotypes were observed in our sample (Table 1). These were not distributed evenly across sites, indicating the existence of population structure along the White Creek gradient. Most haplotypes were found

Table 1. *M. Laminosus* Haplotype Frequencies Along the White Creek Thermal Gradient.

	Frequency at site (%)				
Haplotype[a]	**1**	**2**	**3**	**4**	**5**
2232	8.3	8.3	50.0	29.3	27.9
2122	41.7	16.7	–	–	–
2132	–	–	8.8	2.4	30.2
2121	–	16.7	–	–	–
2231	–	–	–	–	11.6
3131	8.3	8.3	23.5	19.5	2.3
3121	–	8.3	5.9	9.8	7.0
3122	16.7	33.3	–	–	–
3232	–	–	–	17.1	16.3
3331	–	–	–	–	2.3
3321	–	8.3	–	–	–
3421	–	–	11.8	22.0	–
4323	16.7	–	–	–	2.3
4233	8.3	–	–	–	–

[a]Four-digit haplotype designation indicates allelic identities at *narB, devH, nifH,* and *nirA.*

at two or fewer sites, and only two were observed at all sites. Some haplotypes were most abundant at gradient extremes, e.g., 2122 and 2132 are the most frequently observed haplotypes at site 1 and site 5, respectively. An analysis of molecular variance (AMOVA) indicated that roughly 13% of the total variation in the sample was due to differences in genetic composition among sites (unpublished data), which is comparable to the amount of molecular variation observed among geographic regions for the human global population (Excoffier et al., 1992).

The observed nonrandom associations between haplotype frequencies and environmental temperature may be the result of adaptation along a selection gradient. Preliminary investigations suggest that some haplotypes have indeed diverged with respect to thermal performance. For example, haplotype 2122 (only found at lower temperature sites 1 and 2) and haplotype 2132 (abundant only at site 5 and absent from sites 1 and 2) exhibit dramatic differences in thermal dependence of growth despite sharing identical alleles at three of the four loci examined. Whereas 2122 strains outperform 2132 at 37ºC (the approximate temperature of site 1), haplotype 2132 strains double ~2 times faster than 2122 strains at 55ºC (the approximate temperature of site 5; unpublished data). These results suggest that the population structure observed along the White Creek thermal gradient may in part reflect differences in thermal ecology among genetically divergent population members, and that at least some haplotypes have specialized on the local conditions of their respective environments.

Temperature dependence of dry weight-normalized phycobiliprotein content is qualitatively similar to that of growth, although the difference between haplotypes in the amount of these light-harvesting proteins is most striking at

37ºC (unpublished data). At this temperature, haplotype-specific differences in phycobiliprotein content explain a large fraction (86%) of the variation in fitness between 2122 and 2132. By contrast, differences in phycobiliprotein content at 55ºC explained little (7%) of the differences in fitness. These differences in the relationship between light-harvesting pigment composition and fitness at 37 and 55ºC, respectively, suggest that the genetic basis of adaptation to temperature extremes along the White Creek gradient differs. At the former temperature, 2132 strains may be experiencing either chronic nitrogen deprivation and/or light stress. Production of phycobiliprotein-containing phycobilisome complexes requires a heavy investment in protein, and, consequently, degradation of these supramolecular light-harvesting antennae is an early response to perceived nitrogen limitation and light stress in cyanobacteria (e.g., Collier and Grossman, 1992). Elucidating the genetic basis of observed differences in thermal performance of these haplotypes at gradient extremes will be an exciting challenge!

5. Summary

Recent studies of the distribution of genetic diversity in distant hot springs have demonstrated the existence of geographic structure for the cyanobacteria *Synechococcus* and *Oscillatoria cf. amphigranulata* (Papke et al., 2003) and for the acidophile *Sulfolobus solfataricus* (Whitaker et al., 2003). Here, I have described examples of genetic differentiation at distant as well as much finer geographic scales for the cosmopolitan cyanobacterium *M. laminosus*.

With accumulating evidence for microbial biogeographical structure from extreme environments as well as other habitats (reviewed by Hughes et al., 2006), the next step is to better understand how these patterns are generated by evolutionary and demographic processes. For example, gene networks (Posada and Crandall, 2001) can provide insight into the relative ages of alleles at a locus, thereby potentially enabling inference of the centers of dispersal from which younger alleles have spread by migration. Gene networks reconstructed from the patterns of global sequence variation observed for *M. laminosus* at nitrogen metabolism loci suggest that alleles sampled from western North America are older than, and in some cases gave rise to, alleles found in other parts of the world. This suggests that the evolutionary origins of much of the extant *Mastigocladus* diversity on the planet might be traced back to geothermal activity associated with the hot spot that currently lies under Yellowstone National Park. Enhancing our understanding of the evolutionary origins of microbial diversity will also require the linking of genetic and phenotypic variation on a geographic scale. Our evidence for differences in thermotolerance of *M. laminosus* strains at different phylogenetic scales, from between 16S rRNA groups to within a population distributed along an environmental gradient, highlights the likelihood that genetically divergent microorganisms will often exhibit phenotypic differences of potential ecological importance.

6. References

Atwood, K. C., Schneider, L. K. and Ryan, F. J. (1951). Periodic selection in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **37:** 146–155.

Baas-Becking, L. G. M. (1934). *Geologie of Inleiding Tot de Mileau-Kunde*. W. P. van Stokum, The Hague, The Netherlands.

Bennett, A. F. and Lenski, R. E. (1993). Evolutionary adaptation to temperature. II. Thermal niches of experimental lines of *Escherichia coli*. Evolution **47:** 1–12.

Brock, T. D. (1978). *Thermophilic microorganisms and life at high temperatures*, Springer-Verlag, New York.

Castenholz, R. W. (1972). The occurrence of the thermophilic blue-green alga, *Mastigocladus laminosus*, on Surtsey in 1970. The Surtsey Progress Report **VI:** 14–19.

Castenholz, R. W. (1978). The biogeography of hot spring algae through enrichment cultures. Mitt. Int. Verein. Limnol. **21:** 296–315.

Castenholz, R. W. (1996). Endemism and biodiversity of thermophilic cyanobacteria. Nova Hedwigia Beih. **112:** 33–47.

Collier, J. L. and Grossman, A. R. (1992). Chlorosis induced by nutrient deprivation in *Synechococcus* sp. strain PCC 7942: not all bleaching is the same. J. Bacteriol. **174:** 4718–4726.

Excoffier, L., Smouse, P. E. and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131:** 479–491.

Hughes Martiny, J. B., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. Claire, Kane, M., Adams Krumins, J., Kuske, C. R., Morin, P. J., Naeem S., Øvreås, L., Reysenbach, A.-L., Smith, V. H. and Staley, J. T. (2006). Microbial biogeography: putting microorganisms on the map. Nature Rev. Microbiol. **4:** 102–112.

Finlay, B. (2002). Global dispersal of free-living microbial species. Science **296:** 1061–1063.

Hernández-Muñiz, W. and Stevens, S. E. Jr. (1987). Characterization of the motile hormogonia of *Mastigocladus laminosus*. J. Bacteriol. **169**: 218–223.

Miller, S. R. and Castenholz, R. W. (2000). Evolution of thermotolerance in hot spring cyanobacteria of the genus *Synechococcus*. Appl. Environ. Microbiol. **66**: 4222–4229.

Miller, S. R., Purugganan, M. D. and Curtis, S. E. (2006). Molecular population genetics and phenotypic diversification of two populations of the thermophilic cyanobacterium *Mastigocladus laminosus*. Appl. Environ. Microbiol. **72:** 2793–2800.

Papke, R. T., Ramsing, N. B., Bateson, M. M. and Ward, D. M. (2003). Geographic isolation in thermophilic cyanobacteria. Environ. Microbiol. **5:** 650–659.

Posada, D. and Crandall, K. A. (2001). Intraspecific gene genealogies: trees grafting into networks. Trends Ecol. Evol. **16:** 37–45.

Ramírez, M. E., Hebbar, P. B., Zhou, R., Wolk, C. P. and Curtis, S. E. (2005). *Anabaena* sp. strain PCC 7120 gene *devH* is required for synthesis of the heterocyst glycolipid layer. J. Bacteriol. **187:** 2326–2331.

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

Whitaker, R. J., Grogan, D. W. and Taylor, J. W. (2003). Geographic barriers isolate endemic populations of hyperthermophilic Archaea. Science **301:** 976–978.

Wright, S. (1931). Evolution in Mendelian populations. Genetics **16:** 97–159.

Yu, N., Jensen-Seaman, M. I., Chemnick, L., Ryder, O. and Li, W.-H. (2004). Nucleotide diversity in gorillas. Genetics **166:** 1375–1383.