Biogeography of thermophilic cyanobacteria and the importance of isolation to the evolution of microorganisms
by Robertson Thane Papke

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy m Microbiology
Montana State University

Abstract:
Evolutionary theory predicts the divergence of populations when they become geographically isolated. However, Baas Becking's theory that "everything is everywhere and the environment selects" excludes geographic isolation for microorganisms. In previous diversity and distribution studies, the sequencing of 16S rRNA genes acquired from natural Synechococcus populations residing in hot spring mats from Yellowstone National Park revealed that a single morphology concealed a rich 16S rRNA genotypic diversity. Predominating within that diversity is a group of closely related 16S rRNA genotypes (the A/B cluster) that are uniquely distributed along thermal and light gradients. Curiously, the upper temperature limit for cyanobacterial mat formation is different in globally disparate sites suggesting barriers to dispersal for some populations. I hypothesized that either members of the A/B cluster are distributed globally, but the highest temperature adapted forms (A types) are limited in their dispersal capabilities, or alternatively, globally disparate hot springs are dominated by unrelated Synechococcus genotypes. To test these hypotheses, I performed phylogenetic analysis on PCR-amplified, cloned, 16S rDNA genes recovered from Synechococcus populations residing in hot spring mats in Italy, New Zealand, Japan and the northwest U.S.A. The abundance of detected lineages was determined using lineage-specific oligonuleotide probes; low-abundance genotypes were sought using the same probes as PCR primers. I also assessed 20 different hot spring physical/chemical properties to determine whether adaptation was important to the local and global distributions of Synechococcus populations. Results revealed that: (1) A/B cluster 16S rDNA sequences were not detected outside of the U.S., (2) each country had unique dominating Synechococcus genotypes, (3) within the U.S. and Japan there exist local geographic clades for A/B and Cl lineages, respectively, at the 16S rRNA and internal transcribed spacer region loci, (4) Oscillatoria amphigranulata, a filamentous thermophilic cyanobacterial species also demonstrated unique geographical distributions, and (5) genetic variation did not correlate with tested hot spring physical/chemical parameters. The results revealed that all cyanobacterial lineages had a different dispersal capability, but even the most widely dispersed exhibited substantial evidence of geographic isolation. Additional evidence for isolated prokaryotic populations is reviewed and the general importance of isolation in microbial evolution is emphasized.

BIOGEOGRAPHY OF THERMOPHILIC CYANOBACTERIA AND THE IMPORTANCE OF ISOLATION TO THE EVOLUTION OF MICROORGANISMS

by

Robertson Thane Papke

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

of

Doctor of Philosophy

in

Microbiology

MONTANA STATE UNIVERSITY
Bozeman, Montana

February 2002

APPROVAL

of a thesis submitted by

Robertson Thane Papke

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Dr. David M Ward _____ (Signature) 26 February 2002 Date

Approved for the Department of Microbiology

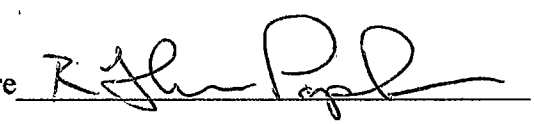
Dr. Cliff Bond _____ (Signature) 26 February 2002 Date

Approved for the College of Graduate Studies

Dr. Bruce McLeod _____ (Signature) 2-27-02 Date

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University-Bozeman, I agree that the Library shall make it available to borrowers under the rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted " the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Signature 

Date 02/27/02

ACKNOWLEDGEMENTS

I would like to thank Dr. David Ward for guiding me through the long process of earning a Ph.D. degree and for teaching me the importance of raising the bar of my own expectations and efforts. Thank you Mary Bateson, your kindness and friendship has made my journey through Bozeman a very happy one. Thank you Kenji Kato for sharing your house, hospitality and friendship and for your supreme efforts in helping me arrange my entire my Japanese hot spring collections. I thank the members of my committee, especially Adam Richman, for their time and effort in making this a better thesis. And thanks to all of the students and fellow scientists at MSU who have enriched my life both personally and scientifically with special thanks to Greg Colores, Mike Franklin, Myke Ferris, Uli Nübel and Marcel van der Meer. I could not have done this without any of you and I am eternally grateful to each and every one of you.

This research was funded by grants from the National Science Foundation, NASA, the Summer Institute in Japan and the Thermal Biology Institute.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Evolutionary theory predicts the divergence of populations when they become geographically isolated. However, Baas Becking's theory that "everything is everywhere and the environment selects" excludes geographic isolation for microorganisms. In previous diversity and distribution studies, the sequencing of 16S rRNA genes acquired from natural *Synechococcus* populations residing in hot spring mats from Yellowstone National Park revealed that a single morphology concealed a rich 16S rRNA genotypic diversity. Predominating within that diversity is a group of closely related 16S rRNA genotypes (the A/B cluster) that are uniquely distributed along thermal and light gradients. Curiously, the upper temperature limit for cyanobacterial mat formation is different in globally disparate sites suggesting barriers to dispersal for some populations. I hypothesized that either members of the A/B cluster are distributed globally, but the highest temperature adapted forms (A types) are limited in their dispersal capabilities, or alternatively, globally disparate hot springs are dominated by unrelated *Synechococcus* genotypes. To test these hypotheses, I performed phylogenetic analysis on PCR-amplified, cloned, 16S rDNA genes recovered from *Synechococcus* populations residing in hot spring mats in Italy, New Zealand, Japan and the northwest U.S.A. The abundance of detected lineages was determined using lineage-specific oligonuleotide probes; low-abundance genotypes were sought using the same probes as PCR primers. I also assessed 20 different hot spring physical/chemical properties to determine whether adaptation was important to the local and global distributions of *Synechococcus* populations. Results revealed that: (1) A/B cluster 16S rDNA sequences were not detected outside of the U.S., (2) each country had unique dominating *Synechococcus* genotypes, (3) within the U.S. and Japan there exist local geographic clades for A/B and C1 lineages, respectively, at the 16S rRNA and internal transcribed spacer region loci, (4) *Oscillatoria amphigranulata*, a filamentous thermophilic cyanobacterial species also demonstrated unique geographical distributions, and (5) genetic variation did not correlate with tested hot spring physical/chemical parameters. The results revealed that all cyanobacterial lineages had a different dispersal capability, but even the most widely dispersed exhibited substantial evidence of geographic isolation. Additional evidence for isolated prokaryotic populations is reviewed and the general importance of isolation in microbial evolution is emphasized.

CHAPTER 1

INTRODUCTION

Development of Evolutionary Theory in Microbiology

Great inroads toward comprehending evolution and the formation of species were made after naturalists and scientists visited locations around the globe, collected plants, animals and fossils and charted the organisms' relatedness against local and/or global distributions and ecological gradients. The independent formation by Darwin and Wallace of the theory of descent with modification via natural selection was completely dependent upon their observations that different yet related species lived in different regions of the world or on separate islands within archipelagos. As biologists searched for and catalogued the diversity of organisms on Earth, the disciplines of biogeography, and more recently phylogeography revealed many more corresponding patterns of organismal relatedness with geography[1]. As a mechanism for speciation geographic isolation is fundamentally different from natural selection, since population differences are driven by neutral genetic drift, not adaptation. Rosenzweig[2] expressed the importance of geographic isolation to the development of species when he articulated that "geographical speciation is the most common mode among most taxa in most places at most times." Indeed, the familiar terms used to describe speciation events, allopatric, parapatric and sympatric speciation all refer to the relative distances (distant, near or

together, respectively) that separate two sister species. Today, it is recognized that populations diverge whenever any kind of barriers to mating success are formed (e.g. different habitats, differential mating periods [day, season or year], anatomical incompatibility, different mating rituals, hybrid death or sterility). In the time since Darwin and Wallace published their great contributions to the science of biology, much has been learned about organismal diversity and mechanisms for speciation.

Unfortunately, evolutionary theory did not have a major impact on the field of microbiology. In 1963, nearly 300 years after van Leeuwenhoek first discovered microorganisms in his microscope and more than 100 years after Darwin and Wallace published, Stanier et al.[3] concluded that, "...any systematic attempt to construct a detailed scheme of natural relationships becomes the purest speculation...." The reasons may be obvious. Macroorganisms can be visualized and collected and morphologically, physiologically, ecologically and genetically described with relative ease. Microorganisms on the other hand are invisible to the naked eye, collections involve cultivation methods that allow recovery of only those that can grow under the conditions presented, their morphologies are exceedingly simple and relatively unvaried, and their diverse phenotypic properties are relatively useless for understanding evolutionary relationships. Because of these limitations, microbiology as a discipline was relegated to the applied side of science (i.e. tools to help the human condition) resulting in countless applications for food science, disease and medicine, genetics, physiology and cellular biology.

Years after the Stanier lament, Woese[4] changed the paradigm of microbiology by describing the three-domain "tree of life" based on sequencing the 16S rRNA molecule of prokaryotes (18S rRNA of eukaryotes). For the first time, the full scope of prokaryotic diversity was placed within the confines of phylogenetic relatedness. Classification based upon evolutionary relationship, once thought impossible, is now possible. The new classification scheme inspired Norman Pace and others[5] to recognize that microorganisms could be identified in situ (without cultivation) by comparing "naturally" occurring 16S rRNA molecule sequences (obtained via molecular techniques) to sequences of cultivated strains in the three-domain tree. Free from the confines of cultivation, microbial ecologists began natural history surveys that further demonstrated the great diversity of microorganisms and stimulated interesting questions about the causes of such diversity. For instance, 16S rRNA analysis of cyanobacterial mats residing in Octopus and Mushroom hot springs in Yellowstone National Park demonstrated that the in situ 16S rRNA gene sequences were different from those of cultivated isolates[6] and that closely related *Synechococcus* (unicellular cyanobacteria) were uniquely distributed across temperature and light gradients[7,8]. It was suggested[9] that the evolutionary/ecological theory, adaptive radiation (i.e. differential adaptation to various environments) could explain the observed relationship of the genotypes to their unique niches, a theory modeled after the adaptation of "Darwin's finches" to different niches on the Galapagos Islands. However, without further distribution analysis (e.g. global sampling) it cannot be determined if the Yellowstone *Synechococcus* radiated within Yellowstone's borders or if they have a wider distribution.

Microbial Biogeography

It is interesting to note that with the new microbial paradigm, lots of problems have been solved, but new problems have arisen. Perhaps the biggest obstacle in the field of microbial biogeography is the question of identity (i.e. how do we know if two populations belong to the same species or if two organisms belong to the same population?). This is of extreme importance when trying to determine the geographic range of a specific species or population. To differentiate species or populations, it is critical to have an established set of criteria, which can be applied to and measured on individuals. In some cases this can be relatively easy. While in New Guinea, Ernst Mayr[10] collected and identified 138 species of birds of which the island's indigenous people identified 137, suggesting that species are not arbitrarily defined but universally accepted regardless of who is counting. However, the myriad of species definitions or concepts, contradicts this notion[11]. Furthermore, it is difficult to identify a single species concept that can be applied to all groups of organisms, extant and extinct, haploid, diploid and polyploid, sexual and asexual or macroorganism and microorganism. Perhaps Darwin[12] expressed the problem best when he wrote "there is no possible test but individual opinion to determine which...shall be considered as species and which as varieties." If species are so difficult to define, then perhaps that unit of identity should not be used, especially with respect to microorganisms where separate species have been arbitrarily defined as organisms with less than 70% similarity in DNA-DNA hybridization,[13] which roughly correlates to 97% similarity at the 16S rRNA locus[14].

Lately, molecular markers (e.g. gene sequence variation) have been extremely successful for linking relatedness with the distribution of organisms[1]. Genetic relatedness can thus be used to define identity. This is appropriate as divergence is really the issue, not what species are. In attempts to determine the biogeography of microorganisms, molecular markers, especially the 16S rRNA gene, have been used in addition to more classical methods of identification (e.g. phenotypic properties). However, there has been little conformity in which measurements should be used to determine the geographic ranges of the studied organisms. The use of conserved genes, like 16S rRNA, to identify and/or define populations may be particularly problematic as they may underestimate the actual diversity and thereby artificially expand our impressions of territorial range. The pitfalls in choosing a wrong level of analysis for determining identity (e.g. morphology, gene restriction enzyme fragment patterns or conserved vs. variable gene sequences) will be considered below.

In a study using microscopy to determine the species diversity of ciliated protozoa (large unicellular eukaryotes whose species are morphologically defined), Fenchel et al.[15] reported (from their study plus others) 181 and 146 species recovered from two ecologically different sediments occurring in a pond (Priest Pot, UK) and shallow bay (Nivå Bay, Helsingør, Denmark). They determined that the diversity discovered was approximately 11% of the total number of free-living ciliate species. Furthermore, they reasoned that similar results would have been found if additional nearby ciliate habitats had been sampled (i.e. 10-20 ecologically different sites). The authors were confident that if the more comprehensive sampling regime had been performed "a very substantial

fraction of all known ciliates" would have been recovered from a relatively small geographical range. From their interpretation of the data, they concluded, "everything is (almost) everywhere". However, the conclusion may be oversimplified. Organisms that live in similar habitats can often have similar morphologies via convergent or parallel evolution thereby concealing genetic diversity within a morphotypically-defined species. Indeed, many planktonic foraminifera species (morphotypically-defined) are comprised of more than one genotype and these geontypes have been considered to be cryptic sibling species[16-18]. In prokaryotes, all unicellular coccoid to rod-shaped cyanobacteria fall within the genus *Synechococcus* [19]. However, this genus is not monophyletic, as the morphology has independently evolved many times[20]. As both examples clearly demonstrate, it is risky to make conclusions about the distribution of microorganisms when identity is based solely upon morphological criteria.

As expressed above, diversity and distribution studies of microorganisms are often performed using the 16S rRNA molecule either by restriction enzyme analysis or by direct sequencing of the molecule. In an attempt to survey the archaeal diversity present in the world's oceans (North Atlantic, Cantabrian Sea [Atlantic Ocean], the Mediterranean Sea, the Santa Barbara Channel [Pacific Ocean], and the Drake Passage [Southern Ocean]) Massana et al.[21] generated 16S rRNA gene libraries from natural samples. They used two restriction enzymes to construct restriction fragment length polymorphism (RFLP) patterns from their clone libraries and interpreted any RFLP patterns that were identical as a single operational taxonomic unit (OTU). The analyses of Massana et al.,[21] revealed that 5 of the 36 OTU's (representing 87% of the analyzed

clones) were "cosmopolitan". The RFLP method is insensitive, as restriction enzymes recognize a very small proportion (e.g. 4-8 nucleotides) of the molecule analyzed. In a computer simulation using prokaryotic 16S rRNA gene sequences, Moyer et al.,[22] tested the efficacy of restriction enzymes in determining the diversity of microorganisms. They found that RFLP could only differentiate among sequences that were at least 3.9% different. This clearly leaves a lot of diversity undetected, especially considering the extremely conserved nature of the 16S rRNA locus. In a study using *Pseudomonas* strains isolated from soil samples collected around the world, Cho and Tiedje [23]compared the effectiveness of 16S rRNA RFLP patterns with repetitive extragenic palindromic-PCR (REP-PCR, a very sensitive method that takes advantage of the entire genomic diversity) for detecting endemic genotypes. In the case of 16S rRNA RFLP pattern analysis, only 4 OTU's were found among 248 isolates and all 4 appeared cosmopolitan in distribution. However, when REP-PCR was applied to each of the strains, 85 genotypes were recovered and identical genotypes were only found in samples from the same geographic sites, indicating high levels of endemism among the strains. Mehta et al.,[24] found similar results when they analyzed *Zylella fastidiosa* isolated from citrus trees in Brazil. It would seem that 16S rRNA RFLP patterns completely underestimate the true diversity of microorganisms and any conclusions as to "cosmopolitan phylotypes" should be avoided when using this technique.

Similar or identical 16S rRNA gene sequences have been used to declare that some organisms have a worldwide distribution. Indeed, Garcia-Pichel et al.,[25] found identical or nearly identical 16S rRNA gene sequences from hypersaline-adapted

cyanobacteria living in microbial mats from Europe, the Middle East and Baha, Mexico. They concluded that the cyanobacterial species *Microcoleus chthonoplastes* is cosmopolitan. Zwart et al., [26] also found nearly identical 16S rRNA genes from lakes located in North America and Europe, and conjectured that the same species has a global distribution. Although 16S rRNA sequence variation is more sensitive than 16S rRNA RFLP pattern analysis for determining identity, 16S rRNA sequence variation may also unnaturally expand our view of population ranges since the 16S rRNA locus is evolutionarily conserved. For instance Ferris and Ward[7] found that two 16S rRNA genes differing by a single nucleotide had unique distributions along a thermal gradient. Because the 16S rRNA genes were found in different habitats, it was argued that genes were retrieved from different species[27]. Since this locus is barely able to detect differentially adapted populations, it may also be too conserved to detect differences in geographic populations. It is also likely that small changes in the 16S rRNA actually reflect major changes in the organism. The average rate of substitution for 16S and 18S rRNA molecules has been calculated to be 1% per 50 million years[28-30]. This translates to one nucleotide substitution per 3.3 million years, suggesting that two organisms with nearly identical 16S rRNA genes have been divergent for a very long time. The evidence suggests that spatially separated organisms should not be interpreted as having a cosmopolitan distribution when slight differences are detected at the 16S rRNA locus. Indeed, the opposite interpretation may be more likely.

It is difficult to cast blame on researchers for using conserved loci to establish identity, because such genes are commonly assayed and there are often databases to

which results can be compared. However, researchers should recognize the limits of the methods before drawing conclusions. If progress is to be made in microbial biogeography, it is likely that more informative molecular markers with greater resolving power will have to be used. For instance, the DNA-dependent RNA polymerase gene (*rpo*C1) evolves much faster than the 16S rRNA gene. *Synechococcus* sp. strains WH7805 and WH8103 differ by 1.4% at the 16S rRNA locus, but differ by 17% at the *rpo*C1 locus[31]. The internal/intervening/intergenic transcribed spacer (ITS) region located on the rRNA operon between the 16S and 23S rRNA genes also has a much higher resolving power [32-34]. However, for in situ analysis, the ITS region has additional benefits. Because the ITS is adjacent to the 16S rRNA gene, it is possible to PCR amplify both loci simultaneously using the 16S rRNA gene to relate the sequence phylogenetically to other known organisms while using the ITS to discriminate between closely related genetic variants with identical 16S rRNA sequences.

Goals of the Thesis

If Rosenzweig and other evolutionary biologists[2,10,35-37] are correct in thinking that geographic isolation is one of the major causes of speciation, then perhaps it is time for microbiologists to understand this biological paradigm and apply it to investigations concerning microbial diversification and distribution, especially since most of the putative evidence (and dogma) that supports the "cosmopolitan" hypothesis is based on observations that can easily be challenged. With this admonishment in mind, it is the

goal of this thesis to provide convincing evidence that microorganisms can become geographically isolated, that isolation can lead to diverging populations and consequently that genetic drift may play an active role in the evolution of microorganism independently of adaptation (via mutation and lateral gene transfer) and natural selection.

My approach to microbial biogeography was to take advantage of the island-like nature of hot springs and previous observations concerning the diversity and distribution of thermophilic cyanobacterial populations from around the globe. Anomalous distributions such as the lack of high-temperature adapted cyanobacteria in regions outside of the U.S.A.[38,39] led to the main hypotheses:

Synechococcus mats in globally separated hot springs are dominated by A/B genotypes, but there is a barrier to the dispersal of higher temperature-adapted A-like genotypes.

Or, alternatively, mats in globally separated hot springs are dominated by *Synechococcus* unrelated to A/B genotypes.

The first hypothesis supports the idea that everything is everywhere, but nature selects. The hypothesis predicts that both B and A-type *Synechococcus* are ubiquitously dispersed; the inability of type-A *Synechococcus* to live above 63C in some hot springs is explained by environmental selection (e.g., sulfide in combination with high temperature is known to prevent the growth of cyanobacteria[40-41]). The alternative hypothesis is consistent with geographic isolation. A test of either hypothesis must also

address the possibility that distribution is patterned according to adaptation to specific physical/chemical parameters.

To test these hypotheses, I made extensive collections from hot springs in Italy, New Zealand, Japan and the northwest United States and analyzed samples by molecular methods of suitable resolution. I developed a 16S rRNA method that allows genetic comparisons to previous 16S rRNA studies while simultaneously sampling a higher resolution genetic marker (ITS region) for detecting sequence variation between identical or nearly identical 16S-rRNA defined genotypes. This is important because 16S rRNA gene sequences are likely to conceal geographical isolation given their conserved nature. I also generated group-specific 16S rRNA probes to quantify populations in their various locations and, using PCR, to detect rare genotypes that may be present but difficult to detect given the limitations of detection methods. Furthermore, to convincingly demonstrate the role of adaptation or niche specialization in determining the distribution of the thermophilic cyanobacteria, in-depth analysis of the physical/chemical parameters of sampled hot springs was performed. Because the results of this work could potentially shift theoretical paradigms in microbiology, chapter 2 was prepared as a research article for the journal *Nature* and the experimental results are thus presented in a condensed style. Furthermore, much additional literature detail is placed intentionally in a minireview (chapter 3) designed to add my results to a growing body of evidence on physical isolation in microbial evolution, an issue that needs to be emphasized to microbiologists.

REFERENCES CITED

1 Avise, J. 2000. *Phylogeography: The History and Formation of Species.* Harvard University Press, Cambridge, MA.

2 Rosenzweig, M. 1996. *Species Diversity in Space and Time.* Cambridge University Press, UK.

3 Stanier, R.Y., Doudoroff, M. and Adelberg, E.A., 1963. *The Microbial World,* 2nd ed., Prentice-Hall, Inc., Englewood Cliffs, NJ.

4 Woese C. 1987. Bacterial evolution. *Microbiol. Rev.* 51:221-271.

5 Olsen, G.L., Lane, D.J., Giovannoni, S.J., and Pace, N.R. 1986. Microbial ecology and evolution: a ribosomal RNA approach. *Ann. Rev. Microbiol.* 40:337-365.

6 Ward, D.M., Weller, R. and Bateson, M.M. 1990. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* 345: 63-65.

7 Ferris, M., and Ward, D. 1997. Seasonal distributions of dominant 16S rRNA-defined populations in a hot spring microbial mat examined by denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* 63:1375-1381.

8 Ramsing, N., Ferris, M., and Ward, D. 2000. Highly ordered vertical structure of *Synechococcus* populations within the one-millimeter-thick photic zone of a hot spring cyanobactrial mat. *Appl. Environ. Microbiol.* 66:1038-1049.

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

10 Mayr, E. 1991. *One Long Argument: Charles Darwin and the Genesis of Modern Evolutionary Thought.* Harvard University Press, Cambridge, MA.

11 Claridge, M., Dawah, H., and Wilson, M. (eds.) 1997. *Species: the Units of Biodiversity.* Chapman & Hall, London.

12 Darwin, C. 1859. *Origin of Species, by Means of Natural Selection of the Preservation of Favoured Races in the Struggle for Life.* Mentor ed., Penguin Books Ltd., Harmondsworth, UK.

13 Wayne, L.G., Brenner, D.J., Colwell, R.R., Grimont, P.A.D., Kandler, O., Krichevsky, M.I., Moore, L.H. Moore, W.E.C. Murray, R.G.E., Stackebrandt, E., Starr, M.P., and Trüper, H.G. 1987. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol.* 37:463-464.

14 Stackebrandt, E., and Goebel B.M. 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. System Bacteriol.* 44:846-849.

15 Fenchel, T., Esteban, G. and Finlay, F. 1997. Local versus global diversity of microorganisms: cryptic diversity of ciliated protozoa. *OIKOS* 80:220-225.

16 Huber, B., Bijma, J. and Darling K. 1997. Cryptic speciation in the living planktonic foraminifer *Globigerinella siphonifera* (d'Orbigny). *Paleobiol.* 23:33-62.

17 de Vargas, C., Norris, R., Zaninetti, L., Gibb, S., and Pawlowski, J. 1999. Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanic provinces. *Proc. Natl. Acad. Sci. USA* 96:2864-2868.

18 Darling, K., Wade, C., Kroon, D., Leigh Brown, A., and Bijma, J. 1999. The diversity and distribution of modern planktic foraminiferal small subunit ribosomal RNA genotypes and their potential as tracers of present and past ocean circulations. *Paleoceanogr.* 14:3-12.

19 Waterbury, J., and Rippka, R. 1989. Subsection I. Order *Chroococcales* Wettstein 1924, emend. Rippka et al., 1979. in *Bergey's Manual of Systematic Bacteriology*. Staley, J., Bryant, M., Pfennig, N., and Holt, J. (eds.) Williams & Williams. Baltimore.

20 Turner, S., Pryer, K. Miao, M. and Palmer, J. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J. Eukaryot. Microbiol.* 46:327-338.

21 Massana, R., DeLong, E., and Pedros-Alio, C. 2000. A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces. *Appl. Environ. Microbiol.* 66:177-187.

22 Moyer, C., Tiedje, J., Dobbs, F., and Karl, D. 1996. A computer-simulated restriction fragment length polymorphism analysis of bacterial small-subunit rRNA genes: efficacy of selected tetrameric restriction enzymes for studies of microbial diversity in nature. *Appl. Environ. Microbiol.* 62:2501-2507.

23 Cho, J.-C., and Tiedje, J. 2000. Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. *Appl. Environ. Microbiol.* 66:5448-5456.

24 Mehta, A., Leite, R. Jr., Rosato, Y. 2001. Assessment of the genetic diversity of *Xylella fastidiosa* isolated from citrus in Brazil by PCR-RFLP of the 16S rDNA and 16S-23S intergenic spacer and rep-PCR fingerprinting. *Antonie Van Leeuwenhoek* 79:53-59.

25 Garcia-Pichel, F., Prufert-Bebout, L., and Muyzer, G. 1996. Phenotypic and phylogenetic analyses show *Microcoleus chthonoplastes* to be a cosmopolitan cyanobacterium. *Appl. Environ. Microbiol.* 62:3284-3291.

26 Zwart, G., Hiorns, W., Methe, B., van Agterveld, M., Huismans, R., Nold, S., Zehr, J., and Laanbroek, H. 1998. Nearly identical 16S rRNA sequences recovered from lakes in North America and Europe indicate the existance of clades of globally distributed freshwater bacteria. *System. Appl. Microbiol.* 21:546-556.

27 Ward, D. 1998. A natural species concept for prokaryotes. *Curr. Opin. Microbiol.* 1:271-277.

28 Ochman, H., and Wilson, A. 1987. Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *J. Mol. Evol.* 26:74-86.

29 Darling, K., Wade, C., Stewart, I., Kroon, D., Dingle, R., and Leigh-Brown, A. 2000. Molecular evidence for genetic mixing of Arctic and Antarctic subpolar poulations of planktonic foraminifers. *Nature* 405:43-47.

30 Baumann, P., Baumann, L., Lai, C., Rouhbakhsh, D., Moran, N. and Clark, M. 1995. Genetics, physiology, and evolutionary relationships of the genus *Buchnera*: intracellular symbionts of aphids. *Annu. Rev. Microbiol.* 49:55-94.

31 Toledo, G. and Palenik, B. 1997. *Synechococcus* diversity in the California current as seen by RNA polymerase (*rpo*C1) gene sequences of isolated strains. *Appl. Environ. Microbiol.* 63:4298-4303.

32 Frothingham, R., and Wilson, K. 1993. Sequence-based differentiation of strains in the *Mycobacterium avium* complex. *J. Bacteriol.* 175:2818-2825.

33 Scheinert, P., Krausse, R., Ullmann, U., Soller, R., and Krupp, G. 1996. Molecular differentiation of bacteria by PCR amplification of the 16S-23S rRNA spacer. *J. Microbiol. Meth.* 26:103-117.

34 Chun, J., Rivera, I., Colwell R. 2002. Analysis of 16S-23S rRNA intergenic spacer of *Vibrio cholerae* and *Vibrio mimicus* for detection of these species. *Methods Mol. Biol.* 179:171-178.

35 Futuyma, D., 1986. *Evolutionary Biology* 2nd ed. Sinauer Associates, Inc. Sunderland, MA.

36 Wilson, E. 1992. *The Diversity of Life.* W.W. Norton & Co. New York.

37 Mayr, E. 1942. *Systematics and the Origin of Species.* Columbia University Press, New York.

38 Castenholz, R.W. 1996. Endemism and biodiversity of thermophilic cyanobacteria. *Nova Hedwigia, Beiheft* 112: 33-47.

39 Castenholz, R.W. 1978. The biogeography of hot spring algae through enrichment cultures. *Mitt. Internat. Verein. Limnol.* 21: 296-315.

40 Castenholz, R.W. 1976. The effect of sulfide on the bluegreen algae of hot springs. I. New Zealand and Iceland. *J. Phycol.* 12: 54-68.

41 Ward, D.M. and Castenholz, R.W. 2000. in *The Ecology of Cyanobacteria.* Whitton, B.A. & Potts, M. (eds.) Kluwer Academic Publishers, The Netherlands.

CHAPTER 2

GEOGRAPHIC ISOLATION AND THE EVOLUTION OF HOT SPRING CYANOBACTERIA*

Introduction

Genomics and comparative molecular phylogeny have fueled intense consideration of molecular mechanisms for the generation of genetic variation in bacteria, especially lateral gene transfer, and of the role these mechanisms may play in microbial evolution[1,2]. However, less attention has been given to environmental factors that act upon such variation to cause divergence and speciation. Ecological and geographic isolation are recognized as major causes of adaptive and allopatric speciation[3,4]. Microbial ecologists have begun to discover evidence suggesting adaptive radiations as they have used molecular methods to assay microbial diversity[5,6] and distribution patterns along well-defined ecological gradients within natural communities[7-13]. There is, however, considerable debate over the importance of geographic isolation in bacterial speciation.

It has been commonly assumed since early in the 20th century that in the case of microorganisms "everything is everywhere and nature selects"[14,15]. This suggests that microorganisms readily disperse and do not become geographically isolated. Support for the ubiquitous dispersal of microorganisms has come from observations of diversity

*This study has been submitted to *Nature* as: Papke, R.T., N.B. Ramsing, M.M. Bateson and D.M. Ward. Geographic isolation and the evolution of hot spring cyanobacteria.

of protists in sediments, suggesting widespread distribution of morphospecies[16]. One important implication is that the absence of allopatric speciation explains why there appear to be fewer microbial species than expected from correlations between body size and number of species[17]. A recent molecular study of protist diversity in polar oceans demonstrated that the same genetic variants, defined by 18S rRNA sequence variation, were present at both poles[18], further supporting the idea of ubiquitous dispersal and the rarity of allopatry in microbial evolution. It was noted, however, that some closely related genetic variants did exhibit unipolar distribution, and concern was raised that higher resolution genetic markers might be needed to discern geographic patterning[19]. Studies of bacterial diversity and distribution in marine[20,21] and near-marine[22] environments also suggest similar mixed patterns (i.e., the presence of identical as well as slightly different 16S rRNA variants in geographically separate sites). Studies of bacterial diversity and distribution in globally separate soil environments have revealed evidence of unique geographic distributions, but only when methods offering more genetic resolution than 16S rRNA sequence variation were employed[23,24]. In the face of conflicting reports it seemed informative to examine environments where geographic isolation is a prominent feature and thus likely to contribute to diversification. As pointed out by MacArthur and Wilson[25] "...in the science of biogeography, the island is the first unit that the mind can pick out and begin to comprehend".

Hot Spring Mats as Island-like Communities

Hot springs are well-isolated habitats occurring as clusters in globally distant regions and the microorganisms that inhabit them are extremophiles adapted to conditions quite different from the ambient milieu through which they would have to disperse. As such, one would expect that geographic isolation might be an important component to the diversification of hot spring microorganisms. Castenholz[26,27] observed anomalous distributions of cyanobacterial morphotypes inhabiting hot springs around the world. In well-studied North American hot springs such mats are formed by rod-shaped unicellular cyanobacteria of the genus *Synechococcus* with an upper temperature limit of 72°C. Ecologically similar strains are apparently absent from cyanobacterial mats in Japanese, New Zealand, Italian and African hot springs, where *Synechococcus* is reported to occur below ca. 63°C, the upper temperature limit for cyanobacterial mat development. *Synechococcus* was not observed at all in hot springs in Iceland, Alaska and the Azores, even though a pure culture of *Synechococcus* would grow in water from Iceland (Castenholz, personal communication).

Molecular analysis has revealed great diversity within the thermophilic *Synechococcus*[7,28] morphotype. Three unrelated phylogenetic lineages (separated by >10% 16S rRNA sequence variation) containing organisms of this morphotype, termed A/B, C1and C9, have been detected (Figure 1). The predominant *Synechococcus* in Yellowstone hot springs detected by direct molecular analysis is the A/B type[29]. On the basis of distribution[7,10] and pure culture studies[30], the A/B lineage appears to have

diverged into high- and low-temperature adapted A-like and B-like clades, respectively (Figure 1). Furthermore, different genotypes occurred at different depths in the mat[8], leading us to suggest that the pattern of diversity in this lineage resulted from an adaptive radiation[7]. *Synechococcus* spp. C1 and C9 genotypes were also detected in the same Yellowstone spring through cultivation and were less abundant and diverse.

The morphological observations of biogeographical anomaly and our molecular observations in Yellowstone hot springs led us to the following alternative hypotheses regarding biogeographical influences on the distribution and evolution of hot spring *Synechococcus*:

Synechococcus mats in globally separate hot springs are dominated by A/B genotypes, but there is a barrier to the dispersal of higher temperature-adapted A-like genotypes.

Or, alternatively, mats in globally separate hot springs are dominated by *Synechococcus* unrelated to A/B genotypes.

The first hypothesis supports the idea that everything is everywhere, but nature selects. The hypothesis predicts that both B and A-type *Synechococcus* are ubiquitously dispersed; the inability of type-A *Synechococcus* to live above 63C in some hot springs