Identification and distributions of dominant bacterial populations in hot spring Synechococcus mats
by Michael Joseph Ferris

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

Abstract:
Molecular analyses of 16S rRNA sequences retrieved from the Octopus Spring cyanobacterial mat revealed the presence of numerous uncultivated bacterial populations. Notably absent was the sequence of Synechococcus lividus, the cyanobacterium believed to construct the mat. Synechococcus-shaped cells, however, were abundant, typically reaching concentrations of 10^10 ml^-1. By diluting mat inocula prior to enrichment, Synechococcus isolates with sequences other than that of S. lividus were obtained. One of these sequences, type-P, matched that of an Octopus Spring clone. This was the first match between 16S rRNA sequences observed in the mat and a pure culture and proved that cloned sequences are representative of viable mat populations. Conversely, only S. lividus isolates were retrieved from enrichments inoculated with undiluted mat. Moreover, the 16S rRNA sequence of S. lividus was the only one detected in a survey of all thermophilic Synechococcus culture collection strains. In contrast, the degree of diversity among Octopus Spring cyanobacterial sequences spanned that of the entire cyanobacterial lineage. Thus, enrichment culture and simple morphology concealed remarkable diversity which exists among thermophilic cyanobacteria.

Molecular detection and classification of bacterial populations has revolutionized the study of bacteria in nature. Ultimately, microbial ecologists seek to describe the distributions of bacterial populations and determine the factors that control their abundance. A new molecular technique, denaturing gradient gel electrophoresis (DGGE), was used to detect and evaluate the distributions of dominant bacterial populations in hot spring mats. The DGGE profiles revealed that closely related populations of cyanobacteria and green nonsulfur-like bacteria were distributed at discrete intervals along thermal and vertical gradients, and suggested that some populations may be colonist species.

Thus, despite their close phylogenetic relatedness, each population represented a distinct ecotype. This suggests that clusters of phylogenetically similar 16S rRNA sequence types, which are commonly detected in many habitats, reflect the evolutionary diversification Of populations to fill environmental niche space.

IDENTIFICATION AND DISTRIBUTIONS OF DOMINANT BACTERIAL POPULATIONS IN HOT SPRING *SYNECHOCOCCUS* MATS

by

Michael Joseph Ferris

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

of

Doctor of Philosophy

in

Microbiology

MONTANA STATE UNIVERSITY - BOZEMAN
Bozeman, Montana

January 1997

APPROVAL

of a thesis submitted by

Michael Joseph Ferris

This thesis has been read by each member of the thesis 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) 21 January 1997 (Date)

Approved for the Department of Microbiology

Dr. Al J. Jesaitis _____ (Signature) 1-21-97 (Date)

Approved for the College of Graduate Studies

Dr. Robert L. Brown _____ (Signature) 1/23/97 (Date)

STATMENT 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 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 Microform 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 microfilm along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Signature Michael J. Ferris

Date 1-21-97

ACKNOWLEDGEMENTS

I thank Dave Ward for his many acts of kindness, for his countless hours of patient instruction, for sharing his knowledge and for inviting me to join him in his research, the rewards of which I can only repay by promising to share what I have learned with future students.

Thanks also to Nancy Ward for sharing her home and wonderful dinners with me and the rest of the Ward lab.

Thank you, to Mary Bateson for your help in the lab, and both Mary and Frodo for simply being good friends.

Thanks to the group of fellow scientists who made my experience here much more enriching, Steve Nold, Niels Ramsing, Michael Friedrich, Sjila Santegoeds, Eric Kopczynski, Eric Kern and Thayne Papke.

Special thanks to Scott Kobayashi for the many (but never enough) enjoyable hours of fly fishing along Montana's rivers together and the kind of memorable days a person keeps for life.

I extend much appreciation to Bob Lindstrom and the Yellowstone National Park service for their support and assistance.

This research was funded by a grant from the National Science Foundation (BSR-9209677).

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

ABSTRACT

Molecular analyses of 16S rRNA sequences retrieved from the Octopus Spring cyanobacterial mat revealed the presence of numerous uncultivated bacterial populations. Notably absent was the sequence of *Synechococcus lividus*, the cyanobacterium believed to construct the mat. *Synechococcus*-shaped cells, however, were abundant, typically reaching concentrations of 10^{10} ml^{-1}. By diluting mat inocula prior to enrichment, *Synechococcus* isolates with sequences other than that of *S. lividus* were obtained. One of these sequences, type-P, matched that of an Octopus Spring clone. This was the first match between 16S rRNA sequences observed in the mat and a pure culture and proved that cloned sequences are representative of viable mat populations. Conversely, only *S. lividus* isolates were retrieved from enrichments inoculated with undiluted mat. Moreover, the 16S rRNA sequence of *S. lividus* was the only one detected in a survey of all thermophilic *Synechococcus* culture collection strains. In contrast, the degree of diversity among Octopus Spring cyanobacterial sequences spanned that of the entire cyanobacterial lineage. Thus, enrichment culture and simple morphology concealed remarkable diversity which exists among thermophilic cyanobacteria.

Molecular detection and classification of bacterial populations has revolutionized the study of bacteria in nature. Ultimately, microbial ecologists seek to describe the distributions of bacterial populations and determine the factors that control their abundance. A new molecular technique, denaturing gradient gel electrophoresis (DGGE), was used to detect and evaluate the distributions of dominant bacterial populations in hot spring mats. The DGGE profiles revealed that closely related populations of cyanobacteria and green nonsulfur-like bacteria were distributed at discrete intervals along thermal and vertical gradients, and suggested that some populations may be colonist species. Thus, despite their close phylogenetic relatedness, each population represented a distinct ecotype. This suggests that clusters of phylogenetically similar 16S rRNA sequence types, which are commonly detected in many habitats, reflect the evolutionary diversification of populations to fill environmental niche space.

CHAPTER 1

IDENTIFICATION AND DISTRIBUTION OF DOMINANT BACTERIAL POPULATIONS IN HOT SPRING *SYNECHOCOCCUS* MATS: INTRODUCTION

Traditional microbiological methods have restricted progress in microbial ecology. Historically, surveys of microorganisms in natural samples have been conducted using microscopy and cultivation. Undeniably, these studies have been productive and invaluable in many respects. From them we have learned that prokaryotic organisms represent the most abundant, physiologically diverse and biotechnologically important assemblage of organisms known. Yet, we also realize that traditional methods are insufficient to discern the majority of microorganisms in natural samples (27), since genetically distinct bacteria often share similar simple morphologies and many bacteria fail to develop under laboratory culture conditions (1). Some microbiologists estimate that as few as 1% of the bacteria in nature have been characterized (14). As a result, *the true genetic diversity of bacteria is unknown.* Without a reliable means of detecting bacterial populations, many fundamental questions about their ecology cannot be answered.

Molecular Detection of Microbial Populations

Molecular techniques now offer a fresh approach to the detection and classification of bacteria (13,24,25). The realization that sequence information contained in protein and nucleic acid molecules can be used to identify organisms and establish phylogenetic relationships among them has revolutionized the study of bacteria in nature (28). Methods based on obtaining specific nucleotide sequence information directly from crude environmental samples can now be used to detect and identify microorganisms directly in their native habitats (18), circumventing the limitations of cultivation and microscopy.

Notable among the molecules used to characterize bacteria at the genetic level has been the 16S rRNA gene (12). Several properties make 16S rRNA especially useful in analyses of diverse bacterial assemblages. These include its ubiquity among all prokaryotic cells, its highly conserved regions which aid in establishing sequence alignments for phylogenetic analyses and its variable regions which distinguish between populations and are used as target sites for oligonucleotide primers or hybridization probes of varying specificities (2, 18). Beyond the phylogenetic utility of sequence information, the ability to detect individual bacterial populations using oligonucleotide probes and primers permits microbiologists to conduct more

ecologically relevant experiments. Finally, thousands of 16S rRNA sequences are conveniently accessible via on-line databases facilitating sequence analyses (8,11).

Discoveries achieved through 16S rRNA sequence analyses have been impressive; some have literally forced us to reevaluate the way we view the microbial world. For example, the direct retrieval of numerous novel 16S rRNA sequences from environmental samples *proved* speculations that the majority of bacteria in nature have not been recognized in culture (6,20). The symbiotic origins of eukaryotic mitochondria and chloroplasts are now generally accepted because their 16S rRNA sequences place them firmly in the proteobacterial and cyanobacterial lineages. The physical limit of prokaryotic cell size must be reevaluated because 16S rRNA analyses have revealed that "macroscopic" symbionts of surgeon fish are actually individual bacterial cells, *Epulopiscium fishelsonii*, representative of the low G+C Gram-positive bacteria (3). Perhaps most significantly, a three domain "tree of life", based on rational evolutionary theory and concrete genetic information has been established and a new domain, Archaea, has been created. The evolutionary distance separating the Archaea from the Bacteria is as great as that separating the Eukarya from the Bacteria (25, 26).

Ecology Defined

Admittedly, these findings have been impressive, but microbiologists are only beginning to utilize molecular approaches to investigate microorganisms in the context of the environments they inhabit. That is, we are just beginning to investigate the ecology of uncultivated, 16S rRNA-defined bacterial populations. Because I feel the traditional definition of ecology, "the study of the interactions of organisms with their environments" (4) is quite vague and does not convey the essential theme of ecological investigations, I am compelled to present a more informative description here. Simply stated, ecologists attempt to explain the distribution and abundance of organisms (4). This is done by examining biotic and abiotic aspects of the environment in which organisms exist, and by gathering information about their physiologies, behavior, genetics, and possible evolutionary histories. Obviously, this task encompasses an array of scientific disciplines, but the focus is narrowed by attempting to identify factors that most influence reproduction, migration, and death (the three elements that control distribution and abundance).

To organize their task, ecologists divide their approach into three levels of study, individuals, populations, and communities (4). Individuals provide information about physiology, behavior, and genetics. Populations are

collections of "interbreeding" individuals, possessing properties beyond those of the component organisms. Communities (or ecosystems) are collections of populations, interacting with each other and their environment. The community can be so complex that it is sometimes studied as an entity, with measurable properties such as complexity, resiliency, or biomass production. Because of the difficulties encountered in studying individual bacterial cells in their native habitats, much of our understanding of microbial communities has come from the study of their functional properties. Groups of organisms in which all members perform the same function, for example photosynthetic carbon dioxide fixation, are described as guilds (4). While studies at the guild level are informative, they can never resolve the complex interplay that occurs among the specific populations. Only the results of their collective activities are observed. In fact, so superficial is our understanding of bacterial populations that the microbial community is often depicted simply as a box labeled "decomposers" or "microbial loop".

Even though applying sound ecological thinking to the study of bacterial communities is wise, microbiologists must realize that a rigid adherence to macroecological principles is unjustified. There are, after all, fundamental differences between macro- and microorganisms, and many concepts that are clear when applied to macroorganisms become

clouded when applied to microbes. For example, what is death among perpetually reproducing clones? What is the life span of an endospore that remains viable for perhaps millions of years? Strictly speaking, even the macrobiotic concept of "species" is not applicable to bacteria since it is defined by the ability to interbreed.

Molecular Microbial Ecology

Armed with the appropriate molecular techniques, microbiologists can begin to address fundamental questions about the composition of bacterial communities and interactions occurring within them by focusing on the member populations. This approach promises to reveal details and intricacies that permit the development of more accurate, realistic models of community structure, and perhaps even reveal interactions at the level of individual cells (2). Such information will greatly increase our understanding of the earth's microbiota and will ultimately allow us to predict the effects of environmental perturbations on microbial communities. This will in turn allow manipulation of environmental conditions for practical benefit and, more importantly, will provide an informed knowledge-base with which to better manage and preserve natural microbial resources.

Certainly, microbial ecologists are far from attaining these goals, but it is exciting simply to be in a position

where we can now think of these as realistic possibilities. The first step in this direction has already begun with numerous molecular surveys of bacterial populations occurring in diverse habitats world-wide. The second step, relating the distributions of these populations to their environments is just beginning. This was the primary focus of my studies.

Even with molecular tools, surveying bacterial populations in nature is not trivial. Achieving a true understanding of bacterial communities will require microbial ecologists to remain objective about the populations they observe and the techniques they use to detect them. Molecular microbial ecology is still somewhat of a burgeoning field. As such, much of the methodology remains technically demanding and is continuously undergoing refinement. As in any method used to sample biological systems, it is important to appreciate the biases that can be present in molecular procedures. For example, in 16S rRNA-based population surveys, chimeric sequences that may form during mixed-template polymerase chain reaction (PCR) amplifications can sometimes be misinterpreted as new populations (7). It is also possible for the amplification of some templates to be favored, skewing PCR product ratios such that they do not reflect initial population abundances (15,17). So, while exciting new information about the ecology of novel bacterial populations is being gathered, an equally critical aspect of every investigation must be to evaluate how representative of

the true community the detected populations are. Since we do not yet fully understand or perhaps even recognize some methodological biases, significant populations may be present that are not being observed.

The Octopus Spring Mat as a Model Microbial Community

Another factor complicating the study of microbial communities is the propensity of bacterial populations to change rapidly and dramatically in response to fluctuating environmental conditions. It is not unreasonable to expect new populations to appear while other populations recede below detectable levels within hours after an environmental perturbation. Obviously, designating the "native" populations under such conditions is problematic. It is sensible then that some microbial ecologists have sought to work in habitats where environmental variations are minimized. Geothermally heated springs, such as those found in Yellowstone National Park, offer these conditions, making them attractive model systems in which to pioneer molecular-based studies of microbial communities (19). One spring in particular, Octopus Spring, in the Lower Geyser Basin of Yellowstone National Park has been investigated as a model. A large portion of the Octopus Spring microbial mat receives a constant input of hot water which restricts the community to exclusively prokaryotic organisms, simplifying community composition and making ambient temperature changes

irrelevant. An added advantage is that this mat has been extensively studied using traditional techniques, and thus provides a unique opportunity to compare populations detected by molecular approaches to those detected using cultivation and microscopy (18, 21). Most importantly, the Octopus Spring mat provides easily measured, environmentally relevant gradients of parameters such as temperature, light, oxygen and pH (21), along which bacterial population distributions can be evaluated. This information is essential to understanding the factors controlling microbial population distributions and abundances.

A 16S rRNA Perspective of the Octopus Spring Mat Community

Initial 16S rRNA cloning and sequencing analyses of the Octopus Spring mat revealed a striking incongruence between the populations detected using molecular approaches and those detected using cultivation or microscopy. Numerous 16S rRNA sequences were retrieved but none matched those of any cultivated isolate, even those believed to inhabit Octopus Spring (20). Notable among the new 16S rRNA sequence types were several cyanobacterial populations defined as types A, B, I, J and P, none of which matched the 16S rRNA sequence of the unicellular cyanobacterium, *Synechococcus lividus*, believed to construct the mat (22). This was true despite the obvious presence of numerous *Synechococcus*-shaped cells in samples used to construct the clone libraries. Subsequent

oligonucleotide probe analyses indicated that the cloned 16S rRNA sequence types were detectable in the mat and that the populations harboring these sequences might be adapted to different environmental parameters (16). The 16S rRNA sequence of *S. lividus* was below detection in the mat, but was readily detected in enrichment cultures generated using undiluted Octopus Spring mat as inoculum. It was also not detected if the inoculum was homogenized and diluted prior to enrichment. These results implied that *S. lividus* was not a predominant species. In contrast, positive probe responses to several of the cloned Octopus Spring cyanobacterial sequence types (e.g. types A and B) were readily observed in enrichments inoculated with the highly diluted mat homogenates, implying that such populations were predominant in the mat.

Denaturing Gradient Gel Electrophoresis (DGGE) to Study Population Distributions

Much of my work was directed toward elucidating the factors that influence distributions of 16S rRNA-defined populations in the Octopus Spring mat, with particular emphasis on the cyanobacterial populations. Originally, this was to be accomplished by expanding upon oligonucleotide probe studies (16). However, before these studies were initiated, denaturing gradient gel electrophoresis (DGGE) analysis of PCR-amplified 16S rRNA gene segments was

introduced as a method to analyze complex microbial communities (9,10). DGGE offered several advantages over probing. For example, only known populations can be detected with oligonucleotide probes, whereas with DGGE, previously unrecognized community members can be detected. This was particularly important since my dilution culture studies had demonstrated the existence of cyanobacterial populations whose 16S rRNA sequences had not been detected in clone libraries (Chapter 2). Also, unlike probing, DGGE offered the potential to identify multiple populations in a single analysis, enabling more ecologically robust population surveys.

Figure 1 illustrates the DGGE procedure. DNA is extracted from an environmental sample and the PCR is used to amplify the gene of interest (in this case a segment of the 16S rRNA gene). Since there are multiple populations within the sample, multiple 16S rRNA genes are amplified (as indicated by different colors in Figure 1). The size of each PCR product is essentially identical, but the sequences contributed by different community members are unique. Two properties of double stranded DNA (dsDNA) facilitate an understanding of the principle used to separate these PCR products. First, resistance of dsDNA to denaturation is dependent upon the number and arrangement of GC base pairs. Second, the electrophoretic mobility of dsDNA decreases dramatically when it becomes partially denatured. In DGGE,

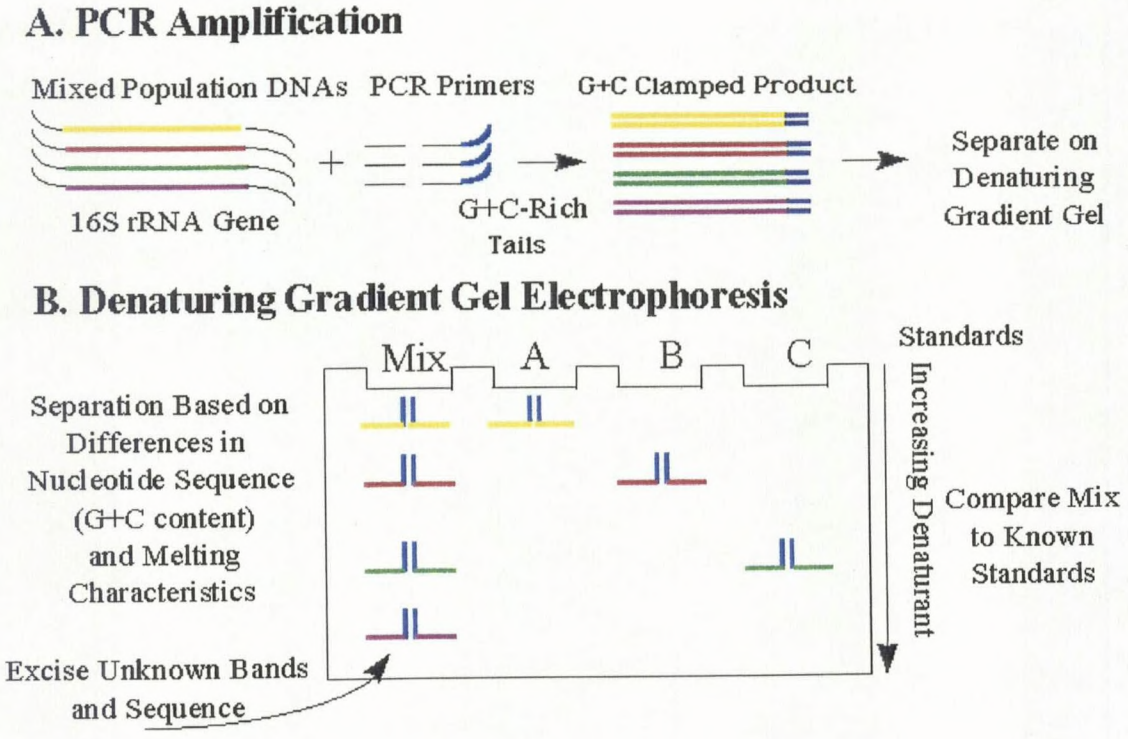

Figure 1. PCR amplification and DGGE analysis of 16S rRNA-defined populations. A, PCR illustrating amplification of mixed populations of 16S rRNA genes using primers complementing conserved sequence domains; one primer has a GC-tail. B, DGGE methodology illustrating separation of PCR products on a gradient gel and comparison to purified standards.

one PCR primer is designed with a 40 bp GC-tail, which generates PCR products having GC-clamp regions (blue in Figure 1). These PCR products are electrophoresed through an acrylamide gel containing a linearly increasing concentration gradient of denaturants (urea and formamide). At some point in the gradient, based on their stability (i.e. sequence), the PCR products denature along their length adjacent to the GC-clamp. This results in the formation of partially denatured molecules, which are essentially immobile. Since no further migration occurs, no increase in denaturant concentration is encountered and the GC-clamp portion of the molecule never dissociates. The end result is that PCR products having different sequences migrate to different positions in the gel and remain there, focused into discrete bands.

An example of DGGE separation is shown in Figure 2. GC-clamped PCR products from five different Octopus Spring cyanobacterial 16S rRNA genes were electrophoresed through a denaturing gradient gel for increasing lengths of time. By 3.5 hours, all the fragments have reached denaturing points where their migrations essentially halt. No further migration occurs even with prolonged electrophoresis. Once separated, the DGGE bands can be excised, reamplified and sequenced. Once a band has been identified by sequence, its unique position in the gel can subsequently be used as