Marine Biological Laboratory
Woods Hole, MA

From Early Biospheric Metabolisms to the Evolution of Complex Systems

Principal Investigator:

Mitchell Sogin sogin@mbl.edu

Co-Investigators/Collaborators:

At Marine Biological Laboratory:
Linda Amaral Zettler
Katrina Edwards
Rebecca Gast
Hilary Morrison
Lorraine Olendzenski
Bertil Olsson
David Patterson
Monica Riley
Stefan Sievert
Norman Wainwright
Jennifer Wernegreen

James Head, *Brown University*
John Mustard, *Brown University*
Andreas Teske, *University of North Carolina, Chapel Hill*

1.

PROPOSAL TITLE PAGE

From Early Biospheric Metabolisms to the Evolution of Complex Systems

Submitted by: The Marine Biological Laboratory at Woods Hole

Principal Investigator: **Mitchell L. Sogin**
Director of the Josephine Bay Paul Center for Comparative Molecular Biology and Evolution.
The Marine Biological Laboratory
7 MBL Street
Woods Hole MA 02543
508-289-7246

Subcontracting Institutions: **The Woods Hole Oceanographic Institution**
Wood Hole MA
02543

Brown University
Box 1862
Providence RI
02912

The University of North Carolina at Chapel Hill
Chapel Hill NC
440 W. Franklin Street, CB #1350
27599-3300

2.

TABLE OF CONTENTS

3. EXECUTIVE SUMMARY

Life flourishes within a thin veneer that corresponds to no more than 0.05% of Earth's diameter and it represents less than 10^{-10} of our global mass. Yet, through metabolic and biogeochemical processes, biology imposes an overwhelming force on planetary change. Whether biological systems similar to those on Earth ever occurred or continue to function on other planets or large satellites is unresolved. Given the imprint of life on Earth's geological record, the presence of comparable biological activity would have a profound effect on the landscapes and atmospheres of other solar system bodies. Recognizing the potential role of biotic systems in modifying planetary environments, astrobiological imperatives include delineating the evolutionary history of life, exploring the nature of early metabolic processes, and defining the limits of the "habitable zone". When and where did life originate? How did biological complexity develop from emergent properties of living organisms? Which metabolic processes were significant drivers in modifying early Earth? What are the limits of extreme environments compatible with life? Answers to these questions will have a direct impact on how to organize and target exploration in search of living organisms beyond our own biosphere.

We cannot be certain about when or how many times prebiotic chemistry crossed the threshold to a microbiological world. Impact histories (4.1-3.8 Ga) constrain the persistence of the earliest evolutionary lineages to the end of the period of intense bombardment although deep subsurface chemoautotrophs at kilometer depths could have survived even the largest impact events. Current paradigms trace the evolutionary history of all extant organisms to a common ancestor or population of ancestors that was exclusively microbial. Cyanobacterial-like fossils suggest that life is at least 3.45 Ga-old [1, 2], but the biogenic origins of these structures are under renewed scrutiny [3]. The chemical record documents prokaryotic metabolisms that may have existed 3.47-3.85 Ga [4, 5] and eukaryotic biosignatures that may be 2.7 Ga-old [6]. Yet, these interpretations are subject to artifacts that might be attributable to microbial contamination and they do not set absolute limits on the possible origins of life on Earth. Much later in Earth's history, the multicellular world emerged [7]. The paleontological record teaches us that microbial life can thrive in diverse environments devoid of multicellular organisms [8]. In contrast, the survival of macroscopic organisms is completely dependent upon the transformations afforded by complex microbial communities. If life exists beyond Earth, compelling arguments hypothesize such a world had its origins and continued manifestation in a microbial world.

Microorganisms of untold diversity are usually members of complex communities that dominate every corner of our biosphere. They orchestrate key processes in geochemical cycling, biodegradation and in the protection of entire ecosystems from major environmental shifts. Their effects can be global in scale. A single group of minute cyanobacteria, the *Synechoccoccus* / *Prochlorococcus* clade, with a planet-wide biomass on the order of one billion metric tons, is responsible for some 10 to 50% of the ocean's primary productivity [9]. Microbial carbon re-mineralization, with and without oxygen, maintains the carbon cycle. Microorganisms control global utilization of nitrogen through nitrogen fixation, nitrification, and nitrate reduction. They drive the bulk of sulfur, iron and manganese biogeochemical cycles [10]. Rarely studied bacterial mutualists provide essential nutrients and other compounds to diverse plant and animal hosts, and thus have a pervasive impact on the distribution, productivity and diversification of multicellular organisms. Processes and activities mediated by microbial populations have continuously imprinted the geological record, which provides a means to date major environmental shifts. Similarly, the genetic record offers insights about biological evolution, the invention of new metabolic capabilities, and increases in biological complexity that led to the emergence of the multicellular kingdoms of plants, animals and fungi.

The MBL/Woods Hole Astrobiology team seeks to understand patterns and mechanisms of genome evolution and metabolic variation that allowed diverse microorganisms to adapt to new environments, generate novel phenotypes, and evolve processes that led to environmental changes on a global scale, some of which can be detected through remote sensing. In addition to providing increased knowledge about evolutionary history of living organisms on Earth, these studies will provide a basis for interpreting telemetry data and developing life detection technology for planetary exploration. Our general strategy emphasizes the integration of molecular approaches to evolutionary biology with studies of metabolic activities in environments that are likely to reflect conditions on early Earth. The physiological and microbial diversity studies will generally focus on the hydrothermally altered sediments of Guaymas basin in the Gulf of California, the acidic, heavy metal laden Río Tinto of southwestern Spain, and isolates from the Juan de Fuca Ridge. Four major research themes (**EARTH'S EARLY BIOSPHERE; TERRESTRIAL ANALOGUES FOR EARLY MARS; BIOSIGNATURES AND LIFE DETECTION** and **EVOLUTION OF GENOME ARCHITECTURE IN PROKAROTES AND EUKARYOTES**) provide a framework for interdisciplinary research. Several of these projects represent new initiatives, while others are logical extensions of efforts supported by our initial membership in the NAI. We also describe an education and outreach program that capitalizes on our strengths in microbiology, molecular evolution and genome sciences.

EARTH'S EARLY BIOSPHERE: EVOLUTION OF MICROBIAL PHYSIOLOGIES. (A. Teske, S. Sievert). Classical, ribosomal (rRNA) phylogenies alone cannot inform us about the early evolution of microbial metabolisms and pathways. Through studies of genes for conserved enzymes with well-defined physiological roles within anaerobic and/or autotrophic pathways, we will explore the physiological history of early microbial life in the context of the early biogeochemical evolution of the biosphere. Individual projects include:

- Exploring the origin and diversification of ancient pathways: sulfate reduction and methanogenesis
- Determining the evolution and diversity of autotrophic CO_2 fixation pathways including reductive TCA pathways in anaerobic and extremophilic microorganisms in modern and paleo-environments.

TERRESTRIAL ANALOGUES FOR EARLY MARS: a) Iron Oxidation – Shaping past and present environments, b) Microbial diversity and population structure studies in the Río Tinto and **c) Life detection through remote sensing.** (L. Amaral Zettler, M. Sogin, K. Edwards, A. Teske, J. Mustard, J. Head). Banded iron formations (BIF) dating back to the Archean and the broad phyletic distribution of proteins with FeS centers indicate the importance of iron and Fe cycling for early life forms. Our astrobiology team will explore microbial Fe metabolism in prokaryotes from metal sulfide deposits at Juan de Fuca Ridge and determine microbial population structures in the acidic, Fe-rich environment of Río Tinto; both are potential analogues for an early, wetter Mars. An interdisciplinary team of biologists and geologists will incorporate new information from these studies into the design of life detection through remote sensing. We will also complete expression profiling experiments for oxygenic cyanobacteria in microbial mats with the goal of understanding the molecular basis of adaptation to extreme environments and application of these techniques to studies of microbial communities in extreme environments. Individual projects include:

- Investigating the fundamental biology of Fe oxidation in mesophilic and extremophilic microorganisms, and establishing the linkages to geochemical processes.
- Community microbial diversity surveys of the microbiota from Río Tinto based upon SAGT ("Serial Analysis Gene Tags", a newly developed technique for high-throughput analyses of rDNA sequences)
- Coupling diversity with biogeochemical data with emphasis on iron oxidation

- Use of minerals from iron-rich, acidic environments as biosignatures for remote sensing
- Expression profiling and the formation of biogeochemical gradients in cyanobacterial mats and in extreme environments

BIOSIGNATURES AND LIFE DETECTION: DETECTING ANCESTRAL PEPTIDES AND PROTEINS (M. Riley, N. Wainwright, R. Gast). Members of protein super families share ancestral conserved functional domains. We will adapt enzyme cascade reaction technology, which detects the presence of bacterial lipopolysaccharides, to develop a miniaturized instrument for the detection of ancestral peptides and nucleic acid sequences. Projects and focus areas include:

- Identifying appropriate key ancestral sequences in *E. coli* gene families
- Designing appropriate probe sequences, hybridization and detection assays
- Developing a miniaturized instrument for amplified life detection

EVOLUTION OF GENOME ARCHITECTURE IN PROKARYOTES: Genome-Genome Integration: Symbiosis, genetic assimilation, and evolutionary innovation. (J. Wernegreen). Beyond single point mutations, genome-genome interactions have been a driving force in the evolution of early life, the origin of eukaryotes and the evolution of multicellular complexity. These interactions foster lateral gene transfer between organisms and the remodeling of genome architecture. Using a genomic scale approach, we will study changes in genome structure and expression patterns for endosymbiotic associations between bacteria and eukaryotic hosts.

- Identifying changes in genome architecture, including lateral gene transfer, that catalyze endosymbiont-host interactions and contribute to their current diversity.
- Examining plasticity of gene expression in genome-genome interactions.

EDUCATIONAL AND PUBLIC OUTREACH PROJECTS AND ENRICHMENT OF THE ASTROBIOLOGY COMMUNITY (L. Olendzenski, D. Patterson, M. Sogin). This Astrobiology proposal will support courses and fellowships in ongoing MBL programs: *Workshop on Molecular Evolution, Advances in Genome Technology and Bioinformatics,* and *Living in the Microbial World*. We are in the final stages of forming a joint Brown-MBL graduate training and research program that will support 30-40 year round students at the MBL. When this new initiative begins (September, 2004), we anticipate accepting PhD students who will pursue astrobiology-relevant research topics in molecular ecology, evolutionary biology and genome science. Our continuing WEB site **micro*scope** will also develop tools for archiving and integrating image-rich data from habitats currently being studied by astrobiologists, including Río Tinto.

The institutional commitment includes dedicated laboratory space, capital equipment for laboratory research and facilities for theoretical analyses. The MBL, WHOI, Brown and UNC are committed to full participation in the virtual Astrobiology Institute. If funded, we intend to cooperate fully with other programs in Astrobiology in terms of sharing resources and new information gained from the proposed investigations. The management team will include PI M.L. Sogin (chair), Lorraine Olendzenski (EPO coordinator at MBL), and Co-Is from the collaborating institutions: Andreas Teske (UNC), Katrina Edwards (WHOI), and John Mustard (Brown University).

4. SUMMARY OF PERSONNEL

Personnel

Mitchell Sogin, PI	MBL
Linda Amaral Zettler, Co-I	MBL
Katrina Edwards, Co-I	WHOI
Rebecca Gast, Co-I	WHOI
James Head, Co-I	Brown
John Mustard, Co-I	Brown
Monica Riley, Co-I	MBL
Stefan Sievert, Co-I	WHOI
Andreas Teske, Co-I	UNC
Norman Wainwright, Co-I	MBL
Jennifer Wernegreen, Co-I	MBL
Lorraine Olendzenski, Co-I	MBL
David J. Patterson, Co-I	MBL

**** Institution Association**

MBL: Marine Biological Laboratory, Woods Hole, MA
WHOI: Woods Hole Oceanographic Institution, Woods Hole, MA
Brown: Brown University, Providence, RI
UNC: University of North Carolina at Chapel Hill

5. RESEARCH AND MANAGEMENT PLAN

This Astrobiology proposal has four primary themes. The first (**5.1**) **EARTH'S EARLY BIOSPHERE: EVOLUTION OF MICROBIAL PHYSIOLOGIES** explores patterns of early microbial physiologies as inferred by analysis of microbial populations, mostly from extreme environments. Technologies and samples are common threads for two projects that will explore sulfate reduction and methanogenesis, and ancient CO_2 fixation. The second theme (**5.2**) **TERRESTRIAL ANALOGUES FOR EARLY MARS: a) Iron Oxidation – Shaping past and present environments, b) Microbial diversity and population structure studies in the Río Tinto,** and **c) Life detection through remote sensing** focuses on isolates from metal sulfide deposits at Juan de Fuca ridge and the acidic, heavy metal environment Río Tinto of southwestern Spain where we will study iron oxidation, microbial diversity and population structures using a novel technique, SAGT (Serial Analysis Gene Tags). We will link this work with life detection studies based upon remote sensing technology. We also describe our progress in an ecogenomic-based study of expression profiling for cyanobacteria. The very same technology will be adapted to studies of microbial populations in extreme environments. The third theme (**5.3**) **BIOSIGNATURES AND LIFE DETECTION: DETECTING ANCESTRAL PEPTIDES AND PROTEINS** will take advantage of our expertise in computational biology and the rapidly expanding genome data bases to design life detection protocols that can either be used for life detection experiments on samples returned to Earth from other solar system bodies or as part of an instrumentation package for *in situ* analyses on other planets or large satellites. The final theme **(5.4) EVOLUTION OF GENOME ARCHITECTURE IN PROKARYOTES: Genome-Genome Integration: Symbiosis, genetic assimilation, and evolutionary innovation** addresses how symbioses influence evolution of genome architecture.

5.1 EARTH'S EARLY BIOSPHERE: EVOLUTION OF MICROBIAL PHYSIOLOGIES.

Life on early Earth developed within an anaerobic environment and evolved physiologies and metabolisms that were ancestral to modern-day sulfate-reducing, methanogenic, and methane-oxidizing microbes. Through molecular evolution studies, most commonly analyses of RNA constituents of the translation apparatus, we have gained important insights about the remarkable diversity of microbial populations in extreme environments where these metabolisms shape biogeochemical processes. These microbial ecosystems may resemble conditions of early Earth. The next chapters about the evolutionary history of early life will emerge from studies of metabolic diversity, distribution and the genetic adaptations of microbiota that participate in sulfate reduction, methanogenesis, methane oxidation, iron oxidation and processes that resembled ancient autotrophic carbon fixation pathways. Through the collaborative use of high-throughput DNA technology and high performance computing, we will explore the molecular evolution of coding regions for proteins that must have been of pivotal importance in the origins and evolution of these early metabolic pathways.

5.1a Origin and diversification of ancient pathways: sulfate reduction and methanogenesis (A. Teske).

Prokaryotic metabolic pathways underlie the biogeochemical cycles that affect geological history and planetary evolution. The anaerobic pathways of sulfate reduction, methanogenesis, and methane oxidation were predominant in ancient microbial communities; their isotopic imprints are pervasive in the carbon- and sulfur-isotopic record, from the present back to the Archean-Proterozoic transition [11]. Anaerobic microorganisms, their physiological interactions and geochemical activities are good model systems for early microbial ecosystems that thrived before the oxygenation of the Earth's biosphere, and for microbiota on planets that never evolved an oxygenated atmosphere or water column. The antiquity and evolutionary significance of

these microbial metabolisms are shown in the high degree of phylogenetic conservation of the key genes of sulfate reduction, methanogenesis, and anaerobic methane oxidation. By comparative sequence analysis of key genes, we can retrace the evolution of these anaerobic microbial pathways and enzymes, and the diversification of microbial life under the conditions of the early biosphere.

Isotopic evidence for widely expressed microbial sulfate reduction, in the form of ^{34}S-depleted sedimentary sulfides, goes back to the middle and early Proterozoic, 2.2 to 2.3 Ga. Only small isotopic fractionations (generally under 10% between sulfate and sedimentary sulfides) occur in sediments older than 2.5 to 2.7 Ga [12]. Laboratory experiments with sulfate-reducing bacterial populations show that, in order to suppress the isotopic fractionation signal of sulfate reduction, the sulfate level of the archean ocean must have been under 200 μM [13]. At this sulfate concentration, methanogenesis becomes the dominant anaerobic microbial process. Thus, it is very likely that methanogenesis predated the onset of sulfate reduction in the biogeochemical evolution of the Earth and dominated the archean biosphere; greenhouse warming of the archean Earth by methanogenesis could have compensated for the lower luminosity of the "faint young sun" [13]. The carbon isotopic imprint of biomass produced by methanogenesis, in the form of highly ^{13}C-depleted kerogen ($\delta^{13}C \bullet$ -60%), is found in late Archean and 2.8 Ga-old Proterozoic kerogens [14]. This isotopic record was originally interpreted as evidence for biomass accumulation by aerobic, bacterial oxidation of biogenic methane [15]. Anaerobic oxidation of biogenic methane and subsequent incorporation into microbial biomass is a more likely explanation, since evidence for the stepwise and pervasive oxygenation of the proterozoic biosphere appears much later, ca. 2.2 Ga [16].

To search for deeply-branching and (possibly) ancestral representatives of sulfate-reducing, methanogenic and methane-oxidizing microorganisms and their key genes in modern environments, we focus on hydrothermal vents and deep subsurface habitats. Hydrothermal vents are some of the earliest and best-protected microbial habitats. They may have survived repeated impact frustration of early life and might occur on other planets with oceans and active plate tectonics, volcanism or tidal heating. On Earth, hydrothermal vents sustain complex microbial ecosystems that utilize inorganic energy sources, such as sulfide, hydrogen, and reduced metals, and geothermal sources of carbon, such as methane, CO_2 and geothermally synthesized low-molecular weight organic compounds. The deep marine and terrestrial subsurface is a vast microbial habitat that houses pervasive microbial life, especially anaerobic bacteria and archaea. Fermentative and sulfate reducing bacteria as well as archaea, and methanogens have been detected down to at least several hundred meters of oceanic sediments [17], and microbial texture alteration has been detected for the basaltic ocean crust [18].

To identify the bacteria and archaea that constitute these diverse methanogenic, sulfate-reducing and methane-oxidizing prokaryotes in hydrothermal vents and in marine deep subsurface habitats, analyses of rRNA sequences are complemented with sequencing surveys of highly conserved marker genes for sulfate reduction and methanogenesis. In this way, it is possible to focus a genomic survey on these specific classes of microorganisms. Screening a microbial community for the highly conserved key genes of these processes, the dissimilatory sulfite reductase (*dsr*) and coenzyme M methyl reductase (*cmr*) genes, results in a diversity census that shows the composition and complexity of this microbial community. At the same time, a survey of these genes in environmental samples shows the evolutionary divergence that has accumulated in these genes since the early Proterozoic or the Archean, and the phylogenetic depth of these metabolisms in the bacterial and archaeal tree of life. With a growing database, homologous and ancestral traits of these genes can be defined that can be significant for an understanding of enzyme evolution and its functional constraints.

Sulfate reduction and *dsrAB* genes. The key gene of dissimilatory sulfate reduction, (*dsrAB*), codes for the alpha and beta subunits of the enzyme dissimilatory sulfite reductase, which catalyses the reduction of sulfite to sulfide, after previous activation and reduction of sulfate to sulfite. Consistent with an early origin of the sulfate-reducing pathway, the *dsrAB* genes are phylogenetically conserved in several deeply branching bacterial and archaeal phyla, and to some extent consistent with 16S rRNA phylogenies [19]. With a growing database of *dsr* genes, 16S rRNA and *dsr* phylogenies turned out to be discordant for some sulfate reducers. For example, the dsr genes of thermophilic members of gram-positive genus *Desulfotomaculum*, and of the thermophilic genus *Thermodesulfobacterium* are apparently related to those of delta-Proteobacterial sulfate reducers, suggesting lateral gene transfer [20]. Since these discordances are known, and limited to a few cases, the phylogeny of the *dsrAB* genes allows a simultaneous phylogenetic and metabolic identification of evolutionary divergent sulfate-reducing microorganisms [19, 20]. The α and β subunits of the *dsr* gene are most likely the results of an early gene duplication, as shown by considerable homology among the two subunits. The homologous subunits include sequence regions harboring the catalytic center that consist of a $[Fe_4S_4]$-cluster coupled to a siroheme prosthetic group [21]. Paralogous gene trees of the two subunits can be used to identify deeply-branching *dsr* genes. At present, the deepest lineage seems to be the thermophilic bacterium *Thermodesulfovibrio yellowstonii,* and not the archaeal genus *Archaeoglobus* [20], which would argue for a bacterial origin of sulfate reduction. More deeply branching dsrAB sequences of ancestral sulfate reducers could be isolated from habitats that are known to harbor a great diversity of sulfate reducers.

We will focus on hydrothermal vents and the deep subsurface, as promising candidate environments for unusual and ancestral sulfate reducers. For example, the Guaymas Basin in the Gulf of California, the northernmost extension of the East Pacific Rise mid-ocean ridge, is covered by organic-rich, hydrothermally altered sediments that show very high sulfate reduction rates over wide temperature ranges [22-25]. The Guaymas sediments harbor highly diverse sulfate-reducing bacterial and archaeal populations, including novel, mesophilic and thermophilic fatty acid oxidizers[26], hyperthermophilic archaea of the genus *Archaeoglobus* [27], and members of the propionate-oxidizing, acetate-producing family *Desulfobulbaceae* [28]. Our recent *dsrAB* gene survey in Guaymas sediments detected *dsrAB* genes related to well-known proteobacterial sulfate-reducing genera (*Desulfobacter* and *Desulfobacterium*), but also a novel, deeply-branching *dsrAB* genes that are unrelated to *dsrAB* genes of any cultured sulfate-reducing microorganism [29]. A phylogenetic examination of the two paralogous *dsrA* and *dsrB* subunits of these deeply-branching Guaymas genes has started. We plan to extend this search for novel, deeply-branching sulfate reducers and their *dsrAB* genes to other hydrothermal and deep subsurface environmental samples, including deep subsurface samples obtained from the Ocean Drilling Program (leg 201, Equatorial Pacific and Peru Margin), microbial mats from Guaymas, sulfidic hydrothermal vent chimney rock, and geothermal fluids from basaltic ocean crust at the flanks of mid-ocean ridges.

Methanogenesis, methane oxidation, and *mrcA* genes. Coenzyme M methyl reductase (*cmr*) is the key enzyme of methanogenesis, which catalyses the terminal and highly exergonic step of the methanogenesis pathway, the reduction and release of the coenzyme-M bound methyl group as free methane. The *cmr* gene occurs exclusively in methanogenic archaea; in contrast to other genes of methanogenic pathways, it is not known to be affected by lateral gene transfer to other prokaryotes, for example aerobic, methanotrophic bacteria [30]. The α subunit (*mrcA*) of the Coenzyme M methyl reductase gene is highly conserved, accessible by PCR, and generally consistent with 16S rRNA phylogenies [31-33]. Thus, the major lineages of methanogenic archaea can be identified on the basis of their *mrcA* genes.

At present, it is not known whether *mrcA* genes also participate in the anaerobic, sulfate-dependent oxidation of methane to CO_2 [34], which is carried out by members of at least two

phylogenetic lineages of recently identified archaea, ANME-1 and ANME-2 [35-38]. Anaerobic methane oxidation by archaea must either proceed through an energy-intensive reversal of this process using *cmr*, or bypass this step completely and have no need for this enzyme [39]. Screening the environmental DNA samples from sites with known anaerobic methane-oxidizing activity by cloning and sequencing will help to decide two different possibilities: Coenzyme M methyl reductase may be restricted to phylotypes of well-known, cultured methanogenic archaea, and is not involved in anaerobic methane oxidation. Alternatively, novel phylogenetic lineages of the *mrcA* gene that do not match known methanogens, and are found specifically in samples with methane-oxidizing activity or 16S rRNA signatures, could play a role in anaerobic methane oxidation. Current genome sequencing studies of purified ANME-1 and ANME-2 archaea from environmental samples are testing whether these methanotrophic archaea carry *mcr* genes [38].

The search for novel *mrcA* genes will focus on anaerobic environments that harbor diverse methanogens, and methane-oxidizing archaeal populations. Guaymas sediments are a promising target; their methanogenic populations include hyperthermophilic, autotrophic methanogens of the genus *Methanococcus* [40, 41] and the deeply branching *Methanopyrus* [42], as well as members of the formate-utilizing, mesophilic or moderately thermophilic family Methanomicrobiales [28]. Methane produced by these diverse methanogenic communities combines with the methane pool originating from pyrolysis of organic matter buried in the Guaymas sediments; the resulting methane concentrations in the Guaymas vent fluids are orders of magnitude higher than at non-sedimented, bare lava vent sites [43]. In this geochemical setting, anaerobic methanotrophs thrive; 16S rRNA analysis and ^{13}C-isotopic analysis of taxonomically informative archaeal lipids has shown that the anaerobic methanotrophic communities in the Guaymas Basin include both ANME-1 and ANME-2 archaea [28].

Methane cycling and sulfate reduction in the deep marine subsurface. Large and very valuable sample sets for this project have become available through collaboration with the Astrobiology group at the University of Rhode Island, led by Stephen D'Hondt. The deep marine subsurface is an ancient microbial habitat of global dimensions that harbors a large proportion of the prokaryotic biomass of the planet; its physiological and evolutionary biodiversity is just beginning to be explored [17]. Marine deep subsurface sediments differ in many respects from hydrothermal vents. They are characterized by low temperatures; the microbial populations are several orders of magnitude less dense and active; suitable electron donors and carbon sources for microbial metabolism are scarce; the geochemical and thermal gradients that are compressed into a few cm at vents are extended over several 100 m [17]. These extended porewater sulfate and methane gradients strongly suggest that sulfate-reducing and methane-cycling microbial populations (analogous to Guaymas) are dominant community members that determine the overall biogeochemical activity patterns in the deep subsurface [44]. Deep subsurface sediments were collected through the Ocean Drilling Program (ODP) cruise Leg 201 in the Equatorial Pacific and on the Peru Margin (ODP sites 1225 – 1231; see preliminary cruise report on www-odp.tamu.edu/publications/prelim/201-prel/201toc.html). For each site, detailed geochemical profiles (sulfate, methane, reduced metals, hydrogen) and cell counts are available; process rates (methanogenesis, sulfate reduction, hydrogen and acetate turnover) are being determined by different cruise participants. The geochemical profiles show clear zones with sulfate reduction, methanogenesis, and the sulfate reduction/methanogenesis transition layers. From each site, selected samples from the methane maximum in the methanogenic zone, from the anaerobic methane oxidation layer, and from the sulfate reduction zone will be examined. First priority is given to sites 1229 and 1230. Site 1229 is located on the Peru Margin at 150 m depth (extending to 200 mbsf) and shows a unique feature, a deep subsurface incursion of sulfate-rich evaporite brines that replenishes sulfate at ca. 90 m depth and thus creates an additional sulfate reduction/methanogenesis transition zone, in addition to the near-surface (ca. 36-38 mbsf) transition [45]. The organic-rich site 1230 is located in 5086 m

water depth in the Peru Trench (extending down to 280 mbsf), and shows a very steep gradient from the sulfate-reducing to the methanogenic layers which was sampled in high resolution. In contrast to the high-organic rich Peru Margin and Peru Trench sites (1227, 1228, 1229, 1230), the organic-poor Central Pacific sites are characterized by low process rates, lower bacterial numbers, and highly extended geochemical gradients. The coexistence of methane and sulfate over several 100 m of sediment suggests slow, concomitant sulfate reduction and methanogenesis in this area (sites 1225, 1226, 1231), contrary to our understanding of the sequential order of microbial redox reactions according to energy yield [46].

In general, the survey for *dsr* and *cmr* marker genes in these subsurface samples has a dual focus. Besides the evolutionary implications of finding novel and deeply branching marker genes for sulfate reduction and methanogenesis (as spelled out for Guaymas), the precisely defined geochemical gradients and process rates will allow the identification of geochemical controls (organic carbon input, *in situ* carbon substrate and hydrogen concentrations) that govern the composition of the sulfate-reducing, methanogenic, and methanotrophic communities. During the course of this project, additional samples will become available, from anaerobic methane-oxidizing sediments of the Hydrate Ridge offshore Oregon (ODP leg 204), deeply buried cretaceous black shales from the Demarara Rise (ODP leg 207), and off-axis ridge flank subsurface fluids made accessible by a corked borehole. These subsurface samples will be investigated in collaboration with the NAI group at URI, which includes biomarker studies (compound-specific isotopic analyses), process rate measurements, and geochemical analyses.

METHODS. In spite of continuing primer development [20], it is not certain whether PCR-based approaches will reliably detect all environmental *dsrAB* and *mrcA* genes of interest, including the most ancestral and deeply-branching lineages that are of particular interest to Astrobiology. New versions of these key genes with non-conserved primer sites would be missed. Primer site conservation can never be taken for granted, as shown by the finding that even 16S rRNA sequence motifs that were regarded as universally conserved showed substantial variation between different bacterial lineages [47]. Therefore, the survey of *dsrAB* and *mrcA* genes will be complemented by alternative approaches.

A possible method is gene cassette PCR [48]. Genes that have been acquired by horizontal gene transfer in bacteria are flanked by integrons, 59-base pair sequence motifs with highly conserved inverted repeats at both ends. The units of DNA captured by integrons, gene cassettes, can accumulate in microbial genomes; multiple insertion events lead to the formation of multicassette arrays. By combining a forward primer for a key gene sequence motif with a reverse primer targeting a neighboring integron inverted repeat, it is possible to circumvent the requirement for two conserved target sites within a key gene, to broaden the amplification range of existing primers, and to sequence terminal regions of genes that would otherwise remain inaccessible [48]. The downside of this approach is that it will selectively detect and amplify genes that have undergone at least one integron-catalyzed gene transfer event.

Shotgun cloning and fosmid library construction for deep subsurface bacteria and archaea could circumvent PCR entirely. To isolate high-quality, non-degraded DNA from deep subsurface sediment material, we will start with established approaches for genomic library construction from soil that include gentle detachment of prokaryotic cells from particles as the first step for library construction [49]. Practical experience with marine environmental samples indicates that fosmid libraries are easier to generate and to screen than BAC libraries, in particular for highly complex communities (California coastal sea water and Antarctic picoplankton; [50]). Inserts that contain phylogenetic anchors can be identified using a multiplex PCR-denaturing gradient gel electrophoresis approach [51]. Just as rRNA sequences can describe phylogenetic context when present in larger DNA fragments, key genes of known function and distribution can serve as “phylogenetic anchors” to identify the source of large

DNA inserts cloned from natural poulations of bacterio- and archaeoplankton [52]. Novel genomic information is not limited to *dsrAB* and *mrcA* genes, but may include functional and structural genes that correspond to physiological demands and environmental stresses of deep subsurface and hydrothermal vent habitats, and allow a better understanding of microbial survival strategies and adaptations to these ancient habitats.

Extensive surveys of ancestral and extreme microbial habitats (hydrothermal vents, deep subsurface environments) for the genetic markers of sulfate reduction, methanogenesis and methane oxidation, will yield new insights into the evolution of these pathways and their key genes. The environmental diversity of key gene sequences and amino acid sequences, their secondary structure motifs and conserved sites, contains information about environmental and functional constraints that have shaped the evolution of these ancient microbial pathways. Deep subsurface and hydrothermal vent environments are good places to start such a survey, for two reasons: They mirror several characteristics of the early Earth, the absence of oxygen, and the necessity for protection against an inhospitable surface environment characterized by cosmic radiation, inclement weather and meteorite bombardments. Further, they provide good analogs to subsurface life under extraterrestrial conditions where a permanently inhospitable surface environment might have kept microbial life underground, and never allowed its evolution towards complex, multicellular life within a phototrophic, oxygenated biosphere [53].

5.1b The evolution and diversity of ancient CO_2-fixation pathways in anaerobic and extremophilic microorganisms: Clues to the early evolution of life on Earth (S. Sievert).

Despite research advances in studies of microbial diversity, life in extreme environments and prebiotic chemistry, the origin of life on Earth remains an enigma. Key unresolved questions include: When did life first appear? What was the nature of the first organisms? and Did early life have a heterotrophic or an autotrophic origin? [54, 55] Carbon isotope data suggest that autotrophic organisms, which preferentially fix isotopically light carbon, may have existed >3.7 Ga ago [56-58], indicating that autotrophy was one of the earliest evolved metabolisms on Earth. Many consider the reductive TCA cycle to be the most ancient autotrophic carbon fixation pathway [59-62]. This cycle occurs in phylogenetically-diverse prokaryotes whose growth conditions fit well with our view of the primitive Earth. One of the two key enzymes of this pathway is ATP citrate lyase (ACL). Limited gene sequence data are available for this enzyme and we have no information about the diversity, distribution and importance of organisms using the reductive TCA cycle in nature. This information is required in order to understand the carbon cycle on our planet today and in the geological past, and the evolution of this ancient carbon fixation pathway.

The objective of this project is to examine the occurrence of the ACL gene in organisms where this activity is known to be present, and whether it occurs in other prokaryotic autotrophs where information about CO_2 fixation pathways is lacking. Newly-acquired sequence data will be used to explore phylogenetic patterns generated by ACL for insights into the evolution of this ancient pathway on Earth. This represents the first step in developing a better understanding of the environmental significance of this potentially important metabolic pathway.

CO_2 fixation and the early evolution of life on Earth. An autotrophic origin of life posits that metabolism evolved before all else [59, 60, 62-65]. Recent data suggest that the atmosphere of early Earth was weakly reduced and contained mainly carbon dioxide, nitrogen, and only trace amounts of CO, H_2, methane, reduced sulfur gasses, and possibly O_2, making it an unlikely environment for the formation of a prebiotic broth that might have fueled the growth of heterotrophic organisms [55, 66, 67]. However, the possibility exists that increased concentrations of organic molecules appeared in evaporating ponds [54] or that selective absorption to minerals led to the concentration and polymerization of active monomers [68]. In

both cases, organic molecules might have been synthesized abiotically on Earth or delivered by meteorites [64]. The first life forms may have been thermophiles [69]. Temperatures on the surface of the early Earth probably had cooled to 100˚C or less by 4.4 Ga [70]. Although these surface temperatures may have existed for only 20 My [71] (a potentially longer time frame for the subsurface), they could have provided suitable conditions for the evolution of thermophilic autotrophs (organisms that use inorganic carbon as the sole source of carbon) [55, 64, 69, 72]. Phylogenetic analyses of 16S rRNA and whole genomes place autotrophic hyperthermophiles at the base of the tree [69, 73-75], although this view has been challenged [76, 77]. Submarine hydrothermal vents might have been sites for the origin of (thermophilic) life or provided refugia for thermophiles to survive the heavy bombardment by meteorites prior to 3.8 Ga, which frequently sterilized the surface of the Earth and vaporized large parts of the ocean [54, 78, 79].

Today, representatives that are able to grow autotrophically exist in almost all major groups of prokaryotes [80]. These organisms play essential roles in ecosystems by providing a continuous supply of organic carbon for heterotrophs. The Calvin cycle may be the most important extant autotrophic carbon fixation pathway [81], but despite its global significance, it has a rather narrow phylogenetic distribution (**Figure 1**). It does not occur in Archaea or

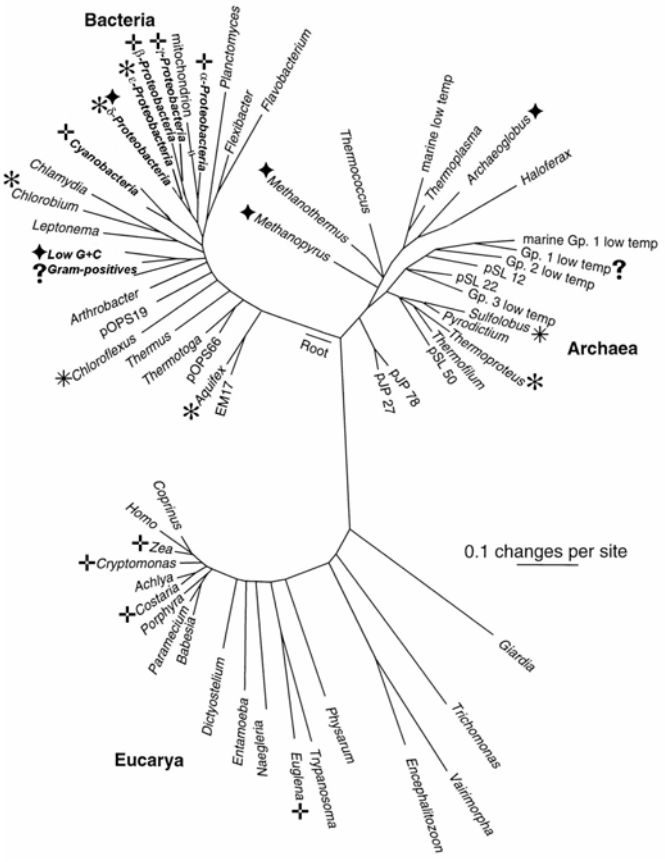

Figure. 1. Universal tree based on 16S rRNA sequences showing the phylogenetic distribution of the different carbon fixation pathways. = Calvin cycle; = reductive TCA cycle; = Hydroxypropionate pathway; = reductive acetyl-CoA pathway **?** = unknown

hyperthermophilic organisms and among the Bacteria has only been identified within the α, β, and γ-Proteobacteria and cyanobacteria, from which the chloroplasts of plants are derived [80]. This phylogenetic distribution argues against the Calvin cycle as being an ancient carbon fixation pathway. Furthermore, microorganisms present in extreme environments that resemble

conditions on the early Earth, e.g., high temperature, anaerobic or acidic conditions, utilize different CO_2-fixation pathways [80, 82]. At present there are three alternative pathways known [80], the reductive tricarboxylic acid cycle, the reductive acetyl-CoA pathway, and the hydroxypropionate pathway. Although fewer organisms are known in which these pathways operate compared to those that utilize the Calvin cycle, the phylogenetic distribution of these alternative pathways is much broader, as they have been identified in several major lines of descent in the bacterial and archaeal domains (**Figure 1**; **Table 1**). In addition, the fact that these pathways occur mainly in anaerobes and (hyper)thermophiles [82-88] supports the idea that these carbon fixation pathways are ancient and were of much greater importance in the geological past before the rise of oxygenic phototrophs about 2.3 Ga, or during large anoxic events such as during the Cretaceous [64, 89-91]. These lines of evidence suggest that if microorganisms exist on other planetary bodies with reducing environments, these pathways are likely to also be present.

Table 1. The phylogenetic distribution of different CO_2-fixation pathways.

		CO_2-Fixation pathway			
Phylogenetic group	**Representative species**	**Calvin**	**Reductive TCA**	**Reductive Acetyl-CoA**	**Hydroxy-propionate**
Proteobacteria					
α	*Rhodosprillum rubrum*	+	-	-	-
β	*Rhodocyclus tenuis*	+	-	-	-
γ	*Thiomicrospira crunogena*	+	-	-	-
δ	*Desulfobacter hydrogenophilus*	-	+	-	-
	Desulfobacterium autotrophicum	-	-	+	-
ε	*Candidatus* Arcobacter sulfidicus	-	+	-	-
Cyanobacteria	*Synechococcus* sp.	+	-	-	-
Chloroflexus	*Chloroflexus aurantiacus*	-	-	-	+
Low G+C Gram-positives	*Chlostridium aceticum*	-	-	+	-
	Ammonifex degensii	?	?	?	?
Chlorobium	*Chlorobium tepidum*	-	+	-	-
Aquificales	*Aquifex pyrophilus*	-	+	-	-
Thermoproteales	*Thermoproteus neutrophilus*	-	+	-	-
Sulfolobales	*Metallosphaera sedulans*	-	-	-	+
Pelagic *Crenarchaeota*	Uncultivated	-	?	-	?
Archaeoglobus	*Archaeoglobus lithotrophicus*	-	-	+	-
Methanogens	*Methanocaldococcus jannaschii*	-	-	+	-
Eukaryotes	Chloroplast containing organisms	+	?	-	-

Reductive TCA cycle. It has been proposed that the first autotrophic pathway was akin to the reductive tricarboxylic acid cycle (TCA) operating in extant autotrophs [59-62]. It has the characteristics of an autocatalytic cycle and leads to a complex cyclic reaction network from which other anabolic pathways could have evolved [59, 60]. The reductive TCA cycle may have preceded the oxidative TCA cycle [61, 92], and based upon biochemical and isotopic analyses appears to operate in phylogenetically diverse autotrophic bacteria and archaea, including anoxic phototrophic bacteria (*Chlorobiaceae*) [93-95], sulfate-reducing bacteria (*Desulfobacter hydrogenophilus*) [83], microaerophilic, hyperthermophilic hydrogen-oxidizing bacteria (*Aquificales*) [84, 85], and hyperthermophilic sulfur-reducing crenarchaeota (*Thermoproteales*) [84, 96]. The two latter groups are important in biogeochemical processes in extant hydrothermal habitats. Hydrothermal systems have prevailed throughout Earth's history and they might resemble sites where life originated. They also may be analogs for extraterrestrial environments where active volcanism occurs [59, 64, 97-99]. Preliminary evidence also suggests that the reductive TCA cycle is operating in *Candidatus* Arcobacter sulfidicus, a sulfide-oxidizer belonging to the ε-proteobacteria that has the unique capability to form sulfur in

filamentous form [100, 101]. Members of the ε-proteobacteria appear to be a major component of microbial communities at deep-sea vents and in the subseafloor [28, 102].

ATP citrate lyase in Prokaryotes. The reductive TCA cycle is essentially the TCA cycle running in reverse, leading to the fixation of 3 molecules of CO_2 and the production of one molecule of triose phosphate. Most of the enzymes between the two pathways are shared, with the exception of two key enzymes that allow the cycle to run in reverse: ATP citrate lyase (ACL) and 2-oxoglutarate:ferredoxin oxidoreductase (OOR). OOR catalyzes the carboxylation of succinyl-CoA to 2-oxoglutarate whereas ACL catalyzes the ATP dependent cleavage of citrate to acetyl-CoA and oxaloacetate, and its presence in prokaryotes is indicative of a functioning reductive TCA cycle. Although several prokaryotes have been shown via specific enzymatic assays to possess ACL activity, the only cultured prokaryotes for which a complete gene encoding ACL has been identified to date are the green-sulfur bacteria *Chlorobium tepidum* and *C. limicola* [103, 104]. The heteromeric ACL of *Chlorobium* is encoded by two separate genes, *acl*A and *acl*B, which are adjacent to one another [104, 105]. The same gene arrangement occurs in eukaryotes such as fungi and plants, whereas in animals a homomeric ACL is encoded by only one gene [104, 105]. The *acl*BA of *Chlorobium*, fungi, and plants show similarity to the N- and C-terminal halves of the ACL from animals, and it has been suggested that ACL in animals is a product of a gene fusion event that occurred early in the evolution of that kingdom [103]. It has been proposed that aclB may have evolved through gene duplication and diversification from the β-subunit of succinyl-CoA synthetase (SCS), whereas aclA may have arisen from a fusion event between the α -subunit of SCS and citrate synthase (CS) after a previous gene duplication and diversification [103, 105] (**Figure 2**). By using primers based on the conserved regions of the gene, a partial ACL sequence from *Candidatus* Arcobacter sulfidicus

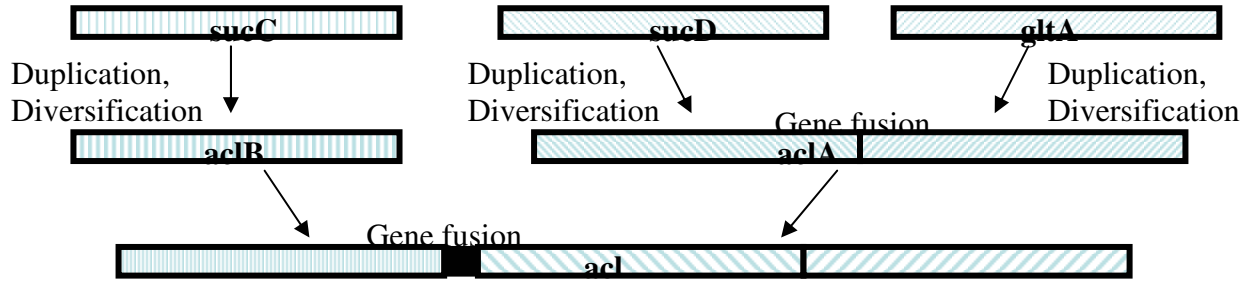

Figure 2. Schematic representation of the possible evolutionary history of ATP citrate lyase. sucC = gene encoding β-subunit of Succinyl-CoA synthetase (SCS); sucD = gene encoding α-subunit of SCS; gltA = gene encoding for citrate synthase; aclB = gene encoding ATP citrate lyase β-subunit; aclA = gene encoding ATP citrate lyase α-subunit; acl = one gene encoding for homomeric ATP citrate lyase in animals. See text for further explanations. Modified after Fatland et al. [103].

was amplified by our laboratory. Amplification products of the expected size were also obtained from *Desulfobacter hydrogenophilus*, but have not yet been sequenced [106]. Recently, two ATP citrate lyase genes were identified on two fosmids also containing the 16S rRNA of the main ε-proteobacterial epibionts of *Alvinella pompejana* [107], a polychaete living on sulfide structures at deep-sea hydrothermal vents on the East Pacific Rise. On these fosmids, the genes corresponding to *acl*A and *acl*B were adjacent to each other in an arrangement identical to that of *Chlorobium*. An ATP citrate lyase has also been purified from the hyperthermophilic hydrogen-oxidizing bacterium *Hydrogenobacter thermophilus*, which is a member of the *Aquificales*, however, no sequence information is available [108]. Although the whole genome of *Aquifex*

aeolicus has been determined [109], no ACL genes similar to the one in *Chlorobium* could be identified in this organism despite detected ATP citrate lyase activity in the closely-related *A. pyrophilus* [84]. The presence of a highly divergent ACL gene in *A. aeolicus* [110], or different metabolic properties of the two *Aquifex* species might explain these observations. Low sequence similarity with *Chlorobium* ACL would be in line with *Aquifex* being a deeply-branching bacterium. Recently, it has also been determined that the genome of *A. aeolicus* does not contain a suitable candidate for OOR and that previous reports indicating its presence are based on an incorrect annotation [109, 110]. However, a partial ACL sequence was recently obtained from *Persephonella marina*, [111] which is also a member of the *Aquificales* and is thought to grow autotrophically by using the reductive TCA cycle [112, 113].

ATP citrate lyase in Eukaryotes. Aside from the above-mentioned prokaryotes, ACL genes occur in yeasts, fungi, green algae, plants, animals, and in photosynthetic protists such as *Chlamydomonas reinhardtii* and the glaucophyte *Cyanophora paradoxa* [114-119]. Phylogenetic analysis places the ACL of the green sulfur bacterium *Chlorobium* basal to the eukaryotic homologues [103], suggesting that eukaryotes derived their ACL from bacteria. Interestingly, phylogenetic trees based on available ACL sequences are congruent with the evolutionary history depicted by rRNA [103], suggesting an ancient origin of this enzyme. In the eukaryotes studied so far, ACL is not involved in carbon fixation or carbon oxidation, but functions in the provision of cytosolic acetyl-CoA for biosynthesis of fats and cholesterols [103, 116], although it has been suggested that the reductive TCA cycle might operate in *Chlamydomonas reinhardtii* [120]. The best candidates for deep-branching eukaryotes include *Giardia* and *Trichomonas* but their nearly completed genomes do not appear to code for ACL. While these parasites might depend on their host for fatty acid synthesis, ACL may be discovered in free-living relatives. Acetyl-CoA, which is synthesized in the mitochondria, must be transported across the mitochondrial membrane to the cytosol where it is used for fatty acid synthesis. Because membranes are impermeable to acetyl-CoA, it must first be converted to citrate, which crosses the membrane. Citrate is then cleaved by ACL into acetyl-CoA and oxaloacetate. Amitochondriate eukaryotes such as *Giardia* and *Trichomonas* may have lost this gene since acetyl-CoA production and fatty acid synthesis are not separated in different compartments. ACL may have initially served to make carbon fixed from CO_2 available for general metabolism, and later this function might have switched with the development of respiratory capabilities, so that ACL tapped into the TCA cycle for carbon derived from catabolic reactions [103].

METHODS. Despite the fact that organisms utilizing CO_2-fixation pathways other than the Calvin cycle might contribute significantly to primary production in their respective habitats, few data exist on the diversity, distribution and importance of the organisms using these pathways in nature. This information is necessary to understand the carbon cycle on our planet today and in the geological past. We propose to focus our studies on the molecular evolution and ecological importance of the reductive TCA cycle, which is generally considered the most ancient autotrophic carbon fixation pathway [59-62].

Sequence diversity of ACL. We will investigate the occurrence of ACL coding regions in organisms known to utilize the reductive TCA cycle. Because of a lack of sequence information for ACL, it will first be necessary to identify the genes coding for the observed ATP citrate lyase activity in the different organisms known to have, or suspected to have, the reductive TCA cycle, including: *Aquifex pyrophilus, Desulfobacter hydrogenophilus, Thermoproteus neutrophilus, Pyrobaculum islandicum, Hydrogenobacter thermophilus, Desulfurobacterium, Persephonella* and *Thermovibrio*. All organisms will be obtained from existing culture collections, cultivated, and harvested for DNA and RNA for later analyses. Biomass will also be frozen in liquid N_2 for later biochemical analysis, e.g., assays diagnostic for ATP citrate lyase activity in cases where

activity is suspected but not yet demonstrated [84, 86, 104]. In addition to biochemical assays, RT-PCR will be used to confirm the expression of ACL.

A database of ACL sequences will be constructed by PCR amplification, cloning and sequencing of ACL genes using existing and newly-developed primers based on conserved regions of the gene. Degenerate primers have been designed by our laboratory based on an alignment of known ACL genes. The G+C content was kept between 40 and 60% and the total degeneracy was kept below 64-fold. The primer length was 18-26 nucleotides, Gs and Cs were preferred at the 3' end to enhance binding, and the G+C content was balanced over the length of the primer. We have used different combinations of primers and tested them under various conditions (temperature gradient, Mg concentration etc.). The primer pairs that gave positive results are depicted in **Figure 3** along with their binding positions and sequence information (aclA-F2/aclA-R5, and aclA-F3/aclA-R4). These primer pairs were designed to amplify a roughly 800 bp long fragment from *acl*A that spans the junction between the N-terminus

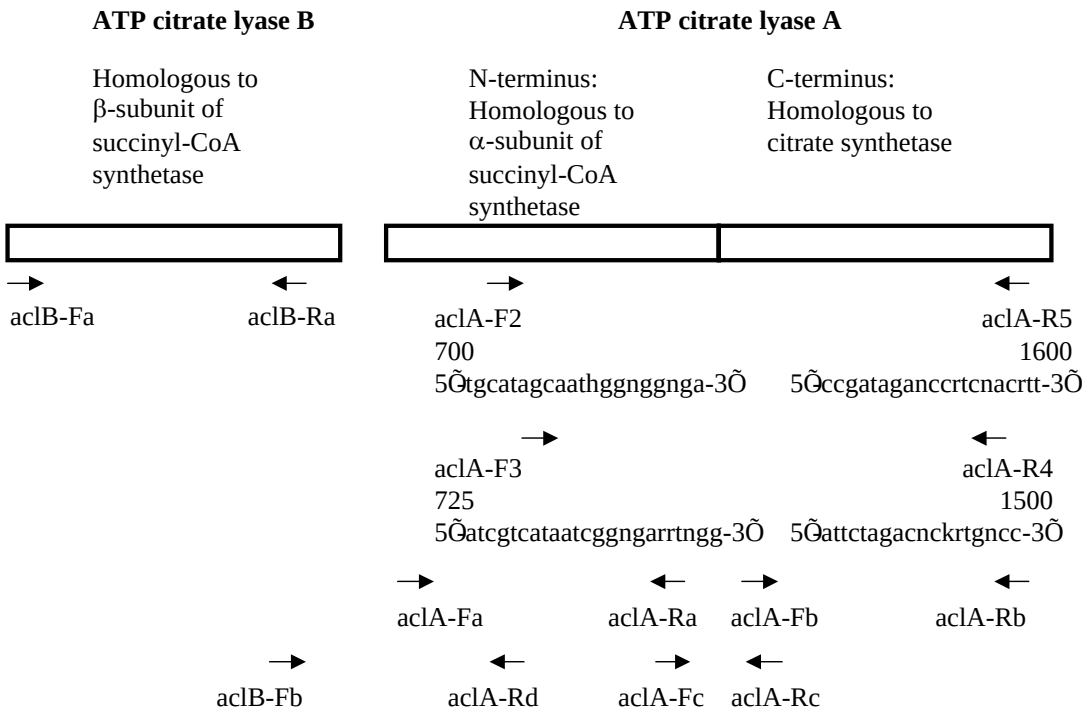

Figure 3. Domain structure of the ATP citrate lyase gene (aclBA) of *Chlorobium* sp., plants, and fungi. aclB and aclA encode for the β- and α-subunit of the ATP citrate lyase, respectively. In *Chlorobium tepidum* both genes are adjacent to each other. Depicted are also the primers with sequence and binding positions that were used to amplify a ~800 bp fragment of aclA from *Candidatus* Arcobacter sulfidicus that spans the region between the domain most similar to the succinyl-CoA synthetase and the domain most similar to citrate synthase (aclA-F2/aclA-R5; aclA-F3/aclA-R4). In addition, primers are depicted that will be used to amplify different regions of the gene(s). See text for further detail.

(homologous to α-SCS) and the C-terminus (homologous to CS). Using these primers we were successful in amplifying an *acl*A fragment from *Candidatus* Arcobacter sulfidicus. A positive amplification result with these primers for other organisms will allow us to confirm the configuration of *acl*A according to the evolutionary scheme in **Figure 3.** A negative amplification result, however, could be due to sequence divergence, and would not necessarily indicate the absence of this gene. We will employ a more rigorous PCR strategy that will first amplify each fragment (aclA-Fa/aclA-Ra, aclA-Fb/aclA-Rb), and then use custom, species-specific primers (designed from sequences for each fragment) that should generate an amplicon

that spans the junction between the N- and C-terminus regions (aclA-Fc/aclA-Rc). We will also design primers for *acl*B (aclB-Fa/aclB-Ra) and primers that should amplify the C-terminus of *acl*B together with the N-terminus of *acl*A if both genes are adjacent to each other (aclB-Fb/aclA-Rd) (see **Figure 3**).

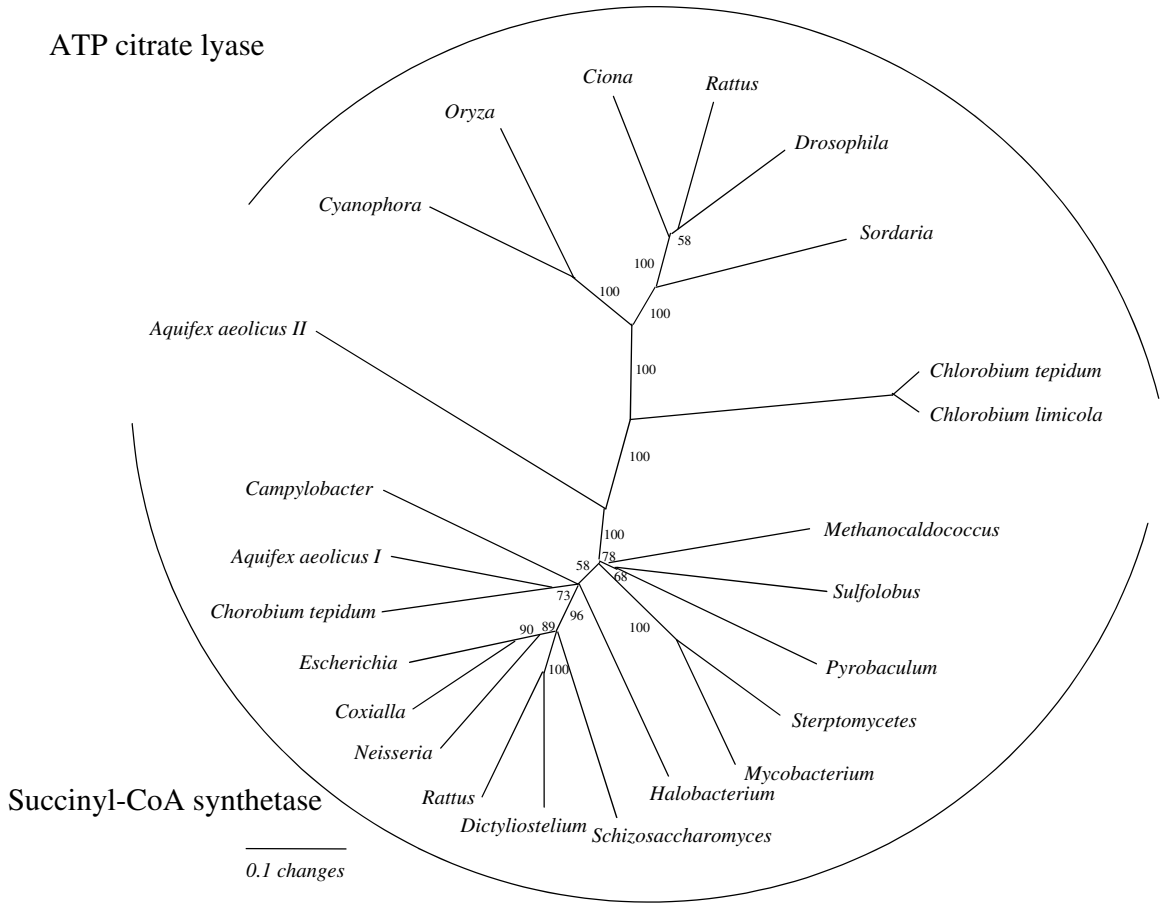

Figure 4. Neighbor joining tree showing the phylogenetic relationship between homologous domains in Succinyl-CoA Synthetase (SCS) and ATP citrate lyase (ACL). Numbers represent bootstrap values after 100 replicates. Note that the phylogenetic position of the second copy of SCS of *Aquifex aeolicus* is between those of SCS and ACL.

Aquifex aeolicus has two distinct copies of succinyl-CoA synthetase [106, 109] (**Figure 4**). Based on preliminary phylogenetic analysis using Neighbor Joining methods [121], this second copy appears to be more closely related to ATP citrate lyases, however more rigorous phylogenetic analysis will be required. It may be that *Aquifex aeolicus* uses a modified succinyl-CoA synthetase for citrate cleavage [109]. Considering the basal position of the Aquificales in phylogenetic trees based on 16S rRNA, this would be in line with the evolutionary scenario proposed by Fatland et al. [103]. First, we will verify the operation of the reductive TCA cycle in *Aquifex aeolicus* by using diagnostic enzyme assays, e.g., for ATP citrate lyase activity. At this point, it has only been assumed that *A. aeolicus* is fixing CO_2 using the reductive TCA cycle because of the operation of this cycle in the related species *A. pyrophilus*. In the case where we obtain a positive result for ATP citrate lyase activity in *A. aeolicus*, we will confirm the hypothesis of Deckert et al. [109] by over-expressing the suspected gene in *E. coli* [122].

Primers based on putative target sequences are used in a PCR reaction to amplify the target gene, which is ligated into an expression vector (e.g., pET 11a, Novagen) and transformed into *E. coli* (e.g., BL21-Gold (DE3), Stratagene). Utilizing the thermophilic nature of the overexpressed enzyme, purification will be carried out using a heating step and subsequent gel filtration.

Survey of ACL genes in other lineages. There are several autotrophic organisms in which carbon fixation pathways have not yet been examined. *Ammonifex degensii* is one of the most interesting from an astrobiological perspective. It is a thermophilic autotroph that reduces nitrate to ammonia, and rRNA sequences from closely related organisms have recently been identified as dominant sequences in surveys of microbial communities living in the oceanic crust [123].

Protists will also be screened for the presence of ATP citrate lyase. The best candidates would be free-living anaerobic or microaerophilic representatives such as *Carpediemonas membranifera*, members of the Euglenozoa, members of the ciliates, and some parabasalids (e.g. free living monocercomonads such as *Monotrichomonas carabina*). Biomass will be harvested and DNA and RNA extracted as for the prokaryotic species. Biochemical assays will be carried out to confirm ATP citrate lyase activity prior to PCR amplifications.

Phylogenetic analysis of ACL genes. We will use our expanded database of ACL sequences to study the molecular evolution of ATP citrate lyase and the reductive TCA cycle. We will determine if phylogenies based on our expanded data set of ACL genes recover well-recognized groups of prokaryotes or whether there is evidence for lateral gene transfer. If phylogenies hold well-recognized groups together, this will provide a framework for future work using ACL as a functional gene marker in specific PCR (or quantitative PCR) to study the diversity and distribution of organisms that utilize the reductive TCA cycle for CO_2 fixation in the environment. In a manner similar to analyses of rRNAs and *dsr* studies described above, samples could be analyzed from deep-sea hydrothermal vents and the deep subsurface, the acidic Río Tinto, and microbial mats occurring in the salt marshes around Woods Hole. These environments can be viewed as analogues for habitats that might have existed on early Earth or that could exist on other planets, and as a result, are likely to be inhabited by organisms with ancient metabolic traits. The discovery of novel sequences with ancestral features will also help us to elucidate the origin and evolution of the reductive TCA cycle.

Organisms capable of autotrophic metabolism were an integral part of early ecosystems. They serve an important function by making inorganic carbon available to other organisms, a central component of the global carbon cycle. The balance between the fixation of inorganic carbon and its release through heterotrophic processes is primarily responsible for the carbon dioxide and oxygen concentrations in the atmosphere. The microbial communities on early Earth most likely existed under conditions that are today confined to what one may describe as extreme environments. Organisms that thrive under those conditions primarily utilize carbon fixation pathways other than the Calvin cycle. This project has the goal to better understand the evolutionary history of one of the most ancient carbon fixation pathways and to lay the foundation for an assessment of its environmental significance. As part of this proposal, environments will be studied that have resemblance to conditions on early Earth and that can also be considered as analogs for possible microbial communities on other planetary bodies. Autotrophy is a requirement for life to exist anywhere, yet very little is known about the organisms responsible for primary production in these microbially dominated ecosystems, making this project an important component to further our understanding about the functioning of these systems.

This project will also foster the interaction between different partners involved in this proposal. Examples include but are not limited to the group of Monica Riley, whose expertise in molecular evolution and comparative sequence analyses will be very valuable. Other ties include links to the Bay Paul Center's DNA sequencing capability and their expertise in phylogenetic

analysis. There will also be extensive collaborations with the other groups working on early metabolisms, namely Andreas Teske (UNC), Katrina Edwards (WHOI), and Linda Amaral Zettler (MBL). Besides sharing samples, the interaction with these groups will lead to a better understanding of the microbial communities inhabiting these systems by combining multiple lines of inquiry.

5.2 TERRESTRIAL ANALOGUES FOR EARLY MARS:

Today, the surface of Mars seems an unlikely place for active biology. It is a cold, oxidizing environment that appears to lack liquid water, although water ice is present in polar regions, and water vapor and ice crystals occur in the atmosphere. There is evidence for a warm and wet early Mars and numerous geological features show that liquid water flowed from the subsurface out onto the surface of Mars at many times throughout its history [124]. Even more exciting are the recent results from Mars Global Surveyor and Odyssey that point to the presence of liquid water and water mobility in the very recent history of Mars. These include the gullies interpreted to represent the flow of liquid water in relatively recent periods [125], evidence for a latitude-dependent layer of water ice and dust deposited from polar regions toward the equator down to about 30 degrees north and south, apparently by recent paleoclimate changes [126-129], abundant near surface water ice detected by Odyssey spectrometers in approximately the same latitude bands [130], and geologically recent mountain glacier deposits seen in equatorial regions along the margins of the Tharsis Montes [131]. This changing view of Mars as a potential abode for life, and discoveries of life whereever liquid water is found on Earth, rekindles enthusiasm for seeking evidence of extant or past life on Mars.

The study of environments enriched in iron may provide insights into the biochemistry and metabolisms of the early Earth and ancient Mars. For example, there are many theories that include pyrite (FeS_2) as a driving force in the origin of life. Pyrite is the most common metal sulfide on Earth [132] and was thought to be involved in prebiotic Earth chemistry in its interactions with phosphate. Furthermore, pyrite is the most stable iron mineral under anaerobic conditions [133] as existed on early Earth. Wächtershäuser [134] proposed that life arose via an anoxic chemoautotrophic pathway described as a "two-dimensional chemiautotrophic surface metabolism in an iron-sulfur world" wherein the reductive citric acid cycle and other nonenzymatic chemistry of the prebiotic Earth occurred on the surface of iron sufide minerals. Russell et al. proposed a model whereby membrane-like accumulations of the iron-sulfur mineral mackinawite allowed the oxidation of hydrogen to be coupled to the reduction of Fe (III) leading to the first protometabolism and the evolution of life [135].

Our theme, **TERRESTRIAL ANALOGUES FOR EARLY MARS** explores the biology and diversity of iron oxidation, which we hypothesize to be relevant to early metabolic processes and physiologies on Earth and possibly Mars. Our studies will focus on two kinds of environments that serve as possible terrestrial analogues for Mars; basalt-rich, mid ocean ridges and iron-sulfide-rich, acidic rivers. We will study the biological basis of iron oxidation and fractionation of Fe isotopes by neutrophilic bacterial isolates from metal sulfide deposits at the Juan de Fuca Ridge and will extend our studies of iron oxidation to eukaryotic cultures (fungal and protistan) from the Río Tinto of southerwestern Spain. These efforts will complement ongoing investigations of iron oxidation by prokaryotic acidophiles of the Río Tinto currently under investigation by the Centro de Astrobiología (CAB, an NAI Associate). We will employ a new technology “SAGT” to efficiently characterize diversity and population structures of microbial communities (both prokaryotic and eukaryotic) in the water column and biofilms of the Río Tinto. These studies of microbial communities will be correlated with variations in biogeochemical parameters. Finally, we will explore linkages between microbial diversity in the Río Tinto and the Juan de Fuca Ridge and the geological processes ([136, 137] and others) that created and modified these environments, and then develop recognition criteria for key

geological settings and minerals to enhance and focus the search for such potential astrobiological sites on Mars. The ultimate objective of this interdisciplinary effort is to develop improved models and methodologies for life detection by remote sensing.

5.2a Iron Oxidation – Shaping past and present environments.

Iron (Fe) is the most abundant element in the Earth [138]. This pervasiveness underscores the importance of Fe, and Fe cycling in the geochemistry of the planet and biogeochemistry of life, including the emergence and development of life on early Earth during the Archean (3.8 to 2.5 Ga). During this eon, life proliferated and achieved extraordinarily levels of complexity in an ocean that was essentially devoid of dissolved oxygen and sulfide [139, 140]. The Archean geologic record contains abundant laminated structures such as stromatolites and banded Fe formations (BIF), both of which have been dated to 3.8 - 3.5 Ga and have arguably been considered as artifacts of the early life on our planet [141, 142], though the precipitation of Fe oxides in BIFs predates the oxygenated atmosphere. Indeed, recent studies have suggested that it is plausible, if not likely that BIFs are the product of biogenic oxidation of Fe in the absence of oxygen [143]. For example, the oldest photosynthetic bacteria were most likely anoxygenic purple non-sulfur bacteria [144] and several of these bacteria have been shown to fix carbon dioxide (CO_2) through the photosynthesis and the oxidation of ferrous iron (Fe^{2+}) rather than the oxidation of H_2O [142]. Along with photosynthesis, the sulfur and oxygen depleted Archean ocean could have favored the evolution of metabolisms that utilize alternative electron donors (e.g. Fe^{2+}, Mn^{4+}). It has even been suggested that residual Fe^{2+} respiration occurs in some *Pedomicrobium* spp. that have evolved an aerobic metabolism, perhaps hinting at an ancestor that was able to conserve energy from this now derelict pathway [145, 146].

Current theories about the Archean ocean indicate it was enriched in dissolved Fe, perhaps as much as 50 μM [147]. At present, free dissolved Fe is largely absent from the world's ocean, because an oxygenated atmosphere prevents buildup of reduced iron. However, local anomalies (e.g. Loihi Seamount, Hawaii, acid mine drainage, groundwater), where dissolved Fe is present in the water column or sediment water interface, persist. However, other sources of Fe potentially available to microorganisms during the Archean and today are present as the solid components of Earth's crust. Fe-rich rocks and minerals are abundant in both terrestrial and oceanic habitats. For example, sulfide ore deposits that are mined for economic metals as copper, gold, zinc, and others, are comprised of Fe-rich minerals such as pyrite (FeS_2), which are known to support the growth of significant populations of Fe-oxidizing bacteria in the modern Earth environment. It has also been hypothesized that Fe-rich sulfide minerals played a role in the evolution and metabolism of Earth's earliest life forms [134]. Despite the limited availability of iron in the modern biosphere, many proteins contain FeS centers. The repeated use of such centers in protein evolution may provide clues to the use of iron in the origin of life and the nature of the early biosphere. Additionally, approximately 70% of the Earth's near-surface crustal material, most of which is covered by our oceans, is comprised of basaltic rock, which is generally about 10 wt. % FeO [148]. Recent studies have described bacteria, archaea, and eukaryotes that thrive in environments similar to the Archean ocean, systems containing high dissolved Fe (e.g. The Loihi Seamount, HI, [149]; Michigan ground water, [150], Iron Mountain, CA, [151]; Río Tinto, Spain, [152, 153]) or in bare-rock systems such as the metal sulfide deposits at Juan de Fuca Ridge [154]. While many of these recent studies are among the first to demonstrate the association of Fe with life on Earth, evidence of such an association has long been recognized – if only qualitatively in many cases.

Though abundant evidence clearly exists that indicated the relative importance of Fe cycling in the early evolution of life and chemoautotrophic metabolism by comparison to any other similarly ancient metabolic pathway (e.g., sulfate reduction, Fe reduction), the fundamental biology of Fe oxidation has received little attention. This may be attributed to the recalcitrance

of Fe-oxidizing microorganisms, particularly non-acidophilic Fe-oxidizers to laboratory growth and manipulation. However, recent strides in our laboratory and several others have now made it possible to interrogate these elusive bacteria in axenic culture. Concomitant with these strides, recent technological advances in isotope biogeochemistry, specifically in the area of stable metal isotopes, enable us to similarly probe the systematics of Fe isotope fractionation in biologically produced rocks and minerals, allowing for the potential development of robust biosignatures for these organisms. Here we propose to conduct detailed laboratory studies aimed at exploring microbial Fe metabolism – from cellular processes to mineralogical and geochemical signatures, with the goals of advancing the fundamental understanding of the biology of Fe-oxidizing microorganisms, and establishing mechanistic linkages to the influence they exert on geochemical processes on Earth, and possibly on other planetary bodies.

Río Tinto. Many examples of Fe-rich aquatic environments arise as a consequence of (or are exacerbated by) industrial activities such as mining, while others occur naturally. Our study site in southwestern Spain focuses on the Río Tinto, a 100-km long, acidic, heavy metal laden river running through the world's largest pyritic belt. Environments like the Río Tinto provide a good terrestrial model for Mars because there is growing evidence that certain mineralogical and sedimentilogical features on Mars could have only formed in an acidic environment [155, 156]. The mining of the Río Tinto region is one of the longest recorded mining operations in the world. The river's name refers to the color of its waters – "tinto" meaning red in Spanish – the result of high levels of dissolved ferric iron (Fe^{3+}) kept in solution by the acidity of the water. In other parts of the river, reduced iron also occurs in high concentrations. Geomicrobiological characterizations show that the Tinto's Fe cycle is operational under both aerobic and anaerobic conditions. Although more than 5,000 years of mining activity (from Chalcolytic times to the present) has altered this region's landscape, the nature of the water – an average pH of 2.0 and heavy metal concentrations several orders of magnitude above normal– has a much longer history. Paleontological studies of iron stromatolites has shown that Fe-oxidizing bacteria existed in the Río Tinto river basin 300,000 years ago [152]. These stromatolitic structures have recorded the geomicrobiology and climate change of the Río Tinto system and are reminiscent in composition and banding pattern of the Archean BIFs. The metabolic byproducts of similar, if not identical, microbial communities maintain current acidity levels of this Fe-rich environment.

Juan de Fuca Ridge. Deep-sea environments are among the most pristine and ancient ecosystems available to study the origin and evolution of life. Recent analyses of remote sensing thermal emission data point to Mars as having an abundance of basalt and possibly basalt that was weathered under submarine conditions [157]. Mid-Ocean Ridge (MOR) systems, such as the Juan de Fuca Ridge, are bare rock ecosystems, which host organisms ranging from hyperthermophilic Archaea to giant *Riftia* sp. tubeworms. However, the biological communities associated with the spreading centers are not limited to the areas directly influenced by the hydrothermal fluids. Newly produced ocean crust, as well as metaliferrous sediments and deposits produced by precipitating hydrothermal fluids, also store an important pool of energy, which can be harvested for cellular processes. It was recently estimated that the production of new crust along the MOR could support a biomass production, from the oxidation of Fe^{2+} within the crust and associated precipitates, of 10^{11}g C /yr using oxygen as the terminal electron acceptor [158]. Much of this oxidation would occur under a fully oxygenated water column, perhaps with microenvironments approaching hypoxia or anoxia.

Evolution and diversity of oxidative Fe metabolisms in neutrophilic and acidophilic prokaryotes. Neutrophilic Fe-oxidizing bacteria (FeOB) have been recognized since the mid 19th century, although until recently evidence of their existence had been circumstantial, reliant upon the production of "diagnostic" morphologically distinct, extracellular Fe oxide minerals [159]. A variety of Fe oxide morphologies have been described and attributed to the activities of FeOB, including "stalks" (presumably from a *Gallionella-like* organism) and "sheaths" (presumably from a *Leptothrix-like* organism), which are the two most widely recognized and cited morphotypes [145]. These morphological markers are often found in habitats were anoxic water containing generous amounts of Fe^{2+} slowly mix with oxygenated water, resulting in the production of a rust-colored precipitate which can be trapped in a biofilm.

Morphological markers have long been observed microscopically in mineral deposits from deep-sea hydrothermal environments and attributed to the activity of FeOB (eg. [160-164]). These diagnostic Fe oxide precipitates have been found in rocks from ancient hydrothermal vents, and a variety of other Fe-rich ancient habitats [165, 166]. In a recent inspection of museum mineral collections and field observations, more than 140 sites were found which contained evidence of biogenic Fe-oxide formation [167]. To date, the best evidence of these distinct biogenic Fe-oxides in a modern Earth habitat comes from observations of microbial mats at the Loihi Seamount. At this site three Fe oxide morphologies have been described: the classical twisted stalks (similar to *Gallionella sp.*), straight sheaths (similar to *Leptothrix sp.*) and amorphous oxides [149]. These authors calculate, assuming the stalks and sheaths are of biogenic origin, that up to 60% of Fe oxide production within microbial mats is the direct result of microbial respiration [149]. A report of biogenic Fe-oxide production from a pure culture estimates that 90% of the oxidation is due to microbial activity [168]. However, the physiological mechanisms that enable these organisms to harness the chemical energy from the oxidation of Fe^{2+} have not yet been elucidated.

Chemolithoautotrophic FeOB have been recently isolated and described from the Juan de Fuca Ridge and shown to be phylogenetically diverse. While previously described Fe-oxidizing microorganisms have been inferred to be members of the Archaea and beta-Proteobacteria, new isolates from the Juan de Fuca Ridge have been shown to be members of the alpha- and gamma-Proteobacteria [169]. Interestingly, the closest known relatives of these Juan de Fuca FeOB are heterotrophic and some (*Pedomicrobium sp.*) display the ability to oxidize Fe^{2+} (for unknown physiological purpose) along with heterotrophic growth.

Despite the fact that evidence for Fe-oxidation was first recognized in neutrophiles, the metabolic role of Fe-oxidation has been more clearly understood, and better studied, among acidophiles. Acidophilic Fe-oxidizing organisms have been recognized in terrestrial sulfide mineral weathering habitats such as the Río Tinto, and part of the intense interest and study of these organisms is linked to the deleterious ramifications their activities have on the environment in the form of "acid mine drainage" (AMD)[170]. For example, *Acidithiobacillus ferrooxidans* generates protons via the oxidation of pyrite to ferrosulfate and sulfuric acid [171]. Fe-oxidizing bacteria such as *Acidithiobacillus ferrooxidans* (formerly known as "*Thiobacillus ferrooxidans*"; [172] and *Leptospirillum ferrooxidans* are examples of well studied acidophilic, mesophilic chemolithotrophs. The newly described Fe-oxidizing archaeon *Ferroplasma acidarmanus* also lives in extreme acidic conditions (pH=0-1) and is thought to be the key player in the mining process. All of these genera can be found in and some of them have been isolated from the Río Tinto [152] and are currently the focus of intense study by members of the CAB.

Iron metabolism/regulation in acidophilic eukaryotes. Iron is vital for life. We know that Fe is required by all living organisms except lactobacilli and a few other bacteria. Fe-associated proteins include hemeproteins, iron sulfur proteins containing Fe/S clusters, mononuclear non-hemeproteins, and diiron oxo-bridged proteins [173, 174]. In humans, excess Fe can be lethal

while Fe deficiencies lead to anemia as well as cognitive impairment and immunodeficiency [174]. Excess Fe is toxic because of its capacity to produce hydroxyl radicals via Fenton chemistry. Organisms like *E. coli* possess bacterioferritin to cope with this problem [175]. Yet, typical environments on Earth have low concentrations of bioavailable Fe and so organisms have developed elaborate ways to scavenge Fe (e.g. bacterial siderophores and phytosiderophores).

What happens to organisms living in Fe-rich environments? We know little about the interactions of acidophilic eukaryotes with Fe, particularly under the high concentrations present in the Río Tinto. Fe oxidation in eukaryotes has never been explored systematically, but observations suggest that this process may occur, particularly among fungi, some of which have been found to oxidize Mn [176]. Mn and Fe oxidation appear to be linked in many environments among prokaryotes, but the same has not been shown yet for eukaryotes.

METHODS. We will examine neutrophilic and acidophilic Fe-oxidizing microorganisms to fully explore the cellular processes and mechanisms that enable microorganisms to harness energy from the oxidation of Fe^{2+}. We currently have 16 axenic cultures of neutrophilic Fe-oxidizing bacteria from the Juan de Fuca Ridge (Edwards Lab). Basic physiological and phylogenetic studies of these strains are complete [169], as well as some field studies that have focused on environmental, geochemical, and mineralogical aspects of their physiology and activities [146, 154]. We are presently well poised to launch studies to examine in detail the biochemical mechanisms used by Fe-oxidizing bacteria and to develop a framework with which to understand the mineralogical and isotopic consequences of their activities that may be recorded in the Earth's rock record. We will generate a 2-4X shotgun genome sequence of a gamma-Proteobacteria neutrophile that we successfully culture in our lab (strain F16). This strain groups phylogenetically with other neutrophiles that we have isolated and with the acidophile *Acidithiobacillus ferrooxidans* (*Thiobacillus ferrooxidans*) whose genome sequence is near completion (http://www.tigr.org/). By comparing the metabolic pathways in neutrophilic and acidophilic strains, we aim to identify the molecular basis of functional and physiological linkages between these two types of Fe-oxidation metabolism. Physiology studies on acidophilic prokaryotic cultures from the Río Tinto are being carried out by our Spanish colleagues at the CAB (R. Amils).

The mechanism of Fe-oxidation has been studied mostly in *A. ferrooxidans* where a complex Fe oxidase system consists of the protein components: Fe (II) oxidase, cytochrome c4-, cytochrome c, rusticyanin and cytochrome oxidase. All these are required for complete Fe oxidase activity. The components may differ in different acidophilic iron oxidizers. For example the Fe oxidase system of *A. ferrooxidans* differs from that of *L. ferrooxidans* and BC1, a moderately thermophilic gram positive Fe-oxidizing bacterial strain [177]. *Leptospirillum ferrooxidans* does not have rusticyanin, but instead has a soluble acid stable cytochrome a. Strain *BC1* also lacks rusticyanin but possesses a novel membrane bound yellow chromophore. These findings suggest a diversity of Fe-oxidation enzyme systems in Fe-oxidizing bacteria. Despite this putative diversity, components of the electron cascade of different acidophilic Fe-oxidizing bacteria may be similar. In *A. ferrooxidans* a periplasmic Fe (II)-oxidase has been purified and the gene cloned [178, 179]. Cytochrome c protein, rusticyanin and cytochrome oxidase have been purified, and the rusticyanin gene has been cloned and sequenced. [180-183]. The cytochrome oxidase protein has also been crystallized in *Paracoccus denitrificans* [184]. Less is known about the molecules involved in Fe-oxidation carried out by neutrophiles. In the sheath-producing neutrophile mentioned above, *Leptothrix discophora*, Fe-oxidation has been attributed to a 150,000 kDa protein [185].

We will also explore the potential for Fe oxidation in acidophilic eukaryotes that we have isolated from the Río Tinto. With the levels of Fe found in the Río Tinto, one would expect that if Fe-oxidizing eukaryotes exist, the Río Tinto is a model environment in which to search for

them. Dr. Ricardo Amils at the CAB has agreed to do prescreening of 1,200 fungal cultures for pH stability and Fe tolerance and will then send us 10 representative cultures to screen for Fe-oxidation activity. We will also screen protist cultures for evidence of iron oxidation, regulation and other Fe interactions. We currently have pure cultures of a chlamydomonad, an euglenid, a chlorellid and a vannellid amoeba (Amaral Zettler culture collection). Purification of mixed cultures of diatoms, ciliates and several phylogenetically distinct amoeboid groups is underway (including a nucleariid amoeba, a vahlkampfiid amoeba, a gromiid, and a heliozoan).

The overarching goal of the proposed work is to understand the evolution and diversity of oxidative iron metabolisms in microorganisms from extreme environments and their relationship to mineral structures that may serve as biosignatures. The aims of this study are as follows:

1. Determine the fractionation of Fe isotopes using neutrophilic cultured FeOB
2. Study the Fe oxide morphology formed by neutrophilic iron oxidizers
3. Identify proteins and genes common to both the acidophilic and neutrophilic FeOB using biochemical and genomic techniques
4. Examine the role of acidophilic eukaryotes in Fe oxidation/regulation

Isotopic fractionation of Fe. The role of biology in cycling of metals in the environment may exert significant control on the fractionation of isotopes. While the fractionation of light isotopes, for example C, N, O, S, is influenced by abiotic processes including changes in temperature and phase transitions, intermediate mass elements like Fe may not be fractionated significantly by abiotic processes as a result of smaller relative mass differences between isotopes [186]. However, the small relative mass difference between ^{56}Fe and ^{54}Fe may result in a significant fraction if metabolic processes incorporate a number of steps at which a fractionation could occur.

Natural ^{56}Fe/^{54}Fe isotopic ratios have been shown to vary in nature over a small range (3.5 per mil) necessitating high precision methods for determining relative ratio changes due to various factors [187]. Recent advancements in the methodology of Fe isotope determination have increased the precision of measurements enabling the analysis of mass dependant fractionations. Fe isotope determination has been measured using both Thermal Ionization Mass Spectrometry (TIMS, eg. [186, 188] and Multi-Collector Inductively Coupled Plasma Mass Spectrometry (MC-ICP-MS, e.g. [187]. Because of the high ionization efficiency and stable mass bias of the latter instrument, low detection limits are attained. However, MC-ICP-MS co-eluding interferences from the argon plasma (primarily ArN and ArO interfere with ^{54}Fe and ^{56}Fe respectively) or other metals (eg. Cr, Ni) can plague the measurement of the Fe isotopes. These difficulties can be circumvented through use of a higher resolution scans, changes in the typical (nitric) acid matrix, using large quantities of Fe (5-25ug) or by adding a reactive collision gas to the plasma [187]. Beard et al. ([187], in press) report the complete removal of ArN through the mixture of Ar and H_2 gas in the collision chamber. These amendments to the technique allow for high precision (±0.3 to ±0.05 per mil) determination of the Fe isotopic ratios.

To date, very few studies have examined biological fractionation of Fe isotopes and include the study of (1) magnetite biomineralization by magnetotactic bacteria [188] (2) fractionation due to Fe(III) reduction by *Shewanella alga* [186, 187] and (3) mineral dissolution by siderophore producing soil bacteria [189]. These studies report negative fractionation from the reduction of Fe(III) of –1.3 and –0.8 per mil ^{56}Fe/^{54}Fe from the *Shewanella* and siderophore experiments respectively. Fractionation from the formation of the oxide mineral magnetite was not demonstrated due to the high detection limit (0.3 per mil) of the method used (TIMS) and pools measured. However, subsequent review of the work concluded that further knowledge of the equilibrium fractionation between the dissolved Fe and the magnetite could reveal a net biological fractionation, which opposes the direction of an equilibrium fractionation [187].

The effect of biological and chemical Fe-oxidation on the Fe isotope ratio will be determined through a set of experiments using strains from our collection of neutrophilic FeOB. Our justification for focusing on the neutrophiles for this portion of the work is in part related to the work described in the next section (**Fe oxide morphology**), and is thus elaborated on there. Experiments will involve growing bacteria in a minimal mineral media using synthetic $Fe(CO_3)_2$ [169] as an Fe^{2+} source. The bacteria will be inoculated into a dialysis bag (10,000 MW), which allows for exchange of dissolved species while containing particulate matter (both cells and Fe oxides) within the bag. The bag and its contents will be harvested and the contents acidified in with 0.5% SeaStar grade nitric acid in a clean Nalgene bottle. The aqueous solution will be treated identically. Relative isotope ratios will be determined using a Thermo Finnigan Neptune MC-ICP-MS using the method of Beard and Johnson [190] including using H_2 gas to eliminate the ArN and ArO co-eluding peaks. Measurement will be taken simultaneously on Faraday collectors using 10^{11} Ω resistors for ^{54}Fe, and ^{52}Cr and 10^{10} Ω resistors for ^{56}Fe. Chromium interference of ^{54}Fe is corrected using the relationship $^{54}Cr/^{52}Cr=0.0282$ [190]. Subsequently, the Fe isotope ratio will be determined under changing culture conditions including temperature, Fe source ($FeCl_{2(aq)}$, basalt, pyrite), pO_2 and pH.

Fe oxide morphology. As noted above, a long standing hallmark of the activities of the neutrophilic FeOB has been the production of morphologically distinct mineral particles. In addition to the differences in pH conditions that are used to classify the two types of Fe-oxidizers we have discussed here, i.e., the acidophiles (generally thriving at pH •4) and the neutrophiles (most often found occurring at pH •6), there are notable differences in the morphologies of Fe oxide particles that are produced as a consequence of their activities. In general (though certain exceptions occur), the activities of acidophilic Fe-oxidizers such as *A. ferrooxidans* and *L. ferrooxidans* result in precipitation of amorphous or poorly crystalline oxyhydroxides that are morphologically indistinct. Occasionally these particles may encase the bacterium that produces them, resulting in a cell "cast" that may be preserved in oxide deposits, but often the Fe oxide deposits that result are largely devoid of recognizable or distinguishable features. In contrast, many if not most neutrophilic FeOB produce morphologically and sometimes mineralogically distinct Fe oxide particles, as discussed above. In keeping with this generality, we have found that strains in our culture collection of neutrophilic FeOB produce a wide variety of Fe oxide morphotypes, some of which are similar to those previously described [145], but many of which appear not to have ever before been observed in laboratory growth cultures [169]. SEM and TEM studies of laboratory cultures and field studies of rocks from which our cultures were derived reveal particle morphologies that include long (20-30 μm), narrow (1-2 μm) angular particles, irregularly coiled particles, and cohesive bundles of straight particles that are composed of individual oxides in the range of 20 nm or less in diameter [154, 169]. Mineralogically, all morphotypes appear to composed of "2-line ferrihydrite" [191], a very finely crystalline form of ferrihydrite [154].

To date we have not undertaken systematic studies of the Fe oxide morphologies we observe in culture, but preliminary observations suggest that the oxide morphologies vary both as function of culture type (species/strain) and culture conditions. We hypothesize that careful study of both the conditions and culture type that gives rise to specific Fe oxide morphologies may be diagnostic to the environmental conditions that give rise to similar particles that we, and many others, have observed in modern and ancient depositional environments on Earth, as well as potentially serve as biomarkers on other planetary bodies. This would be a particularly powerful tool if used in conjunction with diagnostic isotopic signatures, such as proposed above. For these reasons we will focus our mineralogical/morphological studies proposed here, and the Fe isotope work proposed above, entirely on our neutrophilic FeOB culture collection.

In parallel with the Fe isotopic studies designed to explore isotopic fractionation as a function of changing culture conditions (temperature, Fe source, pO_2 and pH), we will conduct

light (Differential Interference Contrast, DIC) and electron (SEM and TEM), and Atomic Force (AFM) microscopic studies on Fe oxide particles produced in culture compared with abiotic controls. The DIC and AFM will be conducted using a Digital Instruments Atomic Force Bioscope. The AFM head is used simultaneously and in-line with the light optical components of an Inverted Zeiss Axiovert 100. This equipment is available and maintained in the Edwards lab. TEM and SEM studies will be conducted at the MBL microscopy facility, which is equipped with a Zeiss 10CA transmission electron microscope and a JEOL 840 scanning electron microscope (supported by a multitude of preparative and ancillary equipment).

Identification of Fe oxidizing genes and proteins in neutrophilic prokaryotes. The important role that FeOB play in geochemical processes is controlled at the cellular level by the genes and ultimately the proteins that participate in terrestrial sulfide dissolution, AMD, rock weathering in the deep-sea, and Fe redox cycling in a variety of aquatic Earth habitats. We propose to conduct molecular biological studies to identify Fe-oxidizing proteins and genes in select strains of neutrophilic FeOB from our Juan de Fuca Ridge culture collection. We will conduct protein purification studies on representatives from α- and γ- proteobacterial lineages with high growth rates and yields. Methodologies for growth and harvest of the neutrophiles will be as described [169]. In Edwards et al. 2003 [169] we detail aerobic growth on several different Fe substrates. For our work here, we plan to use ferrous carbonate, as it appears to facilitate most rapid and abundant growth as well as iron oxide formation.

Physiological studies of *Leptothrix ochracea* and *Gallionella* dominate the research that has been carried out thus far on neutrophilic FeOB. Nothing is known about the genes and little about the proteins involved in Fe-oxidation in neutrophilic FeOB. We propose to use a two pronged approach to identify the proteins and genes involved in neutrophilic iron oxidation. We believe this approach provides the most robust combination of methodologies for answering some of the key questions regarding the basic mechanism of Fe-oxidation.

We propose to initiate our protein work using the methodology of Corstjens et al. and Takai et al. [185, 192]. The former use SDS-PAGE followed by an assay to detect Fe-oxidation activity in the gel. The assay consists of rinsing with 10 mM MES at pH 6.0 for 45 min., followed by incubation in 2 mM $(NH_4)_2Fe(SO_4)_2$ in 10 mM MES at pH 6.0. Development of brown or yellow color indicates Fe-oxidation activity. Using this fairly simple colorimetric assay, it was possible to identify an Fe-oxidizing factor of 150,000 kDa from the neutrophile *Leptothrix discophora*. We will use both SDS and native PAGE to identify the iron oxidizing factor in our study. In a study by Takai et al., several components of the Fe oxidase activity from the cell membrane fraction were detected on SDS PAGE after purification, with masses of 46 kDa (α), 28 kDa (β), 24 kDa (γ), 20 kDa (δ), and 17 kDa (ε) [192]. However, with native PAGE, the homodimeric enzyme was estimated to be 263 kDa $(\alpha\beta\gamma\delta\varepsilon)_2$. We will use a similar approach to isolate the total cell membrane proteins of the neutrophilic cultures. Since the technique that Takai et al. developed was designed for use with enzymes from acidophiles (with low pH optima), we will have to standardize the chromatography for purification of neutrophilic membrane proteins using common gel filtration and ion exchange chromatography (Pharmacia). Purified proteins will then be sent out for N-terminal sequencing to a commercial facility (Rockefeller University). In order to identify the components of the iron oxidation system, we can use the protein sequence to degenerate DNA probes for Southern hybridization or for PCR amplifications. PCR amplicons or genomic fragments that anneal the DNA probes will be excised, cloned, and sequenced to obtain the entire gene sequence of interest. The information generated provided by DNA sequencing will allow us to examine if there is a connection between the phenotype (a given Fe-oxide morphology) and genotype of Fe-oxidizing organisms in our current culture collection – and among any Fe-oxidizers yet to be discovered. This technology is intellectually straight-forward and routine in molecular biology laboratories but we are well aware that this approach can be labor intensive. As an alternative, genomics based

approaches can accerate acquisition of similar kinds of information. When combined with even minimal protein sequence information, it becomes possible to identify relevant coding regions and entire pathways.

We will use information gained from the above work in combination with comparative genomics to determine genes and pathways involved in Fe-oxidation. We will sequence draft coverage (2-4X) of the genome of a relatively well-characterized neutrophile that we successfully culture in our lab (the strain F16). This analysis will provide the first genomic data of a neutrophilic Fe-oxidizing bacterium, and will make a significant contribution to the field of microbial genomics. Among the various neutrophiles we have isolated in our lab, we selected F16 for genome sequencing based on three criteria. First, on a practical level, we selected F16 due to the relative ease of culturing this strain in the lab and our ability to obtain high quality DNA for genomic libraries. Second, F16 is closely related to several other neutrophiles in the gamma-Proteobacteria, and thus may be representative of a group that is abundant in the environment. Third, F16 is a member of a phylogenetic assemblage of γ-proteobacteria that includes the acidophilic Fe-oxidizer *Acidithiobacillus ferrooxidans* (*Thiobacillus ferrooxidans*) whose genome sequencing is near completion (http://www.tigr.org/) and several *Pseudomonas* spp. whose sequences are completed (*Pseudomonas putida* KT2440 and *Pseudomonas aeruginosa* PAO1) or near completion (*Pseudomonas syringae* DC3000, http://www.tigr.org/). This availability of full genome data for closely related lineages will facilitate annotation and will allow us to compare the metabolisms of acidophilic and neutrophilic Fe-oxidizing bacteria. Through those comparisons, we hope to identify any common mechanisms for Fe-oxidation and differences that may reflect their distinct habitats and physiologies.

First, we will determine the genome size of the neutrophilic strain F16 through Pulsed Field Gel Electrophoresis (PFGE) using the Biorad CHEF Mapper XA in the Bay Paul Center. Using this approach, bacterial cells will be embedded in agarose plugs, in which intact genomic DNA will be prepared and subsequently digested with rare cutting enzymes. The resulting large genomic fragments will be sized using PFGE and summed to calculate total genome size. If F16 is greater than ~4 Mb, we will measure chromosome sizes for the six other close relatives of F16 in our culture collection of neutrophilic oxidizers, and select an isolate with the smallest size genome for shotgun sequence analysis. The calculations below assume a genome size in the range of 4 Mb, which is typical for γ-Proteobacteria. We will then use conventional recombinant DNA techniques to construct a genomic library [193] and sequence 10,000 randomly selected, recombinant clones using the 3730 ABI sequencer at our high throughput sequencing facility. Estimating an average genome size of 4 Mb and read lengths of ~900 bp, we will collect a minimum of ~2 X coverage of the genome. If the sampled sequences are Poisson distributed, this coverage will reveal >85% of the genome's coding capacity. Sequences from both ends of each insert will be assembled using ARACHNE [194, 195]. Contigs plus singletons will be analyzed by similarity searches against GenBank using NCBI BLASTX [196-198]. Besides searching for similarities to genes identified as being involved with Fe-oxidation, we will use software such as Pathway/Genome Database (PGDB) designed to help infer gene function, metabolic and genetic networks [199]. For example, in *A. ferrooxidans*, the genes encoding the two c-type cytochromes, cytochrome oxidase and rusticyanin occur on the same operon [200]; the same may be true for the neutrophilic pathway. Depending on the results of this study, future work could include microarray experiments to study upregulation and downregulation of genes during growth on alternative minerals such pyrites and siderites. Genome information will also aid in elucidating other metabolic processes in these extremophiles.

Screening acidophilic eukaryotes for iron oxidase activity. Whether acidophilic eukaryotes living in high iron environments can harness the energy of iron-rich substrates has never been systematically studied. Such metabolic capacities in eukaryotes most likely occur in environments where iron is not a limiting nutrient but these conditions are relatively rare in

contemporary ecosystems. The Río Tinto is unusual in this regard because it is iron rich and supports a broad diversity of eukaryotic protists and fungi. There is a strong correlation between Mn and Fe oxidation. Studies of fungi suggest iron oxidation may occur in eukaryotes. Structures formed in the presence of Mn-oxidizing fungi resemble "*Metallogenium*" particles that are observed at the aerobic-anaerobic interfaces in natural environments where Fe and Mn are deposited [145, 176]. These particles were first described as organisms (hence the genus name) but subsequently, no detectable biological membranes or nucleic acids associated with these particles were found and it was later suspected that they may be formed by diffusible Mn-oxidizing factors secreted by the fungi. Formation of "*Metallogenium"*-like particles is of astrobiological interest because microfossils resembling these structures have been described in precambrian rocks ([145] and references therein). Furthermore, any particles that have direct biogenic links would be important to explore for potential biosignatures.

Unlike the prokaryotic iron oxidizers, assays for iron oxidation in eukaryotes have not been reported in the literature. The potential for photosynthetic activity to catalyze oxidation of Fe^{2+} as seen with Mn^{2+} in *Chlorella* sp. could lead to false positives in assays for iron oxidation [201]. But, low pH may retard such activity. Obviously this potential interference is of little concern for studies of nonphotosynthetic species. Dr. Ricardo Amils at the CAB will prescreen fungal cultures and provide us with ten isolates for further testing. We will include examples from our culture collection of acidophilc protists isolated from the Río Tinto in a systematic survey of iron oxidation. We will survey for proteins involved in possible iron oxidation in eukaryotes using biochemical assays of proteins resolved by SDS- PAGE as described above for neutrophilic prokaryotic iron oxidizers. If we discover evidence of iron oxidation, we may explore options of isolating and characterizing proteins involved in iron oxidation.

With the completion of this work, we will gain a better understanding of some of the microbial physiologies that leave behind a geochemical record. Earth's extreme environments such as the Río Tinto and the Juan de Fuca Ridge serve as possible analogs of environments like that of Mars and therefore search for biosignatures similar to those found on Earth would be of great astrobiological relevance.

5.2b Microbial diversity and population structure studies in the Río Tinto

Declared dead over four centuries ago, the River of Fire (from the Phoenician "Ur-yero"), today known as the Río Tinto, runs 100 km through the world's largest pyritic belt in southwestern Spain and empties its acidic waters into the Atlantic. A combination of microbial ecology and molecular diversity studies show that the Río Tinto is teeming with microbial life [153, 202, 203]. Chemolithotrophic microorganisms (bacteria and archaea) feeding on polymetallic sulfidic ores contribute to the low pH (1.7-2.5) and high iron concentrations in the Río Tinto. Metals including Cu, As, Zn, Cr, Ni, and Ag are present at high concentrations, sulfate can reach 15g/l and redox potentials range from +280 mV to +610 mV [204-207]. Although mining activities have altered the system during the last 5,000 years, evidence of its antiquity older than 0.3 Ma (measured using Uranium-Thorium isotopes) has been found in massive laminated iron beds in three iron rock levels at different elevations above the Río Tinto [152]. Studies of the microorganisms in the Río Tinto may provide clues to ancient pathways or metabolisms no longer used by microbes living under less extreme conditions.

Acidity contributes to the extreme nature of the Río Tinto, but high heavy metal concentrations may present a greater challenge to the microbes living in the river. Yet, chemolithotrophs have learned to harness the energy of the iron-rich substrates found in the river. Pyrite serves as an energy source for chemolithotrophic prokaryotes since it contains both iron and sulfur that can be used by sulfur and iron oxidizing microbes – both found in the Tinto. The major iron-oxidizing prokaryotes in the river include *Acidithiobacillus ferrooxidans*,

Leptospirillum ferrooxidans, and *Ferroplasma* spp. that have been isolated and cultured from different parts of the river. Quantitative fluorescence *in situ* hybridization (FISH) experiments show that *A. ferrooxidans* and *L. ferrooxidans* are the most abundant FeOB [208]. Examples of microbes capable of reductive iron and sulfur metabolisms are also found [203].

A striking aspect of the Río Tinto is the contrast in phylogenetic diversity between the prokaryotic and eukaryotic communities. While the Río Tinto does not support complex metazoan life (although rotifers and insect larvae have been observed in the river; Amaral Zettler, personal observation) unicellular eukaryotes are abundant and more diverse than prokaryotes. Along with chemolithotrophic prokaryotes, photosynthetic protists represent the main primary producers in the river, and even numerically dominate the biomass in some parts of the river [204]. Morphological studies reveal the presence of diverse eukaryotes including fungi, different kinds of algae, various unrelated amoeboid protists, various unrelated flagellated protists, and ciliates. Most recently, molecular studies based upon comparisons of 16S-like rRNA genes PCR amplified from biofilms in the Río Tinto demonstrate even greater eukaryotic diversity than seen at the light microscope level [153] (see **Figure 5**; compare clones in bold from molecular survey with underlined taxa identified directly with light microscopy). To date, molecular surveys of the Río Tinto have been restricted to domain-level investigations of prokaryotes (Bacteria and Archaea) OR eukaryotes (Eukarya) but combined studies have not been undertaken. The nature of samples studied has contributed another significant bias. Eukaryotic studies have thus far only focused on biofilms and not bulk water column samples. In contrast, prokaryotic studies have only targeted the water column. We have only begun to dissect the biodiversity present in this rare ecosystem and the factors that shape community structure in the Río Tinto.

In this project we will take an integrated approach to assess the molecular diversity of both prokaryotic and eukaryotic components of community diversity in selected regions along the Río Tinto. We will employ a new technique developed at the MBL called SAGT [209]) in combination with traditional full-length rRNA gene sequencing of novel taxa not already in our Río Tinto sequence database. We use molecular techniques to investigate microbial diversity and population structure in the water column at both spatial and temporal scales, and we will collect concomitant physico-chemical measurements using water samples taken for the molecular surveys. When possible, determination of as many *in situ* physico-chemical parameters will be made on biofilms as well, although we acknowledge the limitations imposed by microelectrodes available for field measurements. This will allow for the correlation of biological diversity information with physico-chemical parameters of the river. The answers gained through this study will give us a comprehensive view of the microbial ecology of the system, a first step towards establishing an ecological genomics project for the Río Tinto. As part of this effort we seek to answer the following questions:

1. What accounts for the remarkable eukaryotic diversity in the extreme Río Tinto environment?
2. From a molecular perspective, how homogeneous are the blooms of diatoms, euglenids and other algae that form dense communities in various parts of the river?
3. How does phylogenetic diversity compare between the water column and biofilms along the river's edge?
4. Does diversity of a given community correlate in any way with physico-chemical parameters at a given site?
5. Are environments with greater turbidity more homogeneous than less disturbed ones?
6. How are the prokaryotic and eukaryotic microbial communities linked to each other and to what extent do they interact at the gene level?

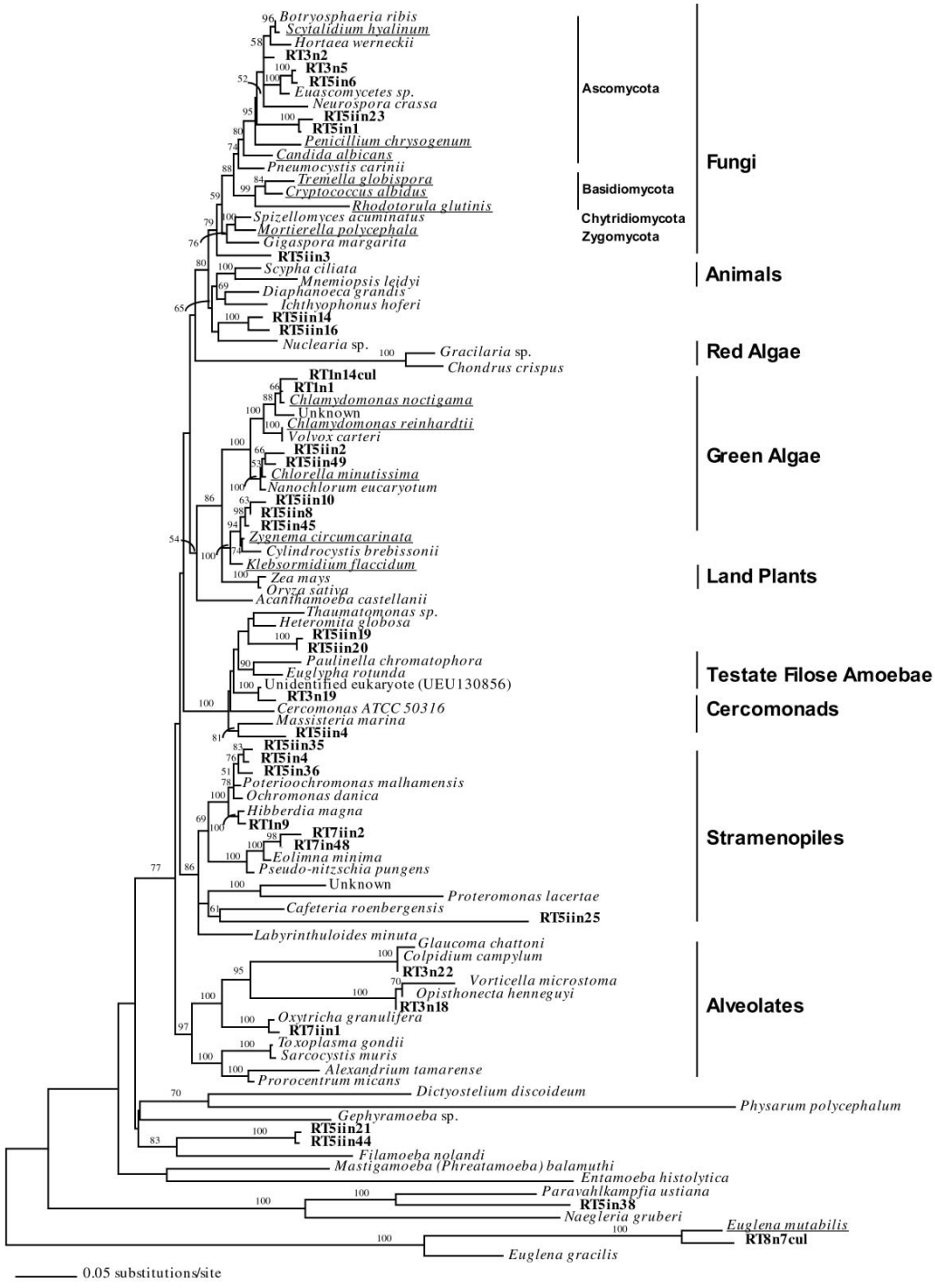

Figure 5. A minimum evolution phylogenetic tree for small subunit rRNAs using a likelihood model. Bold letters indicate environmental clones. "RT" indicates the sequence is from the Río Tinto and "cul" indicates sequences from cultured Río Tinto isolates. Underlined taxa represent genera that have been identified in the river based on microscopic observation. Sampling sites were as follows: RT1, La Palma; RT3, Berrocal Upper; RT5i, the Origin, black filamentous biofilm; RT5ii, the Origin, green filamentous biofilm; RT7i, Anabel's Garden green biofilm; RT7ii, Anabel's Garden yellow biofilm. Bootstrap support values are shown and the scalebar represents the number of substitutions per site.

Selected locations for study. We will establish the phylogenetic complexity and homogeneity of the water column of three chosen locations (henceforth designated "stations") along the river. Our studies will compare phylogenetic diversity in the water column to biofilm communities at a station and between stations along the river. We have chosen our stations based on varying concentrations and oxidation state of iron present (see http://astrobiology.mbl.edu/riotinto/ for more detailed information and photographs of the study locations). For each station we will sample three different sites, with three-fold replication. We will sample during each of the seasonal extremes of the river – one winter and one summer sampling. A brief description of the three locations and rationale for their choice are outlined below:

The "Origin" station where the river begins has several sources of water with very distinct properties and varying iron concentrations. Origin site 1 corresponds to the "Intersection" in Gónzalez-Toril, E., et al [208] where the waters from all three major sources converge before passing under the roadway (pH 2.7, 9.6 g Fe/L). Site 2 is a yellow pool below the wall and corresponds to Gonzalez-Toril site I (izquierda) (pH 2.9, 0.8 g Fe/L). Site 3 is a small deep red stream coming out from Peña del Hierro and corresponds to Gonzalez-Toril site II (izquierda, izquierda) (pH 1.8, 19.8 g Fe/L). Our eukaryotic biodiversity study of both black filamentous and green filamentous biofilms at the Origin yielded a broad range of rRNA phylotypes including sequences related to *Chlorella*, *Zygnema*, *Nuclearia*, unidentified fungi, cercomonads, different stramenopiles, valhkampfiid amoebae and other unidentified amoebae. (See **Figure 5**). We do not know which of these groups also occur in the water column. With respect to prokaryotes, FISH studies of water column samples demonstrated the presence of α–, β– and γ–proteobacteria as well as members of the Nitrospira (including the genus Leptospirillum) and representatives of the Archaea [208]. Sequence analysis of Denaturing Gradient Gel Electorphoresis (DGGE) bands from the same samples confirmed the presence of sequences that were most closely related to cyanobacteria, *Leptospirillum* species, *Acidithiobacillus* species, *Ferromicrobium acidophilus*, *Acidiphilium* species, *Magnetobacterium bavaricum,* and the archaeal species *Ferroplasma acidophilum* and *Thermoplasma acidophilum.* [203]. No attempts were made to characterize eukaryotic populations in these studies.

At the second station, Anabel's Garden, many different types of biofilm flourish and eukaryotic diversity and abundance is high. The average total Fe concentration at Anabel's Garden is 2.3 g/l and it is present in a very reduced form (high Fe^{2+}). Site 1 is in a small pool just below where the water exits from the culvert under the road; on the wall above this pool are black and green filaments attached to the wall in the rushing water. Site 2 is in a small pool to the right of site 1; it appears to fill from water seeping under the roadway and is covered with a thick mat of green algae. Site 3 is downstream of where the seep water from site 2 joins the main flow from site 1. Eukaryotic rDNA diversity studies of biofilms demonstrated the presence of diatoms and ciliates (**Figure 5).** Molecular surveys of prokaryotes in the water column using FISH and DGGE demonstrated similar kinds of species at both the origin and Anabel's Garden. [203].

The Berrocal site is just downstream of a bridge and is characterized by deep red water, localized anoxia, and dramatic concretions of river boulders cemented together with iron precipitates. The average total iron concentration here is 3.1 g/l. While there is little obvious eukaryotic biomass in this area, we have observed conspicuous blooms of *Euglena* in the shallow waters along the shore. We have also seen green and colorless flagellates, lobose amoebae, and heliozoa. Site 1 is just downstream of the rapid section below the bridge, site 2 is underneath the bridge, and site 3 is downstream of the rapids after the river reverts to a calmer flow. Different fungal sequences, cercomonad, and ciliate sequences have been obtained from biofilms at the upper portion of this station (RT3 in **Figure 5**). Water flow at Berrocal is both rapid and turbulent relative to other stations and this may have a significant impact on community structure. We will sample the water column here and localized anoxic sediments. Earlier DGGE

analyses showing domination by *A. ferrooxidans* [203] did not agree with FISH detection of α–, β– and γ–proteobacteria and members of the Nitrospiria group [208].

SAGT study of the whole microbial community in the Río Tinto. We will study the microbial diversity and population distribution in the Río Tinto using the SAGT technique [209]. This technology is modeled after serial analysis of gene expression (SAGE, [210]) which describes relative gene expression patterns. With SAGT, the PCR products from orthologous hypervariable regions in rRNA genes are ligated together to form large concatemers. A single DNA sequencing reaction of a cloned concatemer can include as many as 20-30 orthologous hypervariable regions represented in a population of nucleic acid molecules. In this way, samples loaded onto a 96-channel capillary sequencing machine can provide information about thousands of microorganisms in an analyzed sample.

The first step in the SAGT technique involves PCR amplification of small subunit rRNA genes from genomic DNA extracted from an environmental sample using conserved primers that contain a type II restriction site and flank a hypervariable region (20-100 bp) of the rRNA gene. The primer sequences are removed from the amplicons by digestion with this type II enzyme leaving a short base pair extension. Digested fragments are ligated to form concatemers that are then cloned and sequenced. The concatemer has the following structure:

[cloning vector]--[hypervariable #1]--[hypervariable #2]--...--[hypervariable #N]--[cloning vector]

The primer pairs can be designed to amplify any of the rapidly evolving regions in rRNA genes. Alternatively, we can design primers to amplify regions that are not rapidly evolving and hence restrict the granularity of the analysis, but at the same time identify population members that are more distantly related. It is also possible to design primers to target hypervariable regions from any of the three primary domains of life. Finally the technique offers a strategy for minimizing artifacts of PCR bias in analyses of microbial population structures. A population of PCR products from mixed templates will be biased if there are differences in processivity of the polymerase on different template sequences or if primers bind with different efficiencies in a template-dependent manner. A modest 10% difference in the efficiency of amplication for two starting templates present in equimolar concentrations will produce a 64-fold difference in final product concentrations after 30 rounds of amplification. Two methods for minimizing PCR bias include targeting short sequences for amplification (as inherent in the design of SAGT) and minimizing the number of amplification steps. With our SAGT strategy, we predict it will be possible to generate adequate numbers of amplicons for constructing concatemers of hypervariable regions by using only 5-10 rounds of amplification. After ligation into large concatemers, additional rounds of PCR can be used to obtain adequate concatemer copy numbers for insertion into cloning vectors. We can eliminate PCR bias introduced during this second round of amplification by restricting our data analyses to unique concatemer sequences.

To analyze the high-throughput data we rely upon a data analysis system, which uses PHRED [211, 212] to analyze the output from the sequencing machines and a signature recognition program that we will develop to identify individual hypervariable regions represented in the concatemers. We will build databases that represent the frequency of occurrence of each member of the concatemer and will use FASTA [213, 214] searches of the large rRNA gene database to identify the species to which the individual sequences correspond. We recognize that these short DNA sequences will not be sufficient for reconstructing evolutionary history but if identical or near identical sequences are represented in the large existing database for rRNA genes, it will be possible to obtain a putative identification for each member of the SAGT concatemers. If we discover completely novel hypervariable regions not represented in the databases, we will design primers complementary to these regions and in conjunction with standard conserved primers, amplify and fully sequence these novel amplicons. These novel sequences will then be included in phylogenetic analyses using methods and

programs routinely used in our center. Finally, we will use statistical methods to estimate microbial diversity and species population structure [215]. A special attribute of this technique is it can be "tuned" to address questions about population structure for different phylogenetic ranges. For example, through primer design, we can target genus-specific sequences that flank hypervariable regions. With such a strategy, the representation of amplicons within the concatemers will report relative numbers and diversity within a particular genus. The ability to quickly screen large numbers of hypervariable or moderately conserved regions provides a means to address questions about homogeneity of algal blooms, as well as ascertain the relative molecular diversity of microbial communities found in biofilm versus the water column. By combining diversity information with physical and chemical data (see below) we can explore how geochemical forces shape community structure in the Río Tinto.

Determining relevant physico-chemical parameters of the Río Tinto. The *in situ* measurements of physico-chemical parameters using specific electrodes will include pH, temperature, conductivity, redox potential, and O_2 concentration. Mass spectrometry and X-ray fluorescence will be used to measure S, Fe, Zn, Cu, Al, As, Ni, Mg, Ca, K etc. The speciation (e.g. Fe_2+/Fe_3+, sulfide/sulfate, As (III)/As (V), NO_3-) will require complementary techniques [216]. All physico-chemical parameters and biological characteristics of the river for each sampling site will then be assembled in a single data matrix. To identify the major factors that shape variation of the data we will use principal component analysis, a multivariate method that simultaneously considers many correlated variables and identifies the lowest number needed to accurately represent the structure of the data set. We will determine the significance of associations of the different variables by using the Pearson correlation test for parametric variables (i.e. all except oxygen, sulfate concentration, and redox potential values). When variables considered do not conform to a bivariate normal distribution, we will use non-parametric tests for association, which render correlation coefficients of rank association like the Spearman's Rho coefficient [217]. Our study will reveal potential physico-chemical correlations with distributions of prokaryotes and eukaryotes in the Río Tinto.

Towards ecogenomics in the Río Tinto. In most ecosystems including the Río Tinto, consortia of micro-organisms orchestrate key processes in geochemical cycling, but little is known about ecological and evolutionary responses of these communities to cyclic and transient environmental shifts. Process-oriented studies demonstrate there are feedback loops between biogeochemical gradients and structured microbial populations, yet there are no comprehensive descriptions of underlying biochemical and genetic mechanisms that govern these processes. The next advance in studies of microbial ecology will be to integrate microbial diversity, the entire spectrum of metabolic activities, and complex gene expression patterns that coordinate biogeochemical processes. DNA sequencing coupled with DNA microarrays for massive parallel expression studies or SAGE-like technologies, will provide the foundation to understand how microbial communities govern and respond to formation of biogeochemical gradients. With this advanced technology we can directly measure how microbial gene expression patterns in an entire ecosystem respond to changing chemical and physical parameters. The concurrent measurement of biogeochemical parameters, community-wide gene expression patterns and spatial descriptions of microbial populations offers a means to understand the structure and function of biogeochemical machinery at different levels of biological organization. We seek to discover the links between biological diversity and the resilience and stability of biogeochemical transformations.

Community-wide gene expression studies are a high priority for studies of microbial activities in a variety of extreme environments that are relevant to astrobiology. Examples of important environments include but are not limited to cyanobacterial mats, hydrothermally altered deep sea sediments, deep subsurface communities and iron-rich acidic biofilms and the water column of the Río Tinto. Detailed knowledge about microbial population structures as

described in this proposal are essential to the interpretation of massive, community-wide expression profiling experiments, but the cost of a comprehensive gene expression survey for any microbial habitat is still prohibitive. As an alternative, expression profiles based upon one or two species can provide important insights about how organisms change expression patterns under different conditions. Ultimately we aim to examine gene expression patterns of assemblages of organisms under different conditions in the Río Tinto (such as varying iron concentrations) to distinguish which genes are up-regulated or down-regulated. The objective would be to identify the structure and cooperative regulation of genes from different members that orchestrate biogenic interactions with iron in a microbial community.

As an example of this type of approach, we are currently working with DNA microarray experiments to measure gene expression as it relates to diel variation in microbial mats. We have constructed a genomic library of the mat-forming cyanobacterium *Microcoleus chthonoplastes* using DNA extracted from a culture of a strain collected from the Sippewissett salt marsh, in West Falmouth, Massachusetts. We have sequenced 10,241 clones with insert sizes of 650 – 1150 bp, and used this information to select the sequences to array. Selection criteria for including a gene was based on a BLASTX hit to a known cyanobacterial sequence with an e value <= $1x10^{-4}$ and a Pfam score >= 25. This filter ensured that clones on the array were from *Microcoleus* and that they contained a gene coding region. After removing redundant clones, we selected 1090 sequences to array. These sequences were PCR amplified from the vector, and the DNAs printed onto glass slides for expression profiling studies.

Preliminary experiments have used the array to examine differences in gene expression during the day and night. We took laboratory cultures of *Microcoleus*, placed them into a 14 hr light / 10 hr dark cycle for 3 days, and then collected cells after two hours of light explosure and after two hours of darkness. As expected, a number of genes directly involved in photosynthesis were expressed at higher levels in the day samples as compared to the night samples. A large number of other genes involved in growth, replication, and many other processes were also upregulated during the day. At night, we found "light repressed proteins" showing increased expression, as well as a number of regulatory proteins. These data have also shown us a number of unexpected putative results, such as an increase in expression of an iron transporter at night.

To help better understand and verify these results, we are currently in the process of assembling a time course of expression data across an entire diel cycle. This information should provide us with valuable information about the precise timing and daily expression patterns of many of these genes. Once we have collected these data on diel variations, we plan to take environmental samples directly from the Sippewissett marsh and see how closely our laboratory manipulations can match what is actually happening in the environment. Using a similar strategy, in the future we intend to initiate studies of gene expression underlying biogeochemical processes in the Río Tinto. Such studies would focus on key isolates from both the water column and biofilms as identified by the population structure studies. The goal would be to correlate gene expression patterns with physico-chemical parameter variations. Details are not provided since those studies lie beyond our immediate experimental horizon.

5.2c Life detection through remote sensing (J. Mustard, J. Head)

Deposits in natural Earth environments, such as the primary and secondary minerals at Río Tinto, provide the critical geochemical components for chemolithotrophic microorganisms feeding on polymetallic sulfides and have implications for biological processes at the root level [208]. The Río Tinto deposits and geological environments [218] are thus a unique laboratory for the study of astrobiology at the root level. Understanding the geological processes responsible for the creation and evolution of the Río Tinto environment will provide insight into formation of host environments from which such basic biology might originate, habitate or evolve [208]. The creation of the Río Tinto environment involved key primary and secondary

geological processes such as sedimentation, magmatism, aqueous interaction, mineralization, fluvial alteration, deposition and diagenesis [218, 219]. Quantitative modeling of these key geological processes will provide insight into the conditions that created this unique laboratory, and the ability to predict where other similar environments and laboratories might exist on Earth and on Mars. Recognition of these environments on Mars will inform future exploration objectives and strategies for astrobiological and sample return goals.

Recognition of these environments on Mars involves two fundamental steps: 1) Application of the geological models of the Río Tinto environment to Mars to develop a set of candidate exploration sites, and 2) Development of recognition criteria for key minerals associated with the Río Tinto environment and application to upcoming orbital mineralogical detection experiments. Thus, our proposed work involves: 1) Close biological and geological collaboration to explore the linkages and synergism between the geological conditions creating the deposits and the biological conditions and environments leading to this unique biota. This represents an important step in the dialogue between workers in the biological and planetary geosciences areas. This dialogue will certainly lead to better definition of the interactive role of geological processes and cycles, biological evolution, and the application of this to other planetary environments such as Mars and Europa. 2) Documentation of the geological environments in which this type of biological activity occurs and application to Mars. 3) Identification of remote sensing signatures for key minerals associated with this unique environment. 4) Recommendations to the Mars Exploration Program for identification of potentially similar sites on Mars, and development of strategies for surface exploration and sample return.

The Río Tinto ore body is a supergiant massive sulfide deposit associated with an ancient sill-sediment complex [218]. The magmatic intrusion of sills into wet sediments produced peperitic deposits and hydrothermal fluid/wet sediment interactions, yielding ore-forming metal emplacement (**Figure 6**). The Guaymas Basin appears to represent a somewhat comparable sill-sediment complex [220]. Similar geological environments of shallow intrusion and wet-sediment interaction are thought to be potentially common on Mars [221] (**Figure 6**). Although conceptual models exist for the emplacement events in the Río Tinto [218] and Guaymas Basin [222], detailed quantitative modeling has not been undertaken. In order to provide a better understanding of the emplacement processes at Río Tinto, and to provide a basis for predicting where similar deposits might occur on Mars, we will undertake a quantitative treatment of the interaction of magma and wet sediment, and the resulting alteration, mobilization, and ore emplacement, using as a basis the framework provided in [218] (**Figure 6**). Some treatments of aspects of the processes have been made [223-226] and we will utilize these where appropriate. We will employ techniques similar to those we have used in the study of magma ice/water interactions on the Earth and Mars [221, 227-229]. Using the criteria developed from this analysis, we will then apply them to the Mars environment to document examples of candidate geological environments that might exist there. This will involve treatment of the interaction of magmatic sills with the martian cryosphere (see [221], Figure 12) and with candidate martian oceanic sediments [129].

Specifically, using the geometry of the emplacement processes illustrated in **Figure 6**, we will develop a computer program to model the upward propagation of a dike from a chosen depth in the crust. We will then implement the coupled solution of the equations of motion for the magma in the dike and the stress distribution across the dike wall which determines the dike shape, as described for an elastic host medium [230] and for a visco-elastic host medium [231]. However, instead of solving the equations semi-analytically, which requires the simplifying assumptions that previous workers have been forced to make about the variations of magma density and crustal stress with depth, we will use a completely numerical solution that allows us to choose any magma density and any distribution of crust density and stress. This permits us to

investigate what permutation of source depth, magma density, and mantle-crust density-stress profile lead to dikes that propagate all the way to the surface or stall at some depth in sill-like or

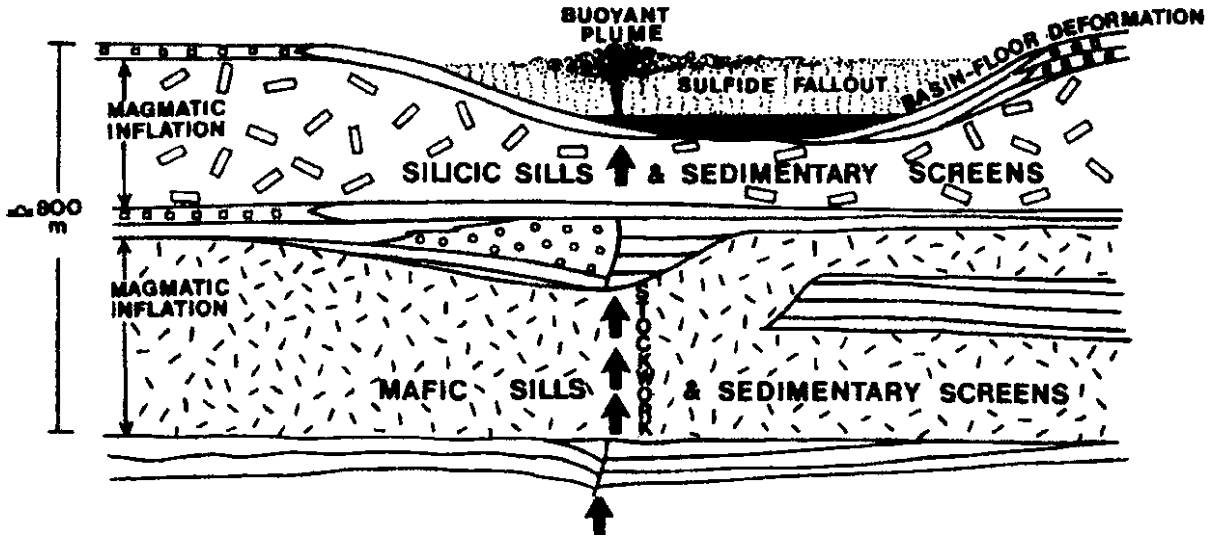

Figure 6. Schematic cross-section showing the Río Tinto deposit stratigraphic relationships and modes of emplacement in the sill-sediment complex. In this version, a major magmatic pulse creates the majority of the sill-sediment complex and basin-floor deformation traps exhalations that are driven by feeder dikes that are focused by the sill complex cap. Modified from [218].

dike-like intrusions. This approach will be particularly important for taking into account the different densities of the magma and the host rock (**Figure 6**). For dike and sill intrusions, we will also determine the final geometry using a numerical model that we recently developed to model dike and sill intrusion geometry on Mars [232]. The modeling will constrain dike and sill geometry and thickness, permiting us to assess the estimates of magmatic inflation predicted from the Río Tinto deposits (e.g., see **Figure 6**). These parameters in turn will allow us to evaluate in detail the near-surface stresses using a numerical stress simulation package (such as TEKTON) and predict in much more detail than previously possible whether features such as faults are expected to form. An additional analysis that we will perform for sill-like intrusions involves our current work on gas production during gas tip propagation. Furthermore, we will use conduction equations from this final configuration to model the mobilization of water-rich sediment to produce peperites, and to expel material to produce the buoyant plumes that are thought to produce the sulfide fallout (e.g., see **Figure 6**). For this latter aspect, we will utilize the quantitative treatment for seafloor submarine eruptions and ejecta dispersal that we have recently published [227]. In summary, the quantitative modeling of these key geological processes will provide insight into the conditions that created the unique Río Tinto deposits (**Figure 6**), and the ability to predict where other similar environments and laboratories might exist on both Earth and on Mars (**Figure 7**).

The deposits of interest in the Río Tinto region include those forming in relation to the Río Tinto itself. The Sinus Meridiani region of Mars has abundant gray hematite at the surface, possibly formed under aqueous conditions [233] and because of this is one of four sites proposed for the Mars Exploration Rovers. The headwaters of the Río Tinto have been proposed as an analog to Sinus Meridiani [219]. Groundwater mobilization of metals from the iron-rich sulfide ore bodies has created a metal-rich fluvial system. This has resulted in the highly acidic sulfate and iron-rich waters of the Río Tinto, and in the unique and flourishing biota [208]. Formation

of iron-bearing deposits in conjunction with the biota occurs presently and more ancient examples are preserved in Río Tinto river terraces. Furthermore, diagenetic alteration has locally modified the character of these deposits [219]. Finally, in the Mars environment, further alteration may occur. For example, on Earth, iron-rich oxidized caps above subsurface sulfide-bearing ore deposits are known as gossans. Gossans represent advanced states of chemical weathering in which sulfide minerals and related rocks are oxidized and leached leaving behind poorly crystalline hydrated iron oxides and sulfates [234].

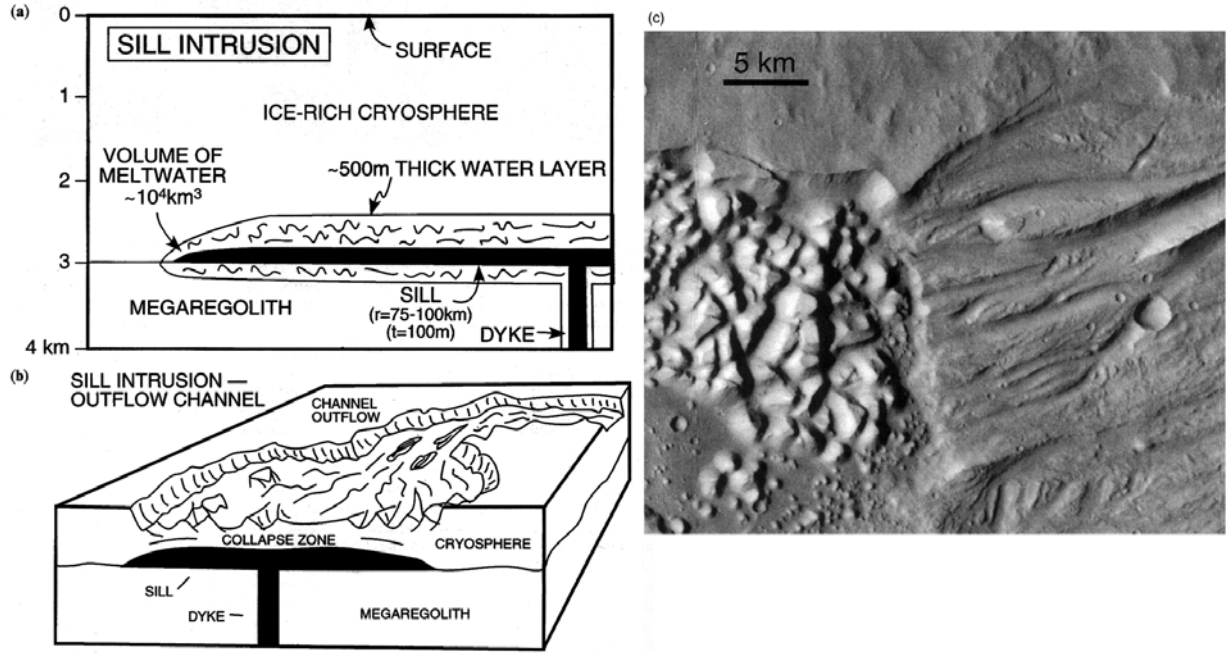

Figure 7. Interaction of a sill intrusion to produce melting, chaos formation and outflow channels on Mars. In (a) a sill intrudes into the megaregolith at the base of the ice-rich cryosphere, and conduction and melting occur, producing a significant volume of meltwater in only a few years. (b) Sill intrusion can plausibly cause surface uplift and fracturing, cryosphere cracking, significant melting, collapse, chaos formation and outflow channels. (c) Region of chaos at the head of any outflow channel on Mars that might result from sill intrusions such as those outlined in (a) and (b). From [221].

Through this investigation we propose to identify the key mineral signatures that are typical of these types of deposits and that might be recognized in future orbital and surface exploration in the Mars Exploration Program. The sulfur and Fe-mineralogy of the Río Tinto region is incredibly diverse [235] and there are multiple deposits related to the evolution of this system. Microbial activity in these environments is strongly coupled to the oxidation and reduction of iron (Fe) as well as sulfur and thus there is a suspected association between the Fe-sulfur mineralogy and biological activity [219]. Ferric and ferrous iron-bearing minerals exhibit a multitude of visible to infrared crystal field absorption features [236] which are diagnostic of mineral composition in natural environments [237]. Ferric oxides, oxyhydroxides, and oxyhydroxysulfates are particularly well characterized by their reflectance spectra (**Figure 8** Ferric oxide spectra). Our goal is to characterize the remotely-sensed signatures of iron-mineralogy in this biologically active environment to prepare a foundation for the interpretation of future spacecraft data.

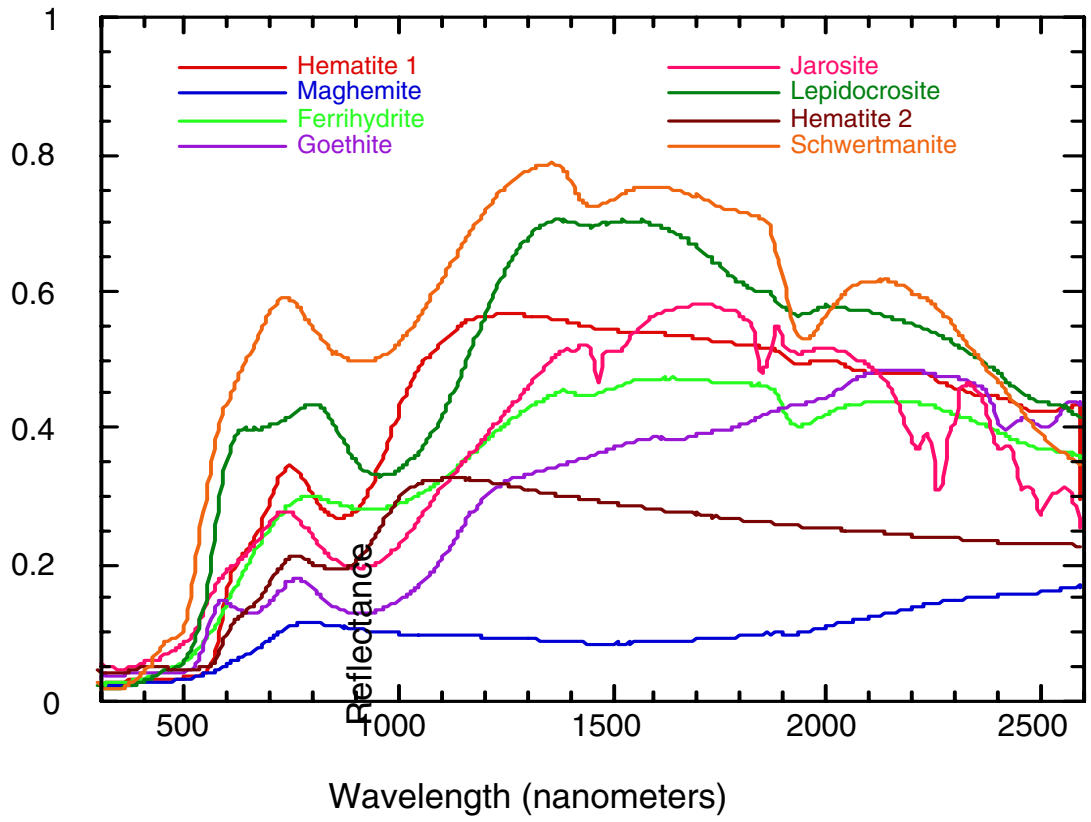

Figure 8. Ferric Spectra: Reflectance spectra of ferric oxides, oxyhydroxides, and oxyhydroxysulfates with a spectral resolution comparable to experiments on forthcoming Mars mission (e.g. OMEGA and CRISM). The position, shape, and strength of ferric and hydroxyl absorptions clearly distinguish these minerals from one another.

Upcoming Mars missions will have a number of sensors that will measure the reflectance of the martian surface in this wavelength region where Fe shows diagnostic mineral absorptions. The ESA Mars Express will have OMEGA, a visible-infrared spectrometer which will acquire global data with a 1 km spatial resolution in hundreds of spectral bands, and HRSC, a high resolution stereo camera (10 m/pixel) with multispectral coverage in the visible. The Mars Reconnaissance Orbiter (MRO) will have the CRISM instrument which will acquire data with spatial resolutions as high as 20 m/pixel and in 500 spectral bands from 0.4-4.1 μm (Mustard is a collaborator on the OMEGA experiment and a co-investigator on the CRISM experiment, Head a co-investigator on HRSC). The deposits in and around the active water courses of the Río Tinto and associated drainage basin span a wide range of iron mineralogy and appear to have distinctive properties with age [219]. Within the river system, seasonal drying and wetting results in the precipitation of amorphous ferric-bearing phases and soluble Fe-sulfate minerals (e.g. hydronium jarosite), though no associations between microbial communities and iron deposition have been established [216]. Increasing pH results in the formation of the mineral goethite and hydronium jarosite is replaced by goethite in older river deposits observed in terraces above the current river bed [219]. In still older deposits, hematite is observed to occur with goethite.

The focus for field sampling at Río Tinto is to address three key questions: What is the origin of eukaryote diversity, how are eukaryote and prokaryote communities linked, and what

relationships exist between the observed community diversity, the physico-chemical parameters and properties of a given site? We will extend the biological physico-chemical characterization of these sites with spectroscopic measurements relevant to the future exploration of Mars. This will be accomplished with both *in situ* and laboratory measurements. The *in situ* measurements will be made with an Analytical Spectral Devices (ASD) FieldSpec FR with a spectral range of 0.35-2.6 μm and a 1 nm sampling (though with coarser spectral resolution). For the laboratory spectroscopic measurements we will use NASA's RELAB facility housed at Brown University. This facility includes a bidirectional spectrometer with similar capabilities to the FieldSpec FR, but with higher precision and accuracy. There is also a Nicolet Nexus 870 Fourier Transform Infrared Spectrometer (FTIR), which can extend the spectral coverage to 200 μm. Therefore with a combination of the two laboratory instruments, the entire wavelength region critical to Mars past and future exploration with remotely sensed data can be covered.

There are four different environments that we will investigate in this context: 1) Primary mineralogy (minerals associated with the Río Tinto ore body itself), 2) secondary mineralogy (minerals associated with the groundwater migration and fluvial emplacement), 3) diagenetic alteration of fluvial deposits, and 4) gossan formation (minerals associated with intense chemical weathering). Thus the investigation tracks the process from unaltered mineralogy through the active biological zone to the older deposits with diagenetic alteration. Our initial focus will be the sites that will be analyzed in detail for microbial diversity. These sites will have associated physico-chemical measurements and we will analyze the sties with *in situ* spectroscopic measurements of reflectance from the visible to 2.6 μm. This will also include older deposits along terraces surrounding the current river bed. We will focus on terrace deposits that show evidence of the same basic depositional environment as the current active river bed. These environments will be sampled to acquire material to bring back to the laboratory for more detailed characterization. The laboratory analyses will include the collection of full spectral range and resolution measurements with the RELAB instruments as well as XRF and XRD measurements for mineralogic and elemental determinations.

This spectroscopic database will be analyzed to determine the nature of the iron mineralogy associated with the various active microbial communities, and how the iron mineralogy evolves with time. While we have measured all the expected minerals as pure specimens previously (e.g. **Figure 8**), the important part of this investigation is to sample the minerals *in situ* and in association with other components of the environment. This is necessary to determine unique associations of minerals that result in remotely sensed signatures diagnostic of the environments of interest. We will integrate these results into the geologic framework to devise a set of recognition criteria for identifying possible similar environments on Mars.

In summary, recognition of these environments on Mars involves two fundamental steps: a) Application of the geological models of the Río Tinto environment to Mars to develop a set of candidate exploration sites, and b) Development of recognition criteria for key minerals associated with the Río Tinto environment and application to upcoming orbital mineralogical detection experiments. The work described in this part of the proposal will complement other ongoing work in the NAI, specifically efforts by the CAB and at Harvard/MIT. We propose to collaborate with these teams to continue the development of Rio Tinto as a guiding terrestrial analog for the exploration of Mars hematite deposits by examining the spectral characteristics of the Río Tinto deposits and their similarities to the formation of the Meridiani hematite deposit on Mars. In our approach, we are developing quantitative models for the broad geological processes involved in the initial emplacement and subsequent evolution of the Río Tinto deposits. A common goal for several NAI teams with leadership provided by the CAB is to examine the Río Tinto deposits at the more detailed textural and mineralogical level and to make predictions that will be useful in the surface exploration of Mars, particularly those analyses performed by the upcoming Mars Exploration Rover Missions. Our effort is very complementary in that we will

be attempting to examine signatures and global geological contexts that will be useful in orbital missions that might examine the global inventory of sites and be useful in plotting out future orbital and surface exploration.

Within this NAI proposal, success in these endeavors will require synergistic collaboration between biologists and geologists to explore the linkages between geological conditions creating the deposits and the biological conditions and environments leading to this unique biota. By initiating a dialogue between the biological and planetary geosciences communities at Brown and MBL, we will gain an enhanced definition of the interactive role of geological processes and cycles, and biological evolution, and the application of this to other planetary environments such as Mars and Europa.

In addition to the benefits that will derive from the type of interdisciplinary interaction described above, this research will provide the following products: 1) Documentation of the geological environments in which this type of biological activity occurs and application to Mars; 2) Identification of remote sensing signatures for key minerals associated with this unique environment, and 3) Recommendations to the Mars Exploration Program for identification of potentially similar sites on Mars, and development of strategies for surface exploration and sample return.

5.3 BIOSIGNATURES AND LIFE DETECTION: DETECTING ANCESTRAL PEPTIDES AND PROTEINS

(M. Riley, N. Wainwright, R. Gast)

The search for signs of life in the Solar System and beyond is a high priority for Astrobiology. While remote sensing of biological signatures has the potential to identify promising sites for future missions, biomarkers provide a basis for the design of life detection experiments. We will combine our expertise in genome analysis and biochemistry to develop sensitive and rapid assay methods to quantify nucleic acid and amino acid sequences that are common to all forms of life including the most primitive organisms. The method will be applicable to both samples returned from future mission as well as miniaturization to be included as flight hardware for *in situ* analysis.

The earliest proteins on Earth would have arisen before cellular life as we know it. Many of the early proteins and fragments of proteins must have been ancestors of the complex protein world that exists today. Generating enough proteins to make a metabolically capable and self-replicating cell must have involved processes similar to gene duplication and divergence. Without assumptions about the process of duplication and divergence at those early times, we can nevertheless ask the question "How many and what kinds of ancestral proteins would be required to catalyze all the basic life processes found in all free-living organisms today?" We can also ask, "Is there a way we could detect features of terrestrial-like life that might exist on other planets?"

In the context of preparing to search for signs of life on other planets, we propose to develop a simple system for detecting either nucleotide or amino acid sequences of commonly occurring elements in proteins found throughout the living world on Earth. The elements we have chosen for detection appear repeatedly in Earth proteins. If we can detect these elements in diverse environments on Earth and if they are also used frequently in living systems elsewhere, then we may be able to detect signs of life on other planets. While we will propose specific sequences to test our methods, the strategies themselves will be flexible enough to accommodate any sequence.

We take as point of departure for the rationale, the characteristics of the proteins of the biochemically best-studied single-cell prokaryote, *Escherichia coli*. There are about 40 frequently occurring domains within *E. coli* proteins that have been characterized both structurally and functionally. Frequently occurring motifs in *E. coli* that might be examined are

listed in **Table 2.** Of these motifs, we have selected the two most frequent. The most common domain is the P-loop- containing nucleotide triphosphate hydrolase (175 instances), which occurs in many proteins that affect hydrolysis of ATP, GTP and other nucleotide triphosphates. The P-loop NTP hydrolase is the most commonly occurring ATPase type in living systems. Since this action is associated both with energy-requiring reactions and with a class of regulatory proteins, the motif is widely used for many separate reactions. The second domain we have chosen, the NAD(P)-binding Rossman fold (73 instances), appears in oxidative proteins. Oxidation is a frequently occurring type of energy-producing reaction, and NAD(P) is the cofactor most often used to capture the electrons to pass them along the electron transport chain to generate energy. The Rossman fold is the most common of several types of NAD(P)-binding sites. These functions are so universal and well-conserved that one would expect them to have evolved at ancestral times in protein development.

Table 2.

Number of Proteins	Domain description	Typical size (aas)
175	P-loop nucleotide triphosphate hydrolase	200
73	NAD(P)-binding Rossman fold	150
39	Winged helix, in regulatory proteins	100
37	Periplasmic binding, transport, type II	200
32	Actin-like ATPase domain ñ contraction	200
30	FAD/NAD(P)-binding domain	400
29	S-adenosyl-L-methionine-dependent methyl group transfer	150
28	Porins (transport)	320
26	Nucleic acid-binding proteins	60
22	Thioredoxin	100
22	Homeodomain-like ATP-binding site	50
22	His kinase of 2-component regulatory system	70
21	α/β Hydrolases	250
21	Periplasmic binding, transport, type I	270
21	ATPase domain in chaperones/DNA topoisomerase	120
21	Trans-glycosidases	200
20	Membrane proteins, all-	250

Since these two motifs appear throughout the entire sweep of living organisms on Earth, from the simplest to the most complex, they seem logical to use as probes to find if there are proteins on other planets that use these motifs. Implementing a search for fragments with either of these specificities on an unmanned mission will require a mechanical way to assay for the chosen targets and the appropriate methods to detect the targets.

Proposed method / Instrument development. Whether one is concerned with collection of data to support life or spacecraft microbial contamination detection, major desirable criteria include accurate and sensitive instrumentation, and the ability to report data near the point of collection and in near real-time. With those criteria as goals, a portable instrument was designed to have the capability of performing kinetic analysis of enzyme-related reaction development. This has been initially customized to the *Limulus* Amebocyte Lysate (LAL) assay for microbial detection; however, it is flexible enough to be adapted to a wide variety of assays that can be analyzed spectrophotometrically, including antibody and nucleic acid assays. The instrument is designed to operate independently, but to have the capability of integration via a serial data interface. The system is comprised of a heating block to control reaction temperature, a pump to

facilitate fluid movement, LED and sensor pairs to quantify optical changes during the test and on-board microprocessor and memory to control all aspects of the instrument during the test and store data results. Power requirements are 10.5 volts DC, with maximum consumption at start-up of 25 watts.

LAL is a sensitive enzyme cascade that is triggered by microbial cell wall material [238, 239]. The test has been miniaturized and run in a simple, linear micro fluidic chamber that sequentially 1) hydrates a dry, stabilized reagent, 2) incubates the sample and 3) reads color development spectrophotometrically. Data from a preliminary run is shown below in Figure 9.

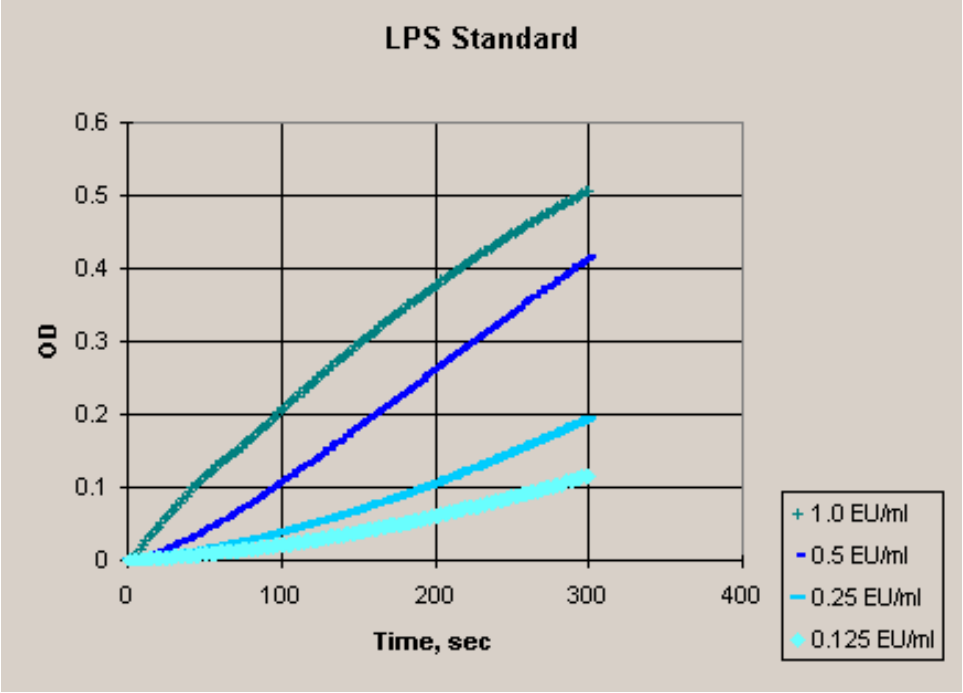

Figure 9. Each line represents the kinetic chromogenic development of one channel of a microfluidic chamber. Four channels are shown at 1.0, 0.5, 0.25 and 0.125 EU/ml (endotoxin units/ml). Data may be stored or interpreted by onboard software.

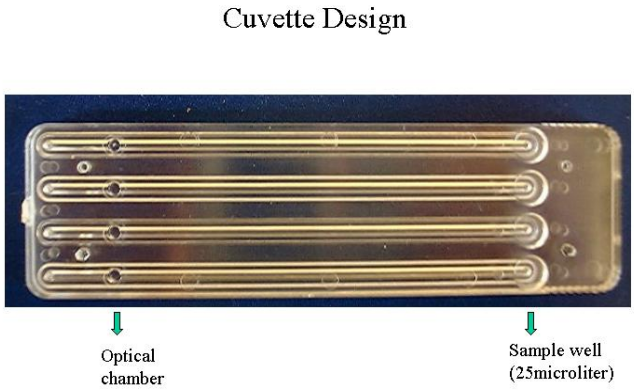

Figure 10. Microfluidic cuvette showing four parallel channels in which four separate tests can be performed. Sample is introduced on the right and is sequentially drawn over defined channel areas. Dried reagent become re-hydrated and chromogenic enzyme reactions are quantified by kinetic readings at the optical chamber.

In this study, we will develop methods specifically to:

1. Use the existing prototype instrument design to collect kinetic spectrophotometric data from competitive labeled antibody assays and competitive hybridization assays.
2. Develop the detection format initially based on fluorogenic and chromogenic antibody and nucleic acid probe label.
3. Increase method sensitivity by coupling the LAL enzyme cascade to LPS-labeled antibodies (or antigens) and nucleic acids.
4. Field test the methods with samples extracted from the environment.

Nucleic acid detection. The DNA detection method that we propose to develop is based upon the release of a labeled fragment in response to competitive hybridization with sample DNA. Research with competitive and strand displacement hybridization has displayed the use and effectiveness of these probe-target reaction systems to enhance the recovery of correctly matched hybrids [240-242]. We intend to modify this method to permit the detection of conserved nucleic acid sequences that correspond to the NAD(P)H -binding site of P-loop nucleoside triphosphate hydrolase and Rossman fold.

The method utilizes a double stranded probe DNA molecule in which the strand to be dissociated and detected is 20-25% shorter than the full-length hybrid [243], 2002). This design makes the incomplete hybrid less stable than the full hybrid, but more stable than full hybrids with clear mismatches. For our initial studies, we will label the strand to be dissociated with a fluorescent molecule. The complementary strand will be labeled with biotin to facilitate its removal from the hybridized solution after reannealing has taken place [244]. The labeled probe duplex will be mixed with double stranded tester DNA and denatured in a solution of buffer. Disassociated sequences are allowed to reanneal, the released fluorescent strand (and any hybrids) will be collected in the solution, and the solution assayed for the presence of fluorescent signal. Of the possible reannealed results in the solution phase, we will only be detecting the fluorescently labeled strands that have been released from the probe duplex by competitive rehybridization with correct matches from the tester DNA (**Figure 11**). The amount of label released can be directly proportional to the number of sequences present in the sample being tested [240].

Our probes must be synthesized to have heterogeneity at bases representing the third codon position due to the redundant nature of the genetic code. We recognize that this degeneracy can have a significant effect on the hybridization behavior of the probes. To counteract these effects, modified nucleotide bases will be incorporated to help reduce the potential melting temperature differences. P nucleotide is 6-(beta-D-ribofuranosyl)-3,4-dihydro-8H-pyrimido[4,5-c][1,2]oxiazin-7-one and it pairs with purines [245], and K nucleotide is 2-amino-9-(2-deoxy-beta-ribofuranosyl)-6-methoxyaminopurine, and it pairs with pyrimidines [246]. Both of these nucleotides can be substituted into polynucleotide sequences instead of degenerate bases, reducing the overall amount of heterogeneity by a factor of two. Most importantly, their pairing is significantly more stable than either mismatches or inosine base pairs [245-247], although it does not result in melting temperatures equivalent to the exactly matching hybrids. Primers using these nucleotide substitutions have been very successfully employed in the analysis of highly divergent sequences by denaturing gradient gel based studies[248, 249]. Such analyses are very sensitive to the change in melting temperature of duplexes introduced by even a single base change in the sequence, so the successful use of these modified nucleotides is very promising for helping to stabilize our degenerate hybrids.

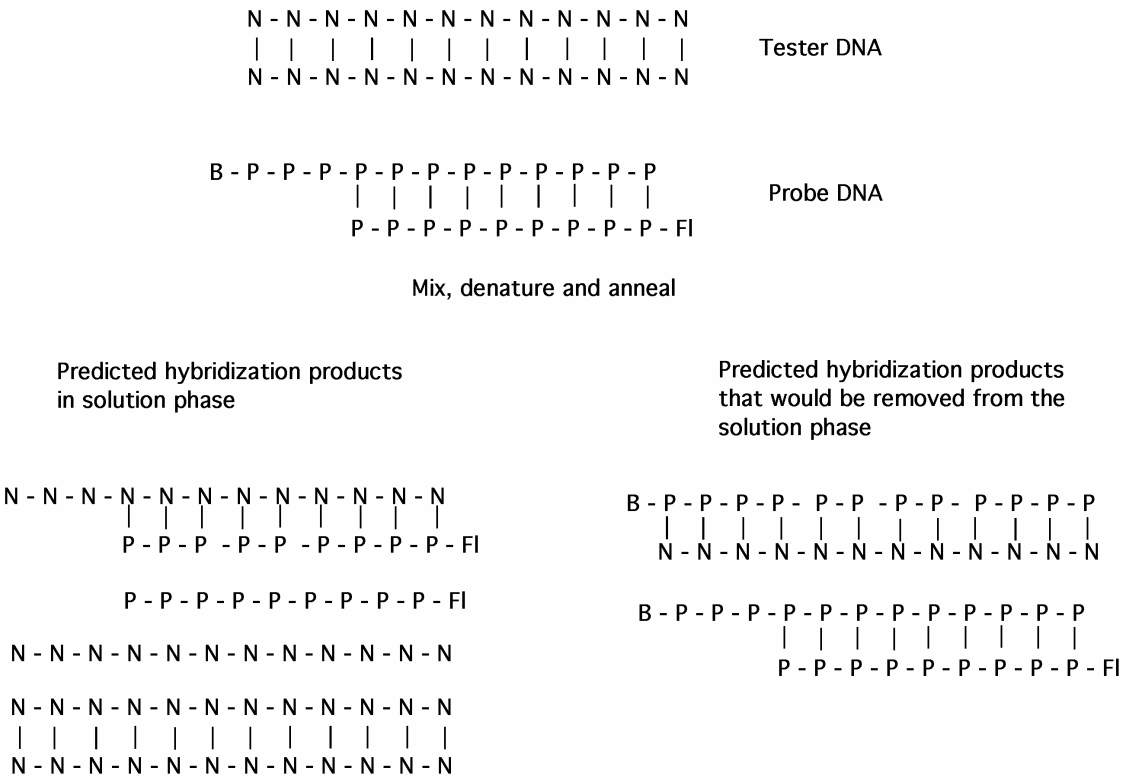

Figure 11. Proposed hybridization scheme and potential products. N = tester DNA nucleotide, P = probe DNA nucleotide, Fl = fluorescent tag, B = biotin tag

We will work with a suite of three probes, each about 50 bases long, representing the entire region of interest. Although competitive hybridization assays have previously been accomplished with probes as large as 131 bp [242], we believe that the use of fragments will increase our likelihood of detecting sequences by reducing the overall degeneracy of each probe resulting in more stable hybrids. Furthermore, the detection of hybridization by any of the probes in the mixture will result in a signal, thereby potentially increasing our ability to detect even highly divergent domains or smaller fragments. Our tester DNA will initially consist of cloned sequences from various organisms. As we move into more complex mixtures and environmental samples, we will explore shearing the tester DNA into uniform pieces of less than 1000 bp (or smaller if possible). This will help to make the hybridization process more homogeneous and manageable. The probe hybridization behavior will be tested at different temperatures, with different concentrations of probe and tester DNAs, and over a range of times to determine the best set of conditions for full hybrid association. The release of the displaced strand will be monitored by accumulation of fluorescence in the solution phase and quantified relative to experimental concentration standards that we create. Our hybridization experiments will also be conducted with the potential number of environmental targets in mind so that our tester DNA will be able to drive the reaction.

We desire to link our detection method with the LAL procedure because of the probable need to increase the sensitivity when working with environmental samples. Although we are targeting the sequences of broadly conserved protein regions, the signal is potentially going to be quite low. The competitive hybridization would be accomplished in solution, followed by

detection with LAL reagents being accomplished in the prototype chamber. The detectable release of labeled probe would suggest that the domain in question was present in the sample assayed. Future work would be directed towards establishing relative numbers of sequences present.

Experiments to determine hybridization activity:

1. Probes will be synthesized with fluorescent label on the 5' end of the strand to be detected and biotin on the 5' end of the strand to be removed from solution.
2. Solution hybridization reactions utilizing varying concentrations of probe and target sequences will be accomplished and monitored to determine optimal conditions for time, temperature and limits of detection.
3. Probes will be modified to have Lipopolysaccharide attached to the 5' end of the detection strand, and LAL detection cascade will be utilized to detect specific hybridization.
4. Detection limits for LAL cascade will be determined.
5. Probes will be tested on various environmental samples.

Peptide detection. Antibodies to selected peptides will be raised either as monoclonals overlapping the sequence of interest, or as affinity purified polyclonals. Antibody detection methods will be adapted to the unidirectional instrument cuvette design in the following manner:

1. Peptide antigens will be labeled by covalent attachment of fluorescein.
2. Mixtures of labeled and unlabeled peptide will be made; some with deliberately different sequences that might bind more competitively to divergent sequences.
3. Antibodies will be immobilized by adsorption to the plastic surface of microtiter wells or instrument cuvettes.
4. Sample peptide / labeled peptide mixtures will be exposed to the immobilized antibody.
5. Conditions of competitive binding will be optimized for antigen concentration, flow rates, antigen – antibody contact time.
6. Bound versus free label will be measured spectrophotometrically and compared to standard peptide concentrations.
7. Limits of detection will be assessed.
8. Detection limits will be increased by coupling peptides to Lipopolysaccharide (LPS) to link detection to the LAL amplification cascade.

Detection limits. This study addresses the need to develop increasingly sensitive methods for life detection. Our goals include a combination of the proven methods; antibody binding, nucleic acid hybridization, with the enzyme amplifying sensitivity of LAL, to create a new method. While the initial experiments to adapt antibody and nucleic acid tests to the existing unidirectional flow of the portable cuvette will use dye labels that do not press the limits of sensitivity, the second phase of the project will use LPS as a label to tie the detection signal to the LAL cascade, gaining the amplification of the enzyme cascade. We predict this linkage will rival the sensitivity of radioactive labels in the rapid, portable format.

Our combined assay format will address the need to analyze samples returned from other bodies in the solar system as well as to be compatible with potential flight hardware for *in situ* life detection. An instrument under development currently partially funded through ASTEP (MASSE, Andrew Steele, PI) is using the proposed instrument platform with LAL enzyme cascade reagents to quantify microbial cell wall material in a variety of minerals and fossilized samples. This effort will be an excellent opportunity to collaborate with NAI partners at the The Carnegie Institute and colleagues at the Johnson Space Center.

5.4 EVOLUTION OF GENOME ARCHITECTURE IN PROKARYOTES: Genome-Genome Integration: Symbiosis, genetic assimilation, and evolutionary innovation (J. Wernegreen)

In both prokaryotic and eukaryotic microorganisms, transformations in genome architecture far surpass point mutations as an important driver of evolutionary novelty. These changes include gene duplication and loss, horizontal gene transfer between distinct organisms, and the melding of two genomes through endosymbiosis. Over the past decade, new molecular capabilities have transformed evolutionary biology and enable us to trace the history of complexity at the whole-genome level. The studies below use comprehensive, genomic approaches to elucidate the forces that drive changes in genome architecture, and to link those genome changes with major evolutionary shifts. We will dissect the early and more recent genomic adaptations of endosymbiosis between bacteria and eukaryotes, resulting in close assimilation of two distinct genomes. This study will illuminate mechanisms by which genome changes contribute to major evolutionary transitions.

Descriptions of early life posit the progenote as a complex network of replicating entities engaged in fluid interactions. Close interactions and DNA exchange between genomes of early microbial forms yielded communities of ever increasing genetic and biochemical complexity. Genome-genome interactions remain an important evolutionary driver for microbial consortia. At a different scale but of equal importance, microbial symbionts that contributed to the formation of stable organelles have dictated the evolution of complex organisms. Endosymbiosis, in which one partner lives within the cells of another, represents the most intimate contact between organisms. Mitochondria and chloroplasts result from endosymbiotic events in which host cells acquired the ability to harness light energy and thrive in the presence of oxygen [250]. These innovations forever changed cellular structure, species diversity, and the range of acceptable habitats of Life.

Unlike single mutations, genome interactions can catalyze the acquisition of entirely new combinations of functions and drive major evolutionary transitions. Through studies of binary interactions between two species - i.e. a symbiont and its host- we can dissect the mechanisms of genome communication and coevolution. The Proteobacteria is especially important because many of its lineages have switched from free-living lifestyles to obligate genome interactions. For example, mitochondria evolved from an α-Proteobacterial endosymbiont. Other extant members of this diverse bacterial phylum live exclusively within eukaryotic host cells, but they represent incipient stages of cellular integration and organelle evolution. As such, they offer valuable insights into the mechanisms that drive genome interactions, including lateral gene transfer, gene loss, and modifications of gene expression. Through a comparative genomic approach, we will decipher the 'language' used to establish and maintain genome-genome interactions represented by stable and dynamic symbioses. Our ultimate objective is to understand how such interactions drive organismal integration and evolutionary novelty.

As model systems, we have selected two Proteobacterial endosymbionts that represent both stable and transient genome interactions. *Blochmannia* is a γ -Proteobacterial endosymbiont that has coevolved with ants for >40 MY [251-254]. Like mitochondria, *Blochmannia* benefits its host and experiences maternal transmission to host offspring. We study this endosymbiont as a model to understand coevolution in a stable genome interaction that may resemble early organelles. In contrast, the host switching of labile endosymbionts can drive lateral gene transfer and phenotypic plasticity. The endosymbiont *Wolbachia* is an excellent model to study evolutionary dynamics of unstable genome-genome interactions. Transfer of this intracellular bacterium among hosts explains its current infection of at least hundreds of thousands of invertebrate species and its status as the most abundant endosymbiotic bacteria on Earth [255]. The diverse lifestyles of *Wolbachia* span modification of host reproduction to enhance its own

spread (reproductive parasitism) to beneficial effects in its long-term association with nematodes (mutualism). The proposed *Wolbachia* projects extend upon the postdoctoral fellowship of Seth Bordenstein, an NAI/NRC Fellow in the Wernegreen lab.

Major research themes and experimental objectives: We aim to elucidate changes in gene content and expression patterns that catalyze the establishment and diversification of genome-genome interactions. The studies below examine invertebrate endosymbionts as model systems to address the following broader questions: How do genome interactions reshape genetic content and architecture? Do genetic differences drive phenotypic and ecological variation within a genome interaction? What specific biochemical functions mediate genome-genome interactions? Do changes in gene expression profiles catalyze the formation of genome interactions? How does plasticity in gene expression help to fine-tune coevolved interactions and drive diversity among them?

Remodeling of genetic architecture in genome-genome interactions. Genome interactions dramatically alter the genetic architecture and gene content of the organisms involved and partly explain the wide variation of genome sizes among taxa. Genome size changes are severe among bacterial endosymbionts, which include the smallest known prokaryotic genomes (as small as 450 kb, and typically < 1 Mb) [256, 257]. By comparing gene contents of endosymbionts and free-living relatives, we can better understand the process of genome reduction and the diverse strategies by which bacteria form intimate associations with eukaryotic cells. Genomes of the stable endosymbionts *Buchnera* (of aphids) [258, 259] and *Wigglesworthia* (of tsetse flies) [260] have shrunk by 80% but retain specific biosynthetic genes required by their hosts. Full genome sequences of two *Wolbachia* strains are now in progress [261, 262] and will allow us to interpret the observed variation in genome sizes of *Wolbachia* strains [263]. Despite the insights provided by current data, discerning general rules of genome remodeling will require a richer dataset across phylogenetically independent endosymbionts. In order to expand this comparative framework, we will explore patterns of gene loss and acquisition during the transition to both stable (*Blochmannia*) and transient (*Wolbachia*) genome interactions.

Genetic remodeling in stable genome interactions. *Blochmannia* is closely related to other endosymbionts and free-living bacterial species, yet acquired an intracellular lifestyle independently from *Buchnera* and many other endosymbionts. We are currently sequencing draft (5X) coverage of the small (810 kb; [264]) *Blochmannia* genome to explore unique opportunities for comparative genome analyses. We will compare the gene repertoires of *Blochmannia*, related insect symbionts (e.g. *Buchnera* and *Wigglesworthia*), and free-living bacteria such as *E. coli.* We will focus on the 'complementarity of gene repertoire' of host and symbiont metabolisms, as evidenced by pathways requiring gene products of both, and the retention of host-beneficial biosynthetic genes in tiny endosymbiont genomes. We will employ well-established approaches such as the program Clusters of Orthologous Groups of proteins (COG) [265] to identify shared orthologs among genomes. In addition, we will trace the process of genome reduction by determining size, genetic content, and timing of deletions events. This general approach has illustrated large deletion events in *Buchnera* that were probably deleterious [266]. By comparing multiple endosymbiont lineages including *Blochmannia,* we will distinguish whether genome integration typically occurs via large deletions that encompass many loci, or the gradual loss of individual genes that are dispensable for life within a host cell.

Endosymbionts of mealybugs represent the only documented example of a bacteria (i.e., γ•subdivision Proteobacteria) living within the cells of another bacterium (a larger β-subdivision Proteobacterial symbiont) [267]. This prokaryote-prokaryote binary association is a unique analog to the type of ancient genome interactions that led to the eukaryotic cell and equipped it with organelles. We will test for genome reduction as a consequence of this prokaryotic integration by sizing the genomes of both bacteria. A mixed bacterial sample will be prepared

from lab-reared insects through a homogenization and filtration technique [268] that we successfully applied to aphids, ants, *Drosophila*, and other insects [264, 269]. Bacterial samples will be embedded into agarose plugs, and DNA will be digested with I-*Ceu1*, a homing enzyme that cuts at a specific site within the bacterial 16S rRNA gene. Linear genomic fragments will be resolved by Pulsed Field Gel Electrophoresis (PFGE) using a Biorad CHEF Mapper XA in the Bavis lab. Genomic fragments will be assigned to the γ• or•β•subdivision by amplifying the 16S rRNA gene from each fragment with group-specific PCR primers and determining the partial sequence of this gene [270]. Fragment sizes will be summed to calculate respective genome sizes. We will test for genome size contraction in the γ-subdivision symbiont compared to the close relative β•subdivision symbiont to parallel the genome size of the β•subdivision symbiont to its close relative *Burkholderia pseudomallei* (7.25 Mb) and by comparing the γ-subdivision with free-living bacterial relatives such as *E. coli* (genome size of 4.5-5.5 Mb; [271]).

Gene content and phenotypic plasticity of *Wolbachia*. *Wolbachia* exhibit the most extensive plasticity in genome size of any intracellular bacteria. The genome of the nematode *Wolbachia* (~1.0 Mb) is 30% smaller than that of arthropod *Wolbachia* (1.66 Mb) [263]. We aim to decode the evolutionary processes that shape this genome plasticity. We will first determine the direction of lifestyle transitions by resolving phylogenetic relationships among the six *Wolbachia* subgroups (A and B – arthropod parasites; C and D – nematode mutualists; E and F – arthropod, unknown lifestyle) and determining the root of this tree. To this end, we will sequence five genes from each of 30 lineages spanning representatives of the A-F *Wolbachia* subgroups. The root of the tree will be determined statistically by comparing the likelihood of phylogenies with alternative root positions.

Previous genome size differences among *Wolbachia* strains are based on just four isolates. We will size genomes of phylogenetically diverse *Wolbachia* using PFGE. DNA from *Wolbachia* will be isolated, digested, and analyzed by a purification technique that is previously published [263] and has been successfully used in our lab. Genome size changes will be mapped across the *Wolbachia* phylogeny to correlate gene loss or acquisition with switches in host taxa or interaction type. We predict that the beneficial and stable *Wolbachia* engaged in long-term associations with nematode hosts will have smaller genomes due to reduced opportunities for lateral transfer, small effective population sizes, and high mutation rates [272], while the parasitic and more transient *Wolbachia* will have larger genomes due to frequent recombination and lateral gene transfer [273]. We will then reconstruct the history of genome fluidity throughout the diversification of *Wolbachia* by mapping gene loss and acquisition events onto the *Wolbachia* phylogeny. Results will advance our understanding of the importance of genome fluidity in this unstable endosymbiosis and the timing of genome-level changes.

Identifying 'key symbiotic genes' that catalyze genome-genome interactions. Functions that mediate genome-genome interactions reflect the evolutionary duration and context of a particular association [274]. Insect endosymbionts rely on distinct evolutionary strategies that span the supply of host-beneficial compounds to horizontal infection and reproductive parasitism. Genomes engaged in such distinct strategies are likely to vary in the types of 'key symbiotic genes' they encode, defined here as functions that catalyze the establishment and/or maintenance of a host-symbiont interaction. We will identify candidates for key symbiotic genes in *Blochmannia* and *Wolbachia* based on the retention of loci in these small endosymbiont genomes, their distribution across strains, and inferred functions. As a second step, we will confirm the expression of these loci by analyzing bacterial mRNA isolated directly from natural host populations.

Key genes shaping stable genome interactions. The hosts of beneficial, stable endosymbionts may provide proteins or cell structures to facilitate symbiont transmission to new cells. In this case, functions of key symbiotic genes may not involve cell invasion, but might

benefit the host and thus provide an evolutionary 'incentive' for stable symbiont transmission [258-260]. We will identify genes that are specifically retained in the small *Blochmannia* genome and may function to maintain its long-term interaction with ants. We will target genes related to proposed functions of *Blochmannia*, such as nutrient biosynthesis, pathogen defense, or pheromone production. Our preliminary data for the *Blochmannia* genome (460 kb of this endosymbiont sequence) indicates a high frequency and diversity of amino acid biosynthetic genes and suggest this symbiont plays a nutritional role. Draft coverage of the *Blochmannia* genome will reveal many putative symbiotic genes, of which we will identify ~20 as top candidates. We will perform real-time RT-qPCR (reverse transcriptase quantitative Polymerase Chain Reaction) to quantify expression patterns of those 20 candidate symbiotic genes, relative to expression of a *Blochmannia* housekeeping gene as an internal standard to control for variation in the number of endosymbionts infecting a host. RT-qPCR allows extremely sensitive detection of relative mRNA levels in a mixed sample.

Even in stable associations, variation at key symbiotic genes may drive ecological diversification within a given host group. We will explore genome variation among *Blochmannia* strains as it contributes to the wide ecological diversity of its host genus, the ant *Camponotus*, a dominant player in many ecosystems [275]. We will use Suppressive Subtractive Hybridization (SSH) to compare gene contents of *Blochmannia* associated with three distinct ant species. SSH is a cost-effective alternative to full genome sequencing when one aims to identify genes that are present in one genome (the 'tester' DNA), but absent from another (the 'driver' DNA)[276]. Briefly, tester and driver genomes are fragmented and hybridized, the hybrid molecules are removed, and the remaining (unhybridized) fragments represent tester-specific DNA. Since relatively low quantities of bacterial DNA can be isolated directly from hosts, we will use a modified SSH in which test-specific DNA is amplified by PCR (Clontech PCR-Select Bacterial Genome Subtraction Kit). We will generate four libraries from two reciprocal crosses (*Blochmannia* of *C. pennsylvanicus*, versus each of *C. festinatus* and *C. castaneus*), and will sequence these libraries to identify genes that are unique to a given *Blochmannia* strain. We predict that *Blochmannia* genomes will differ primarily at key symbiotic genes that allow their respective hosts to exploit distinct niches.

Key symbiotic genes as drivers of *Wolbachia* diversity. Recent data reveal unusual levels of genome fluidity in *Wolbachia*, including large differences in genome sizes [263], presence of mobile genetic elements [277], lateral gene transfer, recombination, and symbiont transfer to new hosts [273]. We will compare gene contents using DNA microarrays, a technique that can circumvent the laborious and expensive sequencing projects of multiple related genomes [278]. We will investigate variability of all *Wolbachia* ORFs (including housekeeping genes and mobile elements) to quantify the extent of gene loss and acquisition at an extensive scale. One particular focus will be the identification of key symbiotic genes that mediate distinct *Wolbachia*-host interactions. Arrays will be constructed from PCR-amplified products of all coding genes in the 1.3 Mb genome of *Wolbachia* infecting *Drosophila melanogaster*, either within the Bay Paul Center NASA-funded microarray facility or through an active collaboration (J. Werren, Univ. Rochester, pers. comm.). Genomic DNA of 20-30 *Wolbachia* strains will be hybridized to arrays at the MBL.

Analyses of arrays will address several components of genome flux in *Wolbachia*. We will identify genes that are shared across the 20-30 *Wolbachia* lineages hybridized against the array, but absent in larger α-Proteobacterial genomes that are not engaged in obligate endosymbiosis (i.e., the 5.7 Mb genome of *Agrobacterium tumefaciens*, the 9.1 Mb genome of *Bradyrhizobium japonicum*), as candidates for functions required in the establishment of this genome-genome interaction, and as targets of further gene expression analysis. In addition, we will explore gene loss and acquisition throughout the evolutionary history of this endosymbiosis, by (i) comparing genome fluidity in harmful and beneficial *Wolbachia* strains (ii) determining

the evolutionary history of specific gene loss or acquisition events across the *Wolbachia* phylogeny (iii) deciphering the size of gene loss and acquisition events (i.e., 1kb or 100kb changes in genome size) and (iv) identifying genes that correlate with interaction type (reproductive parasitism versus mutualism) transmission mode (strict vertical transmission versus occasional horizontal transmission), and distinct host taxa. This will be the first study to compare gene contents across closely related endosymbiont strains engaged in diverse interaction types.

Genome stasis versus lateral transfer in genome-genome interactions. Darwinian evolutionary processes of mutation and selection have shaped all of biology; however, such processes may operate quite differently in microbes. For example, transposons, plasmids, bacteriophages and other GTAs (Gene Transfer Agents) drive frequent lateral gene transfer in many prokaryotes. Such transfer can move genes that are already refined by natural selection in the donor genome and thus allow rapid exploitation of new niches in the recipient [279]. Persistent gene transfer between intracellular prokaryotes or organelles and the eukaryotic nucleus plays a vital role in genome integration [280-282].

Genome stasis versus fluidity in stable associations. In contrast to frequent gene transfer in free-living bacteria and unstable symbionts, genomes of certain stable endosymbionts lack GTAs and show no evidence of lateral gene transfer. Without gene transfer, their genome reduction is irreversible and may impose severe constraints on evolutionary potential. For example, two *Buchnera* strains have undergone no gene rearrangements or acquisitions during the 50–70 MY since their divergence, in contrast to a 2,000-fold greater genome lability in the related enterobacteria [259]. This genome stasis may reflect a lack of molecular tools for transfer (e.g., GTAs, recombination functions, and repeated DNA sequences) and limited opportunities to recombine with genetically distinct bacteria. However, certain features of the *Blochmannia* genome suggest this stable endosymbiont may be prone to rare lateral transfer. The *Blochmannia* chromosome is about 25% larger than related endosymbionts (*Buchnera* or *Wigglesworthia*) and might contain repeated elements, phage, or recombination functions. *Blochmannia*-filled host cells line the midgut where other bacterial species may occur. Other intracellular endosymbionts, including *Wolbachia,* may coexist with *Blochmannia* in the ant ovary.

We will characterize genome fluidity in *Blochmannia* to ascertain if even rare gene transfer influences this stable symbiosis. Based on our full genome analyzes we will assess the potential for gene exchange by comparing recombination functions (*rec* genes) and mobile elements in *Blochmannia* and other small bacterial genomes. Candidates for acquired genes would include those unique to *Blochmannia* or those with unusual codon usage or base compositions [283]. We will also use a phylogenetic approach to test for transfer of key symbiotic genes among *Blochmannia* strains. Congruent phylogenies of *Blochmannia* housekeeping genes and host genes indicate that most bacterial loci are evolutionary stable and track host diversification [284]. However, gene transfer of key symbiotic loci might experience strong selection at the host level if it expands the host's ecological range. We will identify up to 10 symbiotic genes for PCR amplification and sequencing across symbionts of several host species to test for horizontal transfer using phylogenetic approaches.

***Wolbachia* bacteriophage and other GTAs.** In order to understand mechanisms of lateral gene transfer in unstable symbioses, we will identify GTAs that mediate large-scale genomic changes in *Wolbachia*. The bacteriophage 'WO' has been associated with *Wolbachia* for at least 60 MY and undergoes high levels of horizontal transfer [281]. However, we do not know if WO and other GTAs mobilize bacterial and/or host genes and drive evolutionary innovation. We will investigate these features of WO and uncharacterized GTAs (transposable elements, S.L. O'Neill pers. comm.) by comparing the distribution of these mobile elements

across *Wolbachia* lineages with varied lifestyles. We will first determine the abundance and distribution of GTAs across *Wolbachia* genomes by screening 30 *Wolbachia* lineages with PCR. Primers will be designed from mobile elements identified in fully sequenced *Wolbachia* genomes. Two sequenced genomes (of *Wolbachia* group A and D) are currently available via BLAST servers (http://tigrblast.tigr.org/ufmg/index.cgi?database=wolbachia, http://tools.neb.com/wolbachia/search.html) and annotated versions will be publicly available in mid-2003 (S. O'Neill, pers. comm.). We will sequence two ORFs per GTA across all strains. We will then compare phylogenies of GTAs and chromosomal genes to evaluate the frequency of lateral transfer of each GTA. We will assess recombination within GTAs by comparing the gene phylogenies of the two ORFs sampled for each. Furthermore, if these mobile elements undergo lateral transfer between *Wolbachia* genomes, we expect them to be transcriptionally active (i.e., to encode mRNA products). We will use quantitative PCR approaches to determine the gene expression levels and copy numbers of GTAs across six *Wolbachia* subgroups relative to a single-copy housekeeping gene. Expression data will be correlated with distribution data to link GTA expression to its ability to spread throughout *Wolbachia*. We will compare *Wolbachia* strains with respect to GTA copy numbers and the fraction of the genome devoted them.

Lateral gene transfer can also move genes from the endosymbiont to the host nucleus, as exemplified by the transfer of mitochondrial and chloroplast genes to the eukaryotic nuclear genome [285]. If such transfer characterizes intracellular associations, we expect to find signals of similar events in younger endosymbioses. *Wolbachia* is a particularly promising system to investigate this phenomenon given its abundance, the presence of genetic machinery for lateral transfer in the genome, and genetic evidence of a recent lateral transfer event between *Wolbachia* and a beetle host [280]. We will screen host DNA for *Wolbachia* genes in order to test for endosymbiont gene transfer to the eukaryotic nucleus. We have identified 10-15 *Wolbachia*-infected host systems in which eliminating *Wolbachia* does not adversely affect the host. We will eliminate *Wolbachia* from these hosts, isolate host genomic DNA from uninfected individuals, and hybridize this host DNA against the *Wolbachia* microarray to test for endosymbiont gene transfer to the host. *Caveat:* Only recent cases of transfer will be detected, since a *Wolbachia* gene transferred to the host nucleus in the distant evolutionary past might have a different base composition that would prevent its detection through microarray hybridization.

Deciphering the language of genome-genome communication: Modulation of gene expression in endosymbionts. As endosymbiont genomes continuously streamline, changes in gene expression may become increasingly important relative to genome content differences. Currently, we do not understand how such changes in expression are linked to physiological or other changes in the host, nor do we know whether changes in gene copy numbers or regulatory functions typically account for expression changes. To understand the relevance of expression plasticity in the evolution of complex life, we will link symbiont expression patterns to natural environmental variation that the host experiences.

Plasticity of gene expression in stable genome interactions. Stable genome interactions can trigger the over-expression of chaperonins (*groEL*) as a mechanism to stabilize endosymbiont proteins [286], and up-regulation of biosynthetic or other genes according to the particular ecological requirements of the host [287]. We will explore links between *Blochmannia* gene expression and ecological and evolutionary novelty of their ant hosts. Specifically, we will quantify bacterial expression across several levels: distinct host species, host developmental stages, and different castes (queens, workers, etc). Ant castes represent the most extreme morphological, biochemical, and behavioral variation ever documented among genetically identical organisms and result from differential gene expression during larval stages [288]. *Blochmannia* might influence this process in ants, given its close involvement with larval development [289]. We will perform RT-qPCR of up to five candidate symbiotic genes (identified above). For any gene showing differential expression among treatments, we will

quantify the copy number of that locus using qPCR, in order to attribute expression changes to either gene duplication or regulatory changes.

Linking gene expression to variation of interaction types in unstable associations. We will perform the first study to identify gene expression profiles that underlie distinct interaction types among closely related endosymbionts. From the microarray analysis of *Wolbachia*, we will identify genes that are shared among diverse *Wolbachia* subgroups for further expression analysis. Although we will not exclude other promising candidates identified in the microarray analysis, we will focus primarily on genes encoding bacterial Outer Membrane Proteins (OMPs), which experience constant exposure to the eukaryotic host cellular environment and might mediate key host-symbiont interactions. Expression levels of several OMPs from up to 30 *Wolbachia* lineages from the major subgroups will be quantified using RT-qPCR, relative to a housekeeping gene as an internal control. Expression levels will be mapped onto gene phylogenies to determine when expression differences evolved and if changes associate with switches among host taxa or interaction types. Finally, we will examine DNA sequence variation at OMPs that show expression differences to test whether changes in gene regulation correspond with episodes of natural selection, such as positive selection resulting from an "arms race" between *Wolbachia* and the host.

5.5 SAMPLE PROCESSING, DNA SEQUENCING AND PHYLOGENETIC METHODOLOGIES

The projects described under research themes 5.1a, 5.1b and 5.2b require the collection of samples from field sites and isolation of cultures and/or extraction of DNA for molecular analyses. All three of these projects will share access to collected samples, DNA preparations and will work together to build an integrated database. We have established a DNA extraction laboratory and high throughput genomics capability that allows expanded experimental design well beyond what might be achieved in a more traditional laboratory setting. With high-throughput robotics and efficient DNA sequencing capabilities, molecular-based analyses of microbial populations can process 1000s of samples per day rather than 100s of samples analyzed over weeks if not months. With the technology available in the Bay Paul Center an investigator can afford to economically increase resolution and achieve better statistics (through inclusion of a larger number of sequence reads) in sequence-based studies of metabolic and microbial diversity. The very same technology allows for cost-effective genome level analyses of the sort described in themes 5.2a, and 5.4. The direct reagent cost for DNA sequencing in our environment is well below $1/sample and with anticipated improvements to our pipeline it would not be unreasonable to expect those costs to fall below $0.50/sample. One highly skilled technician oversees the operation of the facility but the workflow is performed by technicians associated with individual research projects. This allows for improved throughput without the cost of maintaining a large technical staff.

The acquisition of DNA sequence data is rapidly becoming a minor cost of molecular analyses. There has been an ever-increasing demand for computational capabilities for the purpose of converting DNA sequences into database entries and downstream processing in the form of database searches, annotation, sequence alignment and phylogenetic inferences. These tasks are common to all but one of the projects described in this proposal for continued membership in NASA's Astrobiology Institute. To address these requirements we have established a common bioinformatics "pipeline" that incorporates tools for sequence assembly using ARACHNE, programs such as CRITICA and GLIMMER for finding open reading frames, and scripts for controlling search algorithms that use FASTA, BLAST and other software tools to search both local and remote databases for annotating sequences. We perform

computationally expensive procedures such as the alignment of large datasets with CLUSTALX on a large Beowulf cluster or on a large memory, 8 gigabyte high performance Alpha processor. Similarly, we use a variety of phylogenetic packages including PAUP, MEGA, MrBayes, PUZZLE, etc. to infer phylogenetic histories for the sequences under study. Because different questions in molecular evolution can only be resolved by the most appropriate phylogenetic tool, it is well beyond the scope of this proposal or any other to summary all contingencies that might be encountered in analysis of molecular data.

5.6 MANAGEMENT PLAN.

This Astrobiology program involves four institutions including The Marine Biological Laboratory at Woods Hole, Woods Hole Oceanographic Institution, Brown University and the University of North Carolina. The project will be managed by a core group of investigators who represent each of the four institutions and major research themes. This management structure will include PI M.L. Sogin (PI MBL), Lorraine Olendzenski (EPO coordinator at MBL), and Co-Is from the collaborating institutions; Andreas Teske (UNC), Katrina Edwards (WHOI), and John Mustard (Brown University). All of the participating institutions have convenient access to video conferencing facilities capable of communicating over internet2. The management group will hold bi-monthly video-conferences. Dr. Sogin will serve as Chair and Dr. Andreas Teske will be the Vice-Chair. The management group and invited guests will assess progress and priorities regarding all Woods Hole Astrobiology research projects, make decisions regarding allocation of available funds and monitor budgeted spending**.** The committee will carefully monitor use of Astrobiology funds relative to time and effort for personnel supported by this interdisciplinary project. Other responsibilities of the management committee will include the pursuit of additional funding for new projects that are relevant to Astrobiology, making decisions about including new members or research groups within this Astrobiology program, and coordinating attendance and participation of investigators associated with the Marine Biology Laboratory Astrobiology team at NASA sponsored and national meetings.

On a bi-annual basis, the management committee will meet with all participants in the Woods Hole Astrobiology initiative. During those meetings, research funded by the Astrobiology program will be presented and openly discussed. Each summer, an Astrobiology retreat will be held in the Woods Hole area for all participating research groups and their associated personnel. The objective of this retreat will be to stimulate discussion on all Astrobiology research topics, giving researchers an opportunity to present their results and receive feedback, and nurture emerging opportunities for collaborative research in Astrobiology.

The PI will oversee administrative aspects of the Astrobiology program including management of the high-throughput DNA sequencing facility. Dr. Sogin will direct the workshop **Advances in Genome Sciences and Bioinformatics.** Administrative assistants in the Bay Paul Center provide secretarial and purchasing support for team members.

6. REFERENCES

1. Schopf, J.W. and B.M. Packer, *Early Archean (3.3-billion to 3.5-billion-year-old) microfossils from Warrawoona Group, Australia.* Science, 1987. **237**: p. 70-3.
2. Schopf, J.W., et al., *Laser--Raman imagery of Earth's earliest fossils.* Nature, 2002. **416**(6876): p. 73-6.
3. Brasier, M.D., et al., *Questioning the evidence for Earth's oldest fossils.* Nature, 2002. **416**(6876): p. 76-81.
4. Mojzsis, S.J., et al., *Evidence for life on Earth before 3,800 million years ago.* Nature, 1996. **384**(6604): p. 55-9.
5. Shen, Y., R. Buick, and D.E. Canfield, *Isotopic evidence for microbial sulphate reduction in the early Archaean era.* Nature, 2001. **410**(6824): p. 77-81.
6. Brocks, J.J., et al., *Archean molecular fossils and the early rise of eukaryotes.* Science, 1999. **285**(5430): p. 1033-6.
7. Barnes, R.S.K., *A Synoptic Classification of Living Organisms*. 1984, Oxford: Blackwell Scientific Publications.
8. Lipps, J.H., *Fossil Prokaryotes and Protists.* 1993, Boston: Blackwell Scientific Publications.
9. Garcia-Pichel, F., *Cyanobacteria*, in *Encyclopedia of Microbiology*, J. Lederberg, Editor. 2000, Academic Press: San Diego.
10. Atlas, R.M. and R. Bartha, *Microbial Ecology. Fundamentals and Applications*. 1993, Redwood City: The Benjamin/Cummings Publishing Company, Inc.
11. Knoll, A.H. and D.E. Canfield, *Isotopic inferences on early ecosystems*, in *Isotope Paleobiology and paleoecology. The Palaeontological Society Papers*, R.D. Norris and R.M. Corfield, Editors. 1998, Palaeontological Society. p. 212-243.
12. Canfield, D.E., K.S. Habicht, and B. Thamdrup, *The Archean sulfur cycle and the early history of atmospheric oxygen.* Science, 2000. **288**(5466): p. 658-61.
13. Habicht, K.S., et al., *Calibration of sulfate levels in the archean ocean.* Science, 2002. **298**(5602): p. 2372-4.
14. Strauss, H. and T. Moore, *Abundances and isotopic compositions of carbon and sulfur species in whole rock and kerogen samples*, in *The Proterozoic biosphere*, J.W. Schopf and C. Klein, Editors. 1992, Cambridge University Press: Cambridge. p. 709-789.
15. Hayes, J.M., *Global Methanotrophy at the Archaean-Proterozoic transition*, in *Early Life on Earth*, S. Bengtson, et al., Editors. 1994, Columbia University Press: New York. p. 220-239.
16. Des Marais, D.J., et al., *Carbon isotope evidence for the stepwise oxidation of the Proterozoic environment.* Nature, 1992. **359**(6396): p. 605-9.
17. Parkes, R.J., B.A. Cragg, and P. Wellsbury, *Recent studies on bacterial populations and processes in subseafloor sediments: A review.* Hydrogeology J., 2000. **8**: p. 11-28.
18. Fisk, M.R., S.J. Giovannoni, and I.H. Thorseth, *Alteration of oceanic volcanic glass: textural evidence of microbial activity.* Science, 1998. **281**(5379): p. 978-80.
19. Wagner, M., et al., *Phylogeny of dissimilatory sulfite reductases supports an early origin of sulfate respiration.* J Bacteriol, 1998. **180**(11): p. 2975-82.

20. Klein, M., et al., *Multiple lateral transfers of dissimilatory sulfite reductase genes between major lineages of sulfate-reducing prokaryotes.* J Bacteriol, 2001. **183**(20): p. 6028-35.
21. Larsen, O., T. Lien, and N.K. Birkeland, *Dissimilatory sulfite reductase from Archaeoglobus profundus and Desulfotomaculum thermocisternum: phylogenetic and structural implications from gene sequences.* Extremophiles, 1999. **3**(1): p. 63-70.
22. Elsgaard, L., et al., *Microbial sulfate reduction in deep-sea sediments at the Guaymas basin hydrothermal vent area: Influence of temparture and substrates.* Geochim. Cosmochim. Acta, 1994. **58**: p. 3335-3343.
23. Jorgensen, B.B., L.X. Zawacki, and H.W. Jannasch, *Thermophilic bacterial sulfate reduction in deep-sea sediments at the Guaymas Basin hydrothermal vents (Gulf of California).* Deep-Sea Res I, 1990. **37**: p. 695-710.
24. Jorgensen, B.B., M.F. Isaksen, and H.W. Jannasch, *Bacterial sulfate reduction above 100ºC in deep-sea hydrothermal vent systems.* Science, 1992. **258**: p. 1756-1757.
25. Weber, A. and B.B. Jorgensen, *Bacterial sulfate reduction in hydrothermal sediments of the Guaymas Basin, Gulf of California, Mexico.* Deep-Sea Research I, 2002. **49**: p. 827-841.
26. Rueter, P., et al., *Anaerobic oxidation of hydrocarbons in crude oil by new types of sulphate-reducing bacteria.* Nature, 1994. **372**(6505): p. 455-8.
27. Burggraf, S., et al., *Archaeoglobus profundus sp. nov. represents a new species within the sulfate-reducing archaebacteria.* Syst. Appl. Microbiology, 1990. **13**: p. 24-28.
28. Teske, A., et al., *Microbial diversity of hydrothermal sediments in the Guaymas Basin: evidence for anaerobic methanotrophic communities.* Appl Environ Microbiol, 2002. **68**(4): p. 1994-2007.
29. Dhillon, A., et al., *Molecular characterization of sulfate-reducing bacteria in the Guaymas Basin.* Appl. Environ. Microbiol., 2003. **In Revision**.
30. Chistoserdova, L., et al., *C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea.* Science, 1998. **281**(5373): p. 99-102.
31. Lueders, T., et al., *Molecular analyses of methyl-coenzyme M reductase alpha-subunit (mcrA) genes in rice field soil and enrichment cultures reveal the methanogenic phenotype of a novel archaeal lineage.* Environ Microbiol, 2001. **3**(3): p. 194-204.
32. Ramakrishnan, B., et al., *Archaeal community structures in rice field soils from different geographical regions before and after initiation of methane production.* FEMS Microbiol Ecol, 2001. **37**: p. 175-186.
33. Springer, E., et al., *Partial gene sequences for the A subunit of methyl-coenzyme M reductase (mcrI) as a phylogenetic tool for the family Methanosarcinaceae.* Int J Syst Bacteriol, 1995. **45**(3): p. 554-9.
34. Hoehler, T.M., et al., *Field and laboratory studies of methane oxidation in an anoxic marine sediment: evidence for a methanogen-sulfate reducer consortium.* Global Biogeochem. Cycles, 1994. **8**: p. 451-463.
35. Boetius, A., et al., *A marine microbial consortium apparently mediating anaerobic oxidation of methane.* Nature, 2000. **407**(6804): p. 623-6.
36. Orphan, V.J., et al., *Comparative analysis of methane-oxidizing archaea and sulfate-reducing bacteria in anoxic marine sediments.* Appl Environ Microbiol, 2001. **67**(4): p. 1922-34.
37. Orphan, V.J., et al., *Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis.* Science, 2001. **293**(5529): p. 484-7.

38. Orphan, V.J., et al., *Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments.* Proc Natl Acad Sci U S A, 2002. **99**(11): p. 7663-8.
39. Hoehler, T.M. and M.J. Alperin, *Anaerobic methane oxidation by a methanogen-sulfate reducer consortium: geochemical evidence and biochemical considerations*, in *Microbial growth on C1-compounds*, M.E. Lidstrom and F.R. Tabita, Editors. 1996, Kluwer Academic Publishers: The Netherlands. p. 326-333.
40. Jones, W.J., C.E. Stugard, and H.W. Jannasch, *Comparison of thermophilic methanogens from submarine vent sites.* Arch. Microbiol., 1989. **151**: p. 314-318.
41. Canganella, F. and W.J. Jones, *Microbial characterization of thermophilic archaea isolated from the Guaymas Basin hydrothermal vent.* Current Microbiology, 1994. **28**: p. 299-306.
42. Kurr, M.R., et al., *Methanopyrus kandleri, gen. and sp. nov. represents a novel group of hyperthermophilic methanogens, growing at 110°C.* Arch. Microbiol., 1991. **156**: p. 239-247.
43. Welhan, J.A., *Origins of methane in hydrothermal systems.* Chem. Geol., 1988. **71**: p. 183-198.
44. D'Hondt, S., S. Rutherford, and A.J. Spivack, *Metabolic activity of subsurface life in deep-sea sediments.* Science, 2002. **295**(5562): p. 2067-70.
45. Cragg, B., et al., eds. *Bacterial biomass and activity profiles within deep sediment layers.* Proc. ODP Sci. Results, ed. R.v.H. E. Suess, et al. (eds.). 1990: College Station. 607-619.
46. Cragg, B.A. and A.E.S. Kemp, *Bacterial profiles in deep sediment layers from the eastern Equatorial Pacific Ocean, site 851*, in *Proc. ODP. Sci. Results*, N.G. Pisias, et al., Editors. 1995: College Station. p. 599-604.
47. Daims, H., et al., *The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set.* Syst Appl Microbiol, 1999. **22**(3): p. 434-44.
48. Stokes, H.W., et al., *Gene cassette PCR: sequence-independent recovery of entire genes from environmental DNA.* Appl Environ Microbiol, 2001. **67**(11): p. 5240-6.
49. Rondon, M.R., et al., *Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms.* Appl Environ Microbiol, 2000. **66**(6): p. 2541-7.
50. Murray, A., et al. *Utilization of a large-contig DNA library for exploring Antarctic picoplankton diversity and as a template for microarrays*. in *Microbial Observatory / LeXen. PI workshop*. 2002. Arlington, VA: National Science Foundation.
51. Beja, O., et al., *Comparative genomic analysis of archaeal genotypic variants in a single population and in two different oceanic provinces.* Appl Environ Microbiol, 2002. **68**(1): p. 335-45.
52. Beja, O., et al., *Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage.* Environ Microbiol, 2000. **2**(5): p. 516-29.
53. Ward, P.D. and D. Brownlee, *Rare Earth. Why Complex Life is Uncommon in the Universe.* 2000, New York: Springer Verlag.
54. Bada, J.L. and A. Lazcano, *Origin of life. Some like it hot, but not the first biomolecules.* Science, 2002. **296**(5575): p. 1982-3.
55. Huber, C. and G. Wachtershauser, *Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions.* Science, 1997. **276**(5310): p. 245-7.

56. Ueno, Y., et al., *Ion microprobe analysis of graphite from ca. 3.8 Ga metasediments, Isua suprcrustal belt, West Greenland: Relationship between metamorphism and carbon isotopic composition.* Geochim. Cosmochim. Acta, 2002. **66**: p. 1257-1268.
57. Rosing, M.T., *13C-Depleted carbon microparticles in >3700-Ma sea-floor sedimentary rocks from west greenland.* Science, 1999. **283**(5402): p. 674-6.
58. van Zuilen, M.A., A. Lepland, and G. Arrhenius, *Reassessing the evidence for the earliest traces of life.* Nature, 2002. **418**(6898): p. 627-30.
59. Cody, G.D., et al., *Geochemical roots of autotrophic carbon fixation: Hydrothermal experiments in the system citric acid, H2O-(±FeS)-(±NiS).* Geochim Cosmochim, 2001. **65**: p. 3357-3576.
60. Wachtershauser, G., *Evolution of the first metabolic cycles.* Proc Natl Acad Sci U S A, 1990. **87**(1): p. 200-4.
61. Romano, A.H. and T. Conway, *Evolution of carbohydrate metabolic pathways.* Res Microbiol, 1996. **147**(6-7): p. 448-55.
62. Russell, M.J. and A.J. Hall, *The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front.* J Geol Soc London, 1997. **154**(3): p. 377-402.
63. Martin, W. and M.J. Russell, *On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells.* Philos Trans R Soc Lond B Biol Sci, 2003. **358**(1429): p. 59-85.
64. Nisbet, E.G. and N.H. Sleep, *The habitat and nature of early life.* Nature, 2001. **409**(6823): p. 1083-91.
65. Wachtershauser, G., *Life in a ligand sphere.* Proc Natl Acad Sci U S A, 1994. **91**(10): p. 4283-7.
66. Maden, B.E., *No soup for starters? Autotrophy and the origins of metabolism.* Trends Biochem Sci, 1995. **20**(9): p. 337-41.
67. Kasting, J.F., *Earth's early atmosphere.* Science, 1993. **259**(5097): p. 920-6.
68. Ferris, J.P., et al., *Synthesis of long prebiotic oligomers on mineral surfaces.* Nature, 1996. **381**(6577): p. 59-61.
69. Pace, N.R., *Origin of life--facing up to the physical setting.* Cell, 1991. **65**(4): p. 531-3.
70. Zahnle, K.J., J.F. Kasting, and J.B. Pollack, *Evolution of a steam atmosphere during Earth's accretion.* Icarus, 1988. **74**: p. 62-97.
71. Sleep, N.H., K. Zahnle, and P.S. Neuhoff, *Initiation of clement surface conditions on the earliest Earth.* Proc. Nat. Acad. Sc, 2001. **98**: p. 3666-3672.
72. Farmer, J., *Thermophiles, early biosphere, and the origin of life on Earth: Implications for the exobiological exploration of Mars.* J. Geoph. Res., 1998. **103**: p. 28,457-28,461.
73. House, C.H. and S.T. Fitz-Gibbon, *Using homolog groups to create a whole-genomic tree of free-living organisms: an update.* J Mol Evol, 2002. **54**(4): p. 539-47.
74. Stetter, K.O., *Hyperthermophilic prokaryotes.* FEMS Microbiol Rev, 1996. **18**: p. 149-158.
75. Vargas, M., et al., *Microbiological evidence for Fe(III) reduction on early Earth.* Nature, 1998. **395**(6697): p. 65-7.
76. Brochier, C. and H. Philippe, *Phylogeny: a non-hyperthermophilic ancestor for bacteria.* Nature, 2002. **417**(6886): p. 244.
77. Galtier, N., N. Tourasse, and M. Gouy, *A nonhyperthermophilic common ancestor to extant life forms.* Science, 1999. **283**(5399): p. 220-1.

78. Sleep, N.H. and K. Zahnle, *Refugia from asteroid impacts on early Mars and the early Earth.* J. Geophys. Res., 1998. **103**: p. 28,529-28,544.
79. Gogarten-Boekels, M., E. Hilario, and J.P. Gogarten, *The effects of heavy meteorite bombardment on the early evolution--the emergence of the three domains of life.* Orig Life Evol Biosph, 1995. **25**(1-3): p. 251-64.
80. Madigan, M.T., J.M. Martinko, and J. Parker, *Brock Biology of Microrganisms*. 2003, Prentice Hall, Pearson Education: Upper Saddle River, NJ.
81. Raven, J.A., *Inorganic carbon assimilation by marine biota.* J. Exp. Mar. Biol. Ecol., 1996. **203**: p. 39-47.
82. Fuchs, G., *Alternative pathways of autotrophic CO2 fixation,*, in *Autotrophic Bacteria*, S.H. G. and B. Bowien, Editors. 1989, Springer Verlag: Berlin. p. 365-382.
83. Schauder, R., F. Widdel, and G. Fuchs, *Carbon assimilation pathways in sulfate-reducing bacteria II. Enzymes of a reductive citric acid cycle in the autotrophic Desulfobacter hydrogenophilus.* Arch. Microbiol, 1987. **148**: p. 218-225.
84. Beh, M., et al., *Enzymes of the reductive citric acid cycle in the autotrophic eubacterium Aquifex pyrophilus and in the archaebacterium Thermoproteus neutrophilus.* Arch. Microbiol. 160:306-311., 1993.
85. Shiba, H., et al., *The CO2 assimilation via the redcutive tricarboxylic acid cycle in an obligately autotrophic aerobic hydrogen-oxidizing bacterium, Hydrogenobacter thermophilus.* Arch Microbiol, 1985. **141**: p. 198-203.
86. Menendez, C., et al., *Presence of acetyl coenzyme A (CoA) carboxylase and propionyl-CoA carboxylase in autotrophic Crenarchaeota and indication for operation of a 3-hydroxypropionate cycle in autotrophic carbon fixation.* J Bacteriol, 1999. **181**(4): p. 1088-98.
87. Vorholt, J.A., et al., *Pathways of autotrophic CO2 fixation and of dissimilatory nitrate reduction to N2O in Ferroglobus placidus.* Arch Microbiol, 1997. **167**(1): p. 19-23.
88. Vorholt, J., et al., *Enzymes and coenzymes of the carbon monooxide dehydrogenase pathway for autotrophic CO_2 fixation in Archaeoglobus lithotrophicus and the lack of of carbon monoxide dehydrogenase in the heterotrophic A. profundus.* Archives of Microbiology, 1995. **163**: p. 122-118.
89. Jahnke, L.L., et al., *Signature lipids and stable carbon isotope analyses of Octopus Spring hyperthermophilic communities compared with those of Aquificales representatives.* Appl Environ Microbiol, 2001. **67**(11): p. 5179-89.
90. Kuypers, M.M., et al., *Massive expansion of marine archaea during a mid-Cretaceous oceanic anoxic event.* Science, 2001. **293**(5527): p. 92-5.
91. van der Meer, M.T.J., et al., *Autotrophy of green non-sulphur bacteria in hot spring microbial mats: biological explanations for isotopically heavy organic carbon in the geological record.* Environmental Microbiology, 2000. **2**(4): p. 428-435.
92. Buchanan, B.B. and D.I. Arnon, *A reverse KREBS cycle in photosynthesis: consensus at last.* Photosynth Res, 1990. **24**: p. 47-53.
93. Evans, M.C., B.B. Buchanan, and D.I. Arnon, *A new ferredoxin-dependent carbon reduction cycle in a photosynthetic bacterium.* Proc Natl Acad Sci U S A, 1966. **55**(4): p. 928-34.
94. Fuchs, G., E. Shipperich, and G. Eden, *Autotrophic CO2 fixation in Chlorobium limicola. Evidence for the operation of a reductive tricarboxylic acid cycle in growing cells.* Arch Microbiol, 1980. **128**: p. 64-71.

95. Ivanovski, R.N., N.V. Sintsov, and E.N. Kontratieva, *ATP linked citrate lyase activity in the green sulfur bacterium Chlorobium limicola forma thiosulfatophilum.* Arch Microbiol, 19080. **128**: p. 239-241.
96. Schäfer, S., C. Barkowski, and G. Fuchs, *Carbon assimilation by the autotrophic thermophilic archaebacterium Thermoproteus neutrophilus.* Arch Microbiol, 1986. **146**: p. 301-308.
97. Rasmussen, B., *Filamentous microfossils in a 3,235-million-year-old volcanogenic massive sulphide deposit.* Nature, 2000. **405**(6787): p. 676-9.
98. Baross, J.A. and S.E. Hoffman, *Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life.* Orig. Life, 1985. **15**: p. 327.
99. Corliss, J.B., J.A. Baross, and S.E. Hoffman, *An hypothesis concerning the relationship between submarine hot springs and the origin of life.* Ocean Acta, 1981. **4 (Suppl)**: p. 59-69.
100. Sievert, S.M., et al., *Unpublished data.* 2003.
101. Wirsen, C.O., et al., *Characterization of an autotrophic sulfide-oxidizing marine Arcobacter sp. that produces filamentous sulfur.* Appl Environ Microbiol, 2002. **68**(1): p. 316-25.
102. Longnecker, K. and A. Reysenbach, *Expansion of the geographic distribution of a novel lineage of epsilon-Proteobacteria to a hydrothermal vent site on the Southern East Pacific Rise.* FEMS Microbiol Ecol, 2001. **35**(3): p. 287-293.
103. Fatland, B.L., et al., *Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-coenzyme A in Arabidopsis.* Plant Physiol, 2002. **130**(2): p. 740-56.
104. Kanao, T., et al., *ATP-citrate lyase from the green sulfur bacterium Chlorobium limicola is a heteromeric enzyme composed of two distinct gene products.* Eur J Biochem, 2001. **268**(6): p. 1670-8.
105. Eisen, J.A., et al., *The complete genome sequence of Chlorobium tepidum TLS, a photosynthetic, anaerobic, green-sulfur bacterium.* Proc Natl Acad Sci U S A, 2002. **99**(14): p. 9509-14.
106. Morris, G. and S.M. Sievert, *Unpublished data.* 2003.
107. Campbell, B. and S.C. Cary. *Molecular analysis of key autotrophic enzymes indicate the importance of the reductive TCA cycle in chemoautotrophic communities at deep-sea hydrothermal vents.* in *Aquatic Sciences Meeting of the American Society of Limnology and Oceanography. Poster presentation*. 2003.
108. Ishii, M., Y. Igarashi, and T. Kodama, *Purification and characterization of ATP:citrate lyase from Hydrogenobacter thermophilus TK-6.* J Bacteriol, 1989. **171**(4): p. 1788-92.
109. Deckert, G., et al., *The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.* Nature, 1998. **392**(6674): p. 353-8.
110. Huynen, M.A., T. Dandekar, and P. Bork, *Variation and evolution of the citric-acid cycle: a genomic perspective.* Trends Microbiol, 1999. **7**(7): p. 281-91.
111. Campbell, B., *pers. comm*. 2003.
112. Gotz, D., et al., *Persephonella marina gen. nov., sp. nov. and Persephonella guaymasensis sp. nov., two novel, thermophilic, hydrogen-oxidizing microaerophiles from deep-sea hydrothermal vents.* Int J Syst Evol Microbiol, 2002. **52**(Pt 4): p. 1349-59.
113. Zhang, C.L., et al., *Carbon isotopic fractionations associated with thermophilic bacteria Thermotoga maritima and Persephonella marina.* Environ Microbiol, 2002. **4**(1): p. 58-64.

114. Suzuki, F., et al., *Studies on ATP citrate lyase of rat liver. II. Immunochemical properties.* J Biochem (Tokyo), 1967. **62**(2): p. 170-8.
115. Shashi, K., A.K. Bachhawat, and R. Joseph, *ATP:citrate lyase of Rhodotorula gracilis: purification and properties.* Biochim Biophys Acta, 1990. **1033**(1): p. 23-30.
116. Rangasamy, D. and C. Ratledge, *Compartmentation of ATP:citrate lyase in plants.* Plant Physiol, 2000. **122**(4): p. 1225-30.
117. Nowrousian, M., et al., *The fungal acl1 and acl2 genes encode two polypeptides with homology to the N- and C-terminal parts of the animal ATP citrate lyase polypeptide.* Curr Genet, 2000. **37**(3): p. 189-93.
118. Adams, I.P., et al., *The distinctiveness of ATP:citrate lyase from Aspergillus nidulans.* Biochim Biophys Acta, 2002. **1597**(1): p. 36-41.
119. Ma, Y., et al., *ATP citrate lyase in the glaucocystophyte alga Cyanophora paradoxa is a cytosolic enzyme: characterisation of the gene for the large subunit at the cDNA and genomic levels.* Mol. Genet. Genomics, 2001. **266**(2): p. 231-238.
120. Chen, C. and M. Gibbs, *Some enzymes and properties of the reductive carboxylic acid cycle are present in the green alga Chlamydomonas reinhardtii F-60.* Plant. Physiol., 1991. **98**: p. 535-539.
121. Rzhetsky, A. and M. Nei, *A simple method for estimating and testing minimum-evolution trees.* Mol Biol Evol, 1992. **9**: p. 945-967.
122. Janssen, S., et al., *Ferredoxins from the archaeon Acidianus ambivalens: overexpression and characterization of the non-zinc-containing ferredoxin FdB.* Biol Chem, 2001. **382**(10): p. 1501-7.
123. Cowen, J.P., et al., *Fluids from Aging Ocean Crust That Support Microbial Life.* Science, 2003. **299**(5603): p. 120-123.
124. Carr, M.H., *Water on Mars*. 1996, New York: Oxford University Press. xi, 229.
125. Malin, M.C. and K.S. Edgett, *Mars Global Surveyor Mars Orbiter Camera: Interplanetary cruise through primary mission.* Journal of Geophysical Research-Planets, 2001. **106**: p. 23429-23570.
126. Mustard, J.F., C.C. D., and M.K. Rifkin, *Evidence for recent climate change on Mars from the identification of youthful near-surface ground ice.* Nature, 2001. **412**: p. 411-414.
127. Kreslavsky, M.A. and J.W. Head, *Mars: Nature and evolution of young latitude-dependent water-ice-rich mantle.* Geophysical Research Letters, 2002. **29**(15): p. art. no.-1719.
128. Kreslavsky, M.A. and J.W. Head, *Kilometer-scale roughness of Mars: Results from MOLA data analysis.* Journal of Geophysical Research-Planets, 2000. **105**(E11): p. 26695-26711.
129. Kreslavsky, M. and J. Head, *The fate of outflow channel effluents in the northern lowlands of Mars: The Vastitas Borealis Formation as a sublimation residue from frozen ponded bodies of water.* J. Geophys. Res., 2002. **107**: p. 5121.
130. Boynton, W.V., et al., *Distribution of hydrogen in the near surface of Mars: evidence for subsurface ice deposits.* Science, 2002. **297**(5578): p. 81-5.
131. Head, J.W. and D.R. Marchant, *Cold-based mountain glaciers on Mars: Western Arsia Mons.* Geology, 2003. **in press**.
132. Bebié, J. and M.A.A. Schoonen, *Pyrite and phosphate in anoxia and an origin-of-life hypothesis.* Earth and Planetary Science Letters, 1999. **171**: p. 1-5.

133. Blöchl, E., et al., *Reactions depending on iron sulfide and linking geochemistry with biochemistry.* Proceedings of the National Academy of Sciences, 1992. **89**: p. 8117-8120.
134. Wächtershäuser, G., *Before enzymes and templates: theory of surface metabolism.* Microbiol Rev, 1988. **52**(4): p. 452-84.
135. Russell, M.J., D.E. Daia, and A.J. Hall, *The emergence of life from FeS bubbles at alkaline hot springs in an acid ocean*, in *Thermophiles: The Keys to Molecular Evolution and the Origin of Life?*, J. Wiegel and M.W.W. Adams, Editors. 1998, T. J. International Ltd.: Padstow. p. 77-126.
136. Head, J.W. and L. Wilson, *Deep submarine pyroclastic eruptions: theory and predicted landorms and deposits.* Journal of volcanology and geothermal research, 2003. **12**: p. 155-193.
137. Head, J.W., L. Wilson, and D.K. Smith, *Mid-ocean ridge eruptive vent morphology and substructure: Evidence for dike widths, eruption rates, and evolution of eruptions and axial volcanic ridges.* Journal of Geophysical Research-Planets, 1996. **101**: p. 28265-28280.
138. Marshall, C.P. and R.W. Fairbridge, *Encyclopedia of Geochemistry*. 1999, Boston: Kluwer Academic Pub. 712.
139. Des Marais, D.J., *When did photosynthesis emerge on Earth?* Science, 2000. **289**: p. 1703-5.
140. Knoll, A.H., *Palynology: Principals and Applications*, J. Jansonius and D. Foundation, Editors. 1996, American Association of Stratigraphic Palynologists Foundation: Dallas. p. 51-80.
141. Des Marais, D.J., *Microbial mats and the early evolution of life.* Trends in Ecology and Evolution, 1990. **5**: p. 140-145.
142. Ehrenreich, A. and F. Widdel, *Anaerobic oxidation of ferrous iron by purple bacteria, a new type of phototrophic metabolism.* Appl Environ Microbiol, 1994. **60**(12): p. 4517-4526.
143. Konhauser, K.O., et al., *Could bacteria have formed the Precambrian banded iron formations?* Geology, 2002. **30**: p. 1079-1082.
144. Xiong, J., et al., *Molecular Evidence for the Early Evolution of Photosynthesis.* Science, 2000. **289**(5485): p. 1724-1730.
145. Ghiorse, W.C., *Biology of iron and manganese-depositing bacteria.* Annual Review of Microbiology, 1984. **38**: p. 515-550.
146. Edwards, K.J., W. Bach, and D.R. Rogers, *Geomicrobiology of the ocean crust: A role for chemoautotrophic Fe-Bacteria.* Biol Bull, 2003. **in press**.
147. Anbar, A.D. and A.H. Knoll, *Proterozoic Ocean Chemistry and Evolution: A Bioinorganic Bridge?* Science, 2002. **297**(5584): p. 1137-1142.
148. Francis, P., *Volcanoes: A Planetary Perspective*. 1998, Oxford: Oxford University Press. 443.
149. Emerson, D. and C.L. Moyer, *Neutrophilic Fe-oxidizing bacteria are abundant at the Loihi Seamount hydrothermal vents and play a major role in Fe oxide deposition.* Applied and Environmental Microbiology, 2002. **68**(6): p. 3085-3093.
150. Emerson, D. and C. Moyer, *Isolation and characterization of novel iron-oxidizing bacteria that grow at circumneutral pH.* Applied and Environmental Microbiolgy, 1997. **63**(12): p. 4784-4792.
151. Edwards, K.J., et al., *An archaeal iron-oxidizing extreme acidophile important in acid mine drainage.* Science, 2000. **287**: p. 1796-1799.

152. Amils, R., et al. *Geomicrobiology of an extreme acidic environment: the Tinto River case*. in *Amercian Geophysical Society*. 2000.
153. Amaral Zettler, L.A., et al., *Eukaryotic diversity in Spain's River of Fire.* Nature, 2002. **417**: p. 137.
154. Edwards, K.J., et al., *Seafloor bio-alteration of sulfide minerals: Results from in-situ incubation studies.* Geochim Cosmochim Acta, 2003. **in press**.
155. Benison, K.C. and D.A. LaClair, *Acid sedimentary environments on Mars? Possible terrestrial analogs*. 2002: The Geological Society of America, Dever Annual Meeting.
156. LaClair, D.A. and K.C. Benison, *Acid sedimentary environments on Mars? The physical sedimentology*. 2002: The Geological Society of America, Dever Annual Meeting.
157. Wyatt, M.B. and H.Y.J. McSween, *Spectral evidence for weathered basalt as an alternative to andesite in the northern lowlands of Mars.* Nature, 2002. **417**(6886): p. 263-266.
158. Bach, W. and K.J. Edwards, *Iron and sulfide oxidation within the basaltic ocean crust - extent, processes, timing, and implications for chemolithoautotrophic primary biomass production.* Geochimica Cosmochimica Acta, 2003. **in press**.
159. Emerson, D., *Microbial oxidation of Fe(II) and Mn(II) at circumneutral pH.*, in *Environmental Microbe-Metal Interactions*, D.R. Lovely, Editor. 2000, ASM Press: Washington, D.C.
160. Alt, J.C., *Hydrothermal oxide and nontronite deposits on seamounts in the eastern Pacific.* Mar Geol, 1988. **81**: p. 227-239.
161. Hannington, M.D. and I.R. Jonasson, *Fe and Mn oxides at seafloor hydrothermal vents*, in *Biomineralization processes of ir on and manganese; modern and ancient environments*, H.C. Skinner, Editor. 1992, Catena-Verlag: Cremlingen-Destedt. p. 351-370.
162. Karl, D.M., A.M. Brittain, and B.D. Tilbrook, *Hydrothermal and microbial processes at Loihi Seamount, a mid-plate hot-spot volcano.* Deep Sea Research Part I: Oceanographic Research Papers, 1989. **36**(11): p. 1655-1673.
163. Karl, D.M., et al., *Loihi Seamount, Hawaii: a mid-plate volcano with a distinctive hydrothermal system.* Nature, 1988. **335**: p. 532-535.
164. Verati, C., et al., *Evidence of bacterial activity from micrometer-scale layer analyses of black-smoker sulfide structures (Pito Seamount Site, Easter microplate).* Chem Geol, 1999. **158**: p. 257-269.
165. Little, C.T., et al., *Late Cretaceous hydrothermal vent communities from the Troodos ophiolite, Cyprus.* Geology, 1999. **27**: p. 1027-1030.
166. Little, C.T., et al., *Early Jurassic hydrothermal vent communities from the Franciscan Complex, San Rafael Mountains, California.* Geology, 1999. **27**(2): p. 167-170.
167. Hoffmann, B.A. and J.D. Farmer, *Filamentous fabrics in low-temperature mineral assemblages: are they fossil biomarkers? Implications for the search for a subsurface fossil record on the early Earth and Mars.* Planetary and Space Science, 2000. **48**: p. 1077-1086.
168. Sobolev, D. and E.E. Roden, *Suboxic deposition of ferric iron by bacteria in opposing gradients of Fe(II) and oxygen at circumneutral pH.* Applied and Environmental Microbiology, 2001. **67**(3): p. 1328-1334.
169. Edwards, K.J., et al., *Isolation and characterization of novel psychrophilic, nuetrophilic, Fe-oxidizing chemolithoautotrophic alpha- and gamma-Proteobacteria from the deep sea.* Appl Environ Microbiol, 2003. **in press**.

170. Colmer, A.R. and M.E. Hinkle, *The role of microorganisms in acid mine drainage.* Science, 1947. **106**: p. 253-256.
171. Temple, K.L. and A.R. Colmer, *The autotrophic oxidation of iron by a new Bacterium: Thiobacillus ferrooxidans.* J. Bac, 1951. **62**: p. 605-611.
172. Kelly, D.P., and A. P. Wood., *Reclassification of some species of Thiobacillus to the newly designated genera Acidithiobacillus gen. nov., Halothiobacillus gen. nov., and Thermithiobacillus gen. nov.* J Syst Bacteriol, 2000. **50**: p. 511-516.
173. Crichton, R.R., *Inorganic Biochemistry of Iron metabolism*. 1991: Chichester:Ellis Horwood, pp263.
174. Crichton, R.R. and J.-L. Pierre, *Old iron, young copper: from Mars to Venus.* BioMetals, 2001. **14**: p. 99-112.
175. Bou-Abdallah, F., et al., *Iron detoxification properties of Escherichia coli bacterioferritin. Attenuation of oxyradical chemistry.* J Biol Chem, 2002. **277**(40): p. 37064-9.
176. Emerson, D., R.E. Garen, and W.C. Ghiorse, *Formation of Metallogenium-like structures by a manganese-oxidizing fungus.* Arch Microbiol, 1989. **151**: p. 223-231.
177. Blake, R.C.I., E.A. Shute, and K.J. White, *Enzymology of respiratory iron oxidation.* Biohydrometallurgy, 1989. **1**(391-401).
178. Fukumori, Y., et al., *Fe (II) oxidizing enzyme purified from Thibacillus ferroxidans.* FEMS Microbiol. Lett, 1988. **50**: p. 169-172.
179. Kusano, T., et al., *Molecular cloning of the gene encoding Thiobacillus ferrooxidans Fe(II) oxidase. High homology of the gene product with HiPIP.* J Biol Chem, 1992. **267**(16): p. 11242-7.
180. Tamegai, H., et al., *Two membrane-bound c-type cytochromes of Thiobacillus ferrooxidans: purification and properties.* FEMS Microbiol Lett, 1994. **119**(1-2): p. 147-53.
181. Cox, J.C. and D.H. Boxer, *The purification and some properties of rustcyanin a blue copper protein involved in iron(II) oxidation from T. ferrooxidans.* Biochem. J, 1978. **174**: p. 497-502.
182. Hall, J.F., S.S. Hasnain, and W.J. Ingledew, *The structural gene for rusticyanin from Thiobacillus ferrooxidans: cloning and sequencing of the rusticyanin gene.* FEMS Microbiol. Lett, 1996. **137**: p. 85-89.
183. Kai, M., et al., *Thiobacillus ferrooxidans cytochrome c oxidase: purification, and molecular and enzymatic features.* J Biochem (Tokyo), 1992. **112**(6): p. 816-21.
184. Iwata, S., et al., *Structure at 2.8 A resolution of cytochrome c oxidase from Paracoccus denitrificans.* Nature, 1995. **376**(6542): p. 660-9.
185. Corstjens, P.L., et al., *Enzymatic iron oxidation by Leptothrix discophora: identification of an iron-oxidizing protein.* Appl Environ Microbiol, 1992. **58**(2): p. 450-4.
186. Beard, B.L., et al., *Iron Isotope Biosignatures.* Science, 1999. **285**(5435): p. 1889-1892.
187. Beard, B.L., et al., *Application of Fe isotopes to tracing the geochemical and biological cycling of Fe.* Chemical Geology. **In Press, Corrected Proof**.
188. Mandernack, K.W., et al., *Oxygen and Iron Isotope Studies of Magnetite Produced by Magnetotactic Bacteria.* Science, 1999. **285**(5435): p. 1892-1896.
189. Brantley, S.L., L. Liermann, and T.D. Bullen, *Fractionation of Fe isotopes by soil microbes and organic acids.* Geology, 2001. **29**(6): p. 535-538.
190. Beard, B.L. and C.M. Johnson, *High precision iron isotope measurements of terrestrial and lunar materials.* Geochimica et Cosmochimica Acta, 1999. **63**(11-12): p. 1653-1660.

191. Janney, D.E., J. M. Cowley, and P. R. Buseck., *Structure of synthetic 2-line ferrihydrite by electron nanodiffraction.* Am Mineral, 2000. **85**: p. 1180-1187.
192. Takai, M., K. Kamimura, and T. Sugio, *A new iron oxidase from a moderately thermophilic iron oxidizing bacterium strain TI-1.* Eur J Biochem, 2001. **268**(6): p. 1653-8.
193. Sambrook, J., Fritsch, E.F and Maniatis, T., *Molecular Cloning: a laboratory manual. Cold Spring Harbour Laboratory Press, New York*. 1989.
194. Batzoglou, S., et al., *ARACHNE: a whole-genome shotgun assembler.* Genome Res, 2002. **12**(1): p. 177-89.
195. Jaffe, D.B., et al., *Whole-genome sequence assembly for Mammalian genomes: arachne 2.* Genome Res, 2003. **13**(1): p. 91-6.
196. Altschul, S.F., et al., *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.
197. Altschul, S.F. and E.V. Koonin, *Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases.* Trends Biochem Sci, 1998. **23**(11): p. 444-7.
198. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res, 1997. **25**(17): p. 3389-402.
199. Karp, P.D., S. Paley, and P. Romero, *The Pathway Tools software.* Bioinformatics, 2002. **18 Suppl 1**: p. S225-32.
200. Appia-Ayme, C., Guiliani, N., Ratouchniak, J., and Bonnefoy,V, *Characterization of an operon Encoding Two c-Type Cytochromes, an aa3-Type Cytochrome Oxidase, and Rusticyanin in Thiobacillus ferroxidans ATCC 33020.* Appl Environ Microbiol, 1999. **65**(11): p. 4781-4787.
201. Richardson, L.L., C. Aguilar, and K.H. Nealson, *Manganese oxidation in pH and 0_2 microenvironments produced by phytoplankton.* Limnol. Oceanogr., 1988. **33**(3): p. 352-363.
202. López-Archilla, A.I. and R. Amils, *A comparative ecological study of two acidic rivers in southwestern Spain.* Microbial Ecology, 1999. **38**: p. 146-156.
203. González-Toril, E., *Ecología molecular de la comunidad microbiana de un ambiente extremo: el Rio Tinto*. 2002, Universidad Autónoma de Madrid: Madrid. p. 136.
204. López-Archilla, A.I., I. Marin, and R. Amils, *Bioleaching and interrelated acidophilic microorganisms from Rio Tinto, Spain.* Geomicrobiology Journal, 1993. **11**: p. 223-233.
205. López-Archilla, A.I., et al., *El rio Tinto, un curso de agua vivo pero con mala fama*, in *Quercas*. 1994. p. 19-22.
206. López-Archilla, A.I., I. Marin, and R. Amils, *Microbial ecology of an acidic river: Biotechnological applications.* Biohydrometallurgical Processing, 1995: p. 63-73.
207. Amils, R., A.I. Lopez-Archilla, and I. Marin, *Modelos de vida en condiciones limite como base para la exobiologia.* Revista de la Real Academia de Cienceas Exactas, Fisicos y Naturales, 1997. **91**(2): p. 87-99.
208. Gónzalez-Toril, E., et al., *Geomicrobiology of the Tinto River, a model of interest for biohydrometallurgy*, in *Biohydrometallurgy: fundamentals, technology, and sustainable development*, V.S.T. Ciminelli and O. García, Editors. 2001, Elsevier: Amsterdam. p. 639-650.
209. Palacios, C., D. Kysela, and M. Sogin. *Serial Analysis of rRNAs (SARS): A novel molecular strategy for sampling microbial diversity throughout the universe*. in *Proceedings of the Second European Workshop on Exo/Astrobiology*. 2002. Graz, Austria.

210. Velculescu, V.E., et al., *Serial analysis of gene expression.* Science, 1995. **270**(5235): p. 484-7.
211. Ewing, B., et al., *Base-calling of automated sequencer traces using phred. I. Accuracy assessment.* Genome Res, 1998. **8**(3): p. 175-85.
212. Ewing, B. and P. Green, *Base-calling of automated sequencer traces using phred. II. Error probabilities.* Genome Res, 1998. **8**(3): p. 186-94.
213. Pearson, W.R., *Rapid and sensitive sequence comparison with FASTP and FASTA.* Methods Enzymol, 1990. **183**: p. 63-98.
214. Pearson, W.R., *Flexible sequence similarity searching with the FASTA3 program package.* Methods Mol Biol, 2000. **132**: p. 185-219.
215. Hughes, J.B., et al., *Counting the uncountable: statistical approaches to estimating microbial diversity.* Appl Environ Microbiol, 2001. **67**(10): p. 4399-406.
216. López-Archilla, A.I., I. Marin, and R. Amils, *Microbial community composition and ecology of an acidic aquatic environment: the Tinto River, Spain.* Microbial Ecology, 2001. **41**: p. 20-35.
217. Avaniss-Aghajani, E., et al., *Molecular technique for rapid identification of mycobacteria.* J Clin Microbiol, 1996. **34**(1): p. 98-102.
218. Boulter, C., *Comparison of Rio Tinto, Spain and Guaymas Basin, Gulf of California: An explanation of a supergiant massive sulfide deposit in an ancient sill-sediment complex.* Geology, 1993. **21**: p. 801-804.
219. Fernandez-Remolar, D., et al. *The Tinto River basin: An analog for Meridiani hematite formation on Mars?* in *LPSC 33, Lunar Planet. Inst.* 2002. Houston.
220. Einsele, G., et al., *Intrusion of basaltic sills into highly porous sediments, and resulting hydrothermal activity.* Nature, 1980. **283**: p. 441-445.
221. Head, J. and L. Wilson, *Mars: A review and synthesis of general environments and geological settings of magma-H2O interactions*, in *Volcano-Ice Interactions on Earth and Mars*, J. Smellie and M. Chapman, Editors. 2002, Geol. Soc. London. p. 27-57.
222. Gieskes, J., et al., *Hydrothermal activity in the Guaymas Basin, Gulf of California: a synthesis*, in *Initial Reports of the Deep Sea Drilling Project*, J. Curray and D. Moore, Editors. 1982, US Govt. Printing Office. p. 1159-1168.
223. Kokelaar, B., *Fluidization of wet sediments during the emplacement and cooling of various igneous bodies.* Geol. Soc. London Journ., 1982. **139**: p. 21-33.
224. Zimanowski, B. and R. Buttner, *Dynamic mingling of magma and liquified sediments.* J. Volc. Geotherm. Res., 2002. **114**: p. 37-44.
225. Wohletz, K., *Water/magma interaction: Some theory and experiments on peperite formation.* J. Volc. Geotherm. Res., 2002. **114**: p. 19-35.
226. Skilling, I., J. White, and J. McPhie, *Peperite: A review of magma-sediment mingling.* J. Volc. Geotherm. Res., 2002. **114**: p. 1-17.
227. Head, J. and L. Wilson, *Deep submarine pyroclastic eruptions: Theory and predicted landforms and deposits.* J. Volc. Geothermal Res., 2003. **121**: p. 155-193.
228. Wilson, L. and J. Head, *Mars: Review and analysis of volcanic eruption theory and relationships to observed landforms.* Rev. Geophysics, 1994. **32**: p. 221-263.
229. Wilson, L. and J. Head, *Heat transfer and melting in subglacial basaltic volcanic eruptions: Implications for volcanic deposit morphology and meltwater volumes*, in *Volcano-Ice Interactions on Earth and Mars*, J. Smellie and M. Chapman, Editors. 2002, Geol. Soc., London. p. 5-26.

230. Lister, J., *Buoyancy-driven fluid fracture: Similarity solutions for the horizontal and vertical propagation of fluid-filled cracks.* J. Fluid Mechanics, 1990. **217**: p. 213-229.
231. Rubin, A., *Dikes vs. diapirs in viscoelastic rock.* Earth Planet. Sci. Lett., 1993. **119**: p. 641-659.
232. Wilson, L. and J. Head, *Tharsis-radial graben systems as the surface manifestation of plume-related dike intrusion complexes.* J. Geophys. Res. Planets, 2002. **107**: p. 10.
233. Christensen, P., et al., *Global mapping of Martian hematite mineral deposits: Remnants of water-driven processes on early Mars.* J. Geoph. Res., 2001. **106**: p. 23873-23885.
234. Burns, R. *Gossans on Mars*. in *18th Lunar Planet. Sci. Conf.* 1988. Cambridge.
235. Hudson-Edwards, K., C. Schell, and M. Macklin, *Mineralogy and geochemistry of alluvium contaminated by metal mining in the Rio Tinto area, southwest Spain.* Applied Geochemistry, 1999. **14**: p. 1015-1030.
236. Burns, R., *Mineralogical Applications of Crystal Field Theory*. 1993, New York: Cambridge Univ. Press. 551.
237. Morris, R., et al., *Mineralogy, composition, and alteration of Mars Pathfinder rocks and soils: Evidence from multispectral, elemental, and magnetic data on terrestrial analogue, SNC meteorite, and Pathfinder samples.* J. Geophys. Res., 2000. **105**: p. 1757-1817.
238. Cooper, J.F., M.E. Weary, and F.T. Jordan, *The impact of non-endotoxin LAL-reactive materials on Limulus amebocyte lysate analyses.* PDA Journal of Pharmaceutical Science and Technology, 1997. **51**: p. 1-7.
239. Wainwright, N.R. and A. Child. *Biomarkers in carbonate thermal spring deposits: Implications for Mars.* in *Lunar and Planetary Science Conference XXX*. 1999.
240. Ellwood, M.S., et al., *Strand displacement applied to assays with nucleic acid probes.* Clinical Chemistry, 1986. **32**(9): p. 1631-1636.
241. Morrison, L.E., T.C. Halder, and L.M. Stols, *Solution-phase detection of polynucleotides using interacting fluorescent labels and competitive hybridization.* Analytical Biochemistry, 1989. **183**: p. 231-244.
242. Terouanne, B., et al., *Quantitative and qualitative analysis of amplified DNA sequences by a competitive hybridization assay.* Analytical Biochemistry, 1992. **205**: p. 193--199.
243. Li, Q., et al., *A new class of homogeneous nucleic acid probes based on specific displacement hybridization.* Nucleic Acids Research, 2002. **30**(2).
244. Soderlund, J., *DNA hybridization: comparison of liquid and solid phase formats.* Annals of Biological Chemistry, 1990. **48**(7): p. 489-491.
245. Lin, P.K.T. and D.M. Brown, *Synthesis and duplex stability of oligonucleotides containing cytosine-thymine analogues.* Nucleic Acids Research, 1989. **17**(24): p. 10373-10383.
246. Brown, D.M. and P.K.T. Lin, *Synthesis and duplex stability of oligonucleotides containing adenine-guanine analogues.* Charbohydrate Research, 1991. **216**: p. 129-139.
247. Lin, P.K.T. and D.M. Brown, *Oligonucleotides containing degenerate bases: Synthesis and uses.* Methods in Molecular Biology, 1994. **26**: p. 187-206.
248. Lovell, C.R., et al., *Recovery and phylogenetic analysis of nifH sequences from diazotrphic bacteria associated with dead aboveground biomass of Spartina alterniflora.* Applied and Environmental Microbiology, 2001. **67**(11): p. 5308-5314.
249. Lovell, C.R., et al., *Molecular analysis of diazotroph diversity in the rhizosphere of the smooth cordgrass, Spartina alterniflora.* Applied and Environmental Microbiology, 2000. **66**(9): p. 3814--3822.

250. Margulis, L., *Symbiotic theory of the origin of eukaryotic organelles; criteria for proof.* Symp Soc Exp Biol, 1975(29): p. 21-38.
251. Buchner, P., *Endosymbiosis of animals with plant microorganisms*. 1965, New York: interscience Publishers, Inc.
252. Dasch, G.A., *Morphological and molecular studies on intracellular bacterial symbiotes of insects*. 1975, Yale University: New Haven, CT. p. 329.
253. Schroder, D., et al., *Intracellular endosymbiotic bacteria of Camponotus species (carpenter ants): systematics, evolution and ultrastructural characterization.* Mol Microbiol, 1996. **21**(3): p. 479-89.
254. Wernegreen, J., et al., *Genome evolution in an insect cell: Distinct features of an ant-bacterial partnership.* Biol. Bull., 2003. **in press (April 2003 issue)**.
255. Werren, J.H. and J. Jaenike, *Wolbachia and cytoplasmic incompatibility in mycophagous Drosophila and their relatives.* Heredity, 1995. **75 (Pt 3)**: p. 320-6.
256. Gil, R., et al., *Extreme genome reduction in Buchnera spp.: toward the minimal genome needed for symbiotic life.* Proc Natl Acad Sci U S A, 2002. **99**(7): p. 4454-8.
257. Moran, N.A., *Microbial minimalism: genome reduction in bacterial pathogens.* Cell, 2002. **108**(5): p. 583-6.
258. Shigenobu, S., et al., *Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS.* Nature, 2000. **407**(6800): p. 81-6.
259. Tamas, I., et al., *50 million years of genomic stasis in endosymbiotic bacteria.* Science, 2002. **296**(5577): p. 2376-9.
260. Akman, L., et al., *Genome sequence of the endocellular obligate symbiont of tsetse flies, Wigglesworthia glossinidia.* Nature Genetics, 2002. **32**(3): p. 402-7.
261. Slatko, B.E., et al., *The Wolbachia Genome Consortium.* Microb Comp Genomics, 1999. **4**(3): p. 161-5.
262. O'Neill, S.L., *Wolbachia: why these bacteria are important to genome research.* Microb Comp Genomics, 1999. **4**(3): p. 159.
263. Sun, L.V., et al., *Determination of Wolbachia genome size by pulsed-field gel electrophoresis.* J Bacteriol, 2001. **183**(7): p. 2219-25.
264. Wernegreen, J.J., A.B. Lazarus, and P.H. Degnan, *Small genome of Candidatus Blochmannia, the bacterial endosymbiont of Camponotus, implies irreversible specialization to an intracellular lifestyle.* Microbiology, 2002. **148**: p. 2551-2556.
265. Tatusov, R.L., et al., *The COG database: new developments in phylogenetic classification of proteins from complete genomes.* Nucleic Acids Res, 2001. **29**(1): p. 22-8.
266. Mira, A., H. Ochman, and N.A. Moran, *Deletional bias and the evolution of bacterial genomes.* Trends Genet, 2001. **17**(10): p. 589-96.
267. von Dohlen, C.D., et al., *Mealybug beta-proteobacterial endosymbionts contain gamma-proteobacterial symbionts.* Nature, 2001. **412**(6845): p. 433-6.
268. Charles, H., et al., *Genome size characterization of the endocellular symbiotic bacteria of the weevil Sitophius oryzae, using pulse field gel electrophoresis.* Insect Biochem Mol Biol, 1997. **27**: p. 345-350.
269. Wernegreen, J.J., et al., *Decoupling of genome size and sequence divergence in a symbiotic bacterium.* J Bacteriol, 2000. **182**(13): p. 3867-9.
270. Buckley, D.H. and T.M. Schmidt, *The Structure of Microbial Communities in Soil and the Lasting Impact of Cultivation.* Microb Ecol, 2001. **42**(1): p. 11-21.

271. Bergthorsson, U. and H. Ochman, *Distribution of chromosome length variation in natural isolates of Escherichia coli.* Mol Biol Evol, 1998. **15**(1): p. 6-16.
272. Lawrence, J. and J. Roth, *Genomic flux: Genome evolution by gene loss and acquisition*, in *Organization of the prokaryotic genome (*R. Charlesbois, Editor. 1999, ASM Press: Washington, D. C.
273. Werren, J.H. and J.D. Bartos, *Recombination in Wolbachia.* Curr Biol, 2001. **11**(6): p. 431-5.
274. Hentschel, U., M. Steinert, and J. Hacker, *Common molecular mechanisms of symbiosis and pathogenesis.* Trends Microbiol, 2000. **8**(5): p. 226-31.
275. Hölldobler, B. and E.O. Wilson, *The Ants*. 1990, Cambridge, Mass: Belknap Press of Harvard University Press.
276. Nesbo, C.L., K.E. Nelson, and W.F. Doolittle, *Suppressive subtractive hybridization detects extensive genomic diversity in Thermotoga maritima.* J Bacteriol, 2002. **184**(16): p. 4475-88.
277. Miao, E.A. and S.I. Miller, *Bacteriophages in the evolution of pathogen-host interactions.* Proc Natl Acad Sci U S A, 1999. **96**(17): p. 9452-4.
278. Kim, C.C., et al., *Improved analytical methods for microarray-based genome-composition analysis.* Genome Biol, 2002. **3**(11): p. RESEARCH0065.
279. Ochman, H. and N.A. Moran, *Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis.* Science, 2001. **292**(5519): p. 1096-9.
280. Kondo, N., et al., *Genome fragment of Wolbachia endosymbiont transferred to X chromosome of host insect.* Proc Natl Acad Sci U S A, 2002. **99**(22): p. 14280-5.
281. Masui, S., et al., *Bacteriophage WO and virus-like particles in Wolbachia, an endosymbiont of arthropods.* Biochem Biophys Res Commun, 2001. **283**(5): p. 1099-104.
282. Michl, D., et al., *Phylogenetic transfer of organelle genes to the nucleus can lead to new mechanisms of protein integration into membranes.* Plant J, 1999. **17**(1): p. 31-40.
283. Lawrence, J.G. and H. Ochman, *Molecular archaeology of the Escherichia coli genome.* Proc Natl Acad Sci U S A, 1998. **95**(16): p. 9413-7.
284. Degnan, P., et al., *Cospeciation of Blochmannia and their Camponotus hosts.* Systematic Biology. **in prep**.
285. Adams, K.L., et al., *Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants.* Nature, 2000. **408**(6810): p. 354-7.
286. Ishikawa, H., *Biochemical and molecular aspects of endosymbiosis in insects.* Int Rev Cytol, 1989. **116**: p. 1-45.
287. Baumann, P., et al., *Genetics, physiology, and evolutionary relationships of the genus Buchnera: intracellular symbionts of aphids.* Annu Rev Microbiol, 1995. **49**: p. 55-94.
288. Evans, J.D. and D.E. Wheeler, *Differential gene expression between developing queens and workers in the honey bee, Apis mellifera.* Proc Natl Acad Sci U S A, 1999. **96**(10): p. 5575-80.
289. Sauer, C., et al., *Tissue Localization of the Endosymbiotic Bacterium "Candidatus Blochmannia floridanus" in Adults and Larvae of the Carpenter Ant Camponotus floridanus.* Appl Environ Microbiol, 2002. **68**(9): p. 4187-93.

7. STRENGTHENING OF THE ASTROBIOLOGY COMMUNITY

The efforts of our interdisciplinary group of research scientists, EPO, and IT specialists, contribute to the overall enhancement of the science and education initiatives of NAI. By maintaining a vigorous education program at the K-14, advanced undergraduate and graduate level, and providing unique content-rich web resources for education and public outreach, we bring understanding to those outside our research program, and train the next generation of scientists. Our facilities for high-throughput DNA sequencing and spectral analysis offer support to research initiatives across NAI. Many of our PI and CoIs have formed collaborations within the broader scope of the NAI and they bring an astrobiology perspective to other endeavors of the Space Science Community. Through the implementation of innovative IT solutions for collaboration, we deepen the relationships among members of our team and foster communication with other members of the virtual institute. Elements of our program are:

7.1 EDUCATION AND PUBLIC OUTREACH (L. Olendzenski, D. Patterson, M. L. Sogin).

Our research themes **EARTH'S EARLY BIOSPHERE; TERRESTRIAL ANALOGUES FOR EARLY MARS; BIOSIGNATURES AND LIFE DETECTION** and **EVOLUTION OF GENOME ARCHITECTURE IN PROKARYOTES** address the goals and objectives of NASA's Office of Space Science (OSS) to probe the origin and evolution of life on Earth and to determine if life exists elsewhere in our solar system. Our EPO program and plan for strengthening the Astrobiology community extend the OSS education goals (to enhance the quality of science, math and technology education at the pre-college level) by integrating our scientific expertise in microbiology, molecular evolution, and planetary geology into intensive workshops for both teachers (Gr. 6-14) and research scientists. The NAI team at MBL offers the workshops *Living in the Microbial World, Life and Living in Space*, and *Astrobiology Mini-Workshops* for teachers. Two MBL courses, the *Workshop in Molecular Evolution* and *Advances in Genome Technology and Bioinformatics,* provide advanced training for graduate students, Postdoctoral fellows and Principle Investigators. We will continue to provide web-based resources for educators and we will expand our web-site **micro*scope** to provide integrated access to data and results from our research projects.

Teacher Enhancement at the MBL. During our initial membership in the Astrobiology Institute, we developed three opportunities for teacher enhancement at the MBL. Middle, high school and community college teachers receive first-hand exposure to current astrobiology topics presented through lecture and discussions with active research scientists and spend significant time in the laboratory gaining hands-on experience with essential lab and teaching techniques, available classroom curricula and inquiry learning. Our program fulfills the following recommendation of the National Science Education Standards for Professional Development for Teachers of Science that:

> "…prospective and practicing teachers must take science courses in which they learn science through inquiry, having the same opportunities as their students will have to develop understanding", and that providers of professional development "must design courses that are heavily based on investigations, where current and future teachers have direct contact with phenomena…. and are involved in groups working on real, open-ended questions."

By partnering with expert teacher Barbara Dorritie of the Cambridge, MA Public School system (see letter of support), we incorporated appropriate models of inquiry into our workshop (*Living in the Microbial World*). We will adopt similar strategies for all of our workshops. Time is devoted to formulation of testable questions, pursuit of individual and group investigations, and to a thorough understanding of the scientific process. We also provide methods for classroom implementation and for the sharing of teacher's experiences while doing inquiry.

Our workshop series includes "*Living in the Microbial World*", a one-week intensive summer workshop for middle and high school students, focusing on microbial diversity and evolution. "*Life and Living in Space*", a four-day workshop for high school and community college teachers, convenes in the fall and is offered in conjunction with the Center for Advanced Studies of Space Life Sciences (CASSLS) at the MBL. This workshop combines NASA Life Science topics and areas of Astrobiology content including planetary protection and life detection. Astrobiology Mini-Workshops, offered during the spring, focus on a single topic and set of activities that change each year. To date these one and a half day workshops have included: *Mapping Time and Space with Forams*, *Life on Earth and Elsewhere*, and *Inquiry into Extremophiles*. *Mapping Time and Space with Forams* focuses on a series of activities from the Living Sands curriculum (www.neosci.com). "*Life on Earth and Elsewhere*" makes use of the activities in the Astrobiology Institute Educators Resource Guide, supplemented by lectures on life in extreme environments and evidence for water on other planets. In "*Inquiry into Extremophiles*", we use a format that encourages teachers to directly observe results from experiments on thermophiles and hyperthermophiles, halophiles, acidophilic eukaryotes and UV-resistant microorganisms. The participants then design and carry out their own investigations.

As part of our continuing EPO program we will improve the curricula through enhanced participation of science team members from MBL, WHOI and Brown. Our Astrobiology Mini-Workshop offerings will include a new workshop that capitalizes on the geologic expertise offered by the Brown node; one or more workshops will employ the SETI Institute's "Voyages Through Time" curriculum pieces; and a third workshop will focus on the **micro*scope** web-based resource in microbial diversity and astrobiology (see below).

Dissemination and expected impact. *Living in the Microbial World* and *Life and Living in Space* collectively accommodate 40 teachers yearly. Participants in either of these workshops have the opportunity to attend *Astrobiology Mini-Workshops* in the spring. We preferentially recruit educator teams that are made up of at least one classroom science teacher and another science teacher, administrator, science coordinator, support staff or math teacher from the same school or school district. We ask teachers to share content and activities with their colleagues through a short workshop or meeting upon return to their school. These mechanisms help to insure that workshop elements will be propagated and incorporated into existing classroom curricula by other teachers in the same school or district. We recruit our teachers from Cape Cod and surrounding Massachusetts areas, through presentations at national meetings (e.g. NABT, NSTA) and through Astrobiology Institute contacts. The MBL Education Office participates annually in the Society for Advancement of Chicano and Native Americans in Science (SACNAS) national conference to promote our education programs. Working in partnership with Marge Anderson, the SMART! Education Program Coordinator of Pfizer Inc., we directly recruit teachers from school districts associated with Pfizer Research Centers and support their travel to our workshops in Woods Hole. In partnership with Pfizer, we will target minority teachers and educators from underserved communities including New York City and New Jersey, Puerto Rico, San Diego/La Jolla, and the Midwest (see letter of support). We have established contacts with Cambridge and Boston, Massachusetts's school district science coordinators to attract teachers from these urban centers. To engage Hispanic minority teachers and students, we will incorporate spanish language astrobiology resources into our programs as they become available. Two or our CoIs (Morrison and Amaral-Zettler) are Spanish speaking Hispanic minorities and postdoctoral researcher Carmen Palacios is a citizen of Spain.

Assessment. We use mid-course feedback and comprehensive end-of-course questionnaires to assess the content, activities and pedagogy of our workshops. Through these mechanisms we can make immediate improvements in workshop structure and content. Teachers who return to subsequent workshops report orally on their success in incorporating astrobiology and microbial diversity related activities and content into their curricula. As part of our ongoing efforts to

improve the effectiveness of our EPO program, we will comprehensively survey all teachers who have participated in our programs to determine what content has been incorporated into their curricula and which activities and modifications they have found most successful. The survey will include written questionnaires, follow-up oral interviews, and continuing interaction on web-based bulletin boards served from MBL. We will publish resultant summaries and assessments on the education pages of the web site **micro*scope** (see below) to assist others who seek to implement astrobiology content in the classroom.

Traveling Collaborative Workshop Team. The Collaborative Workshop Team integrates EPO leads, educators and researchers from five Astrobiology Lead Teams. The proposed team of Catherine Tsairides, (EPO lead, NASA Ames Research Center), Lorraine Olendzenski, (EPO lead, MBL), Lisa Brown and Angela Phelps (EPO leads, Pennsylvania State University), Jackie Allen and Kay Tobola (EPO leads, NASA Johnson Space Center), has developed effective professional development centered around existing astrobiology educational products. Team members have expertise as researchers, principals, university instructors, publishers and classroom educators, and they offer a mix for professional development that includes utilizing members of the research team as speakers, panel members, and hands-on presenters. Education specialists provide expertise in understanding the needs of teachers, classroom management, pedagogy and modeling of activities. The Team consistently models science as inquiry and connects educators to the world of scientific research through the presentation of various programs that directly involve astrobiologists in professional improvement nationwide. We will provide hands-on science experiences and activities through presentations, workshops and short courses, at venues including, but not limited to NSTA, NABT, Astrobiology meetings, and meetings organized by the American Society for Microbiology, Geological Society of America, and other relevant professional societies. These workshops will provide another distribution mechanism for upcoming astrobiology products while continuing to promote the Life on Earth and Elsewhere Educator Resource Guide, the **micro*scope** web based resource in microbial diversity, Johnson Space Center lessons and materials associated with the Yellowstone National Park Astrobiology Initiative.

7.2. WEB RESOURCES FOR ASTROBIOLOGY EDUCATION AND PUBLIC OUTREACH (L. OLENDZENSKI, D. PATTERSON, M. L. SOGIN, B. OLSSON).

The **micro*scope** web site (http://www.mbl.edu/microscope) is an innovative resource for education and public outreach. **Micro*scope** is an image-rich resource providing descriptions and pictures of all categories of microorganisms, currently containing over 4500 downloadable high-resolution images. The images and accompanying information can be accessed by a number of methods, including by habitat, cell shape, alphabetically by genus name and by hierarchical taxonomic classification. Surveyed habitats include marine environments, saltwater marshes, fresh water ponds, and others including the extreme environments, Yellowstone National Park, Río Tinto, and Lake Toolik in Arctic Alaska.

Currently, we are developing **micro*scope** into a central educational repository for students and teachers interested in astrobiology and microbial diversity. Educational resources under development for the site include an Astrobiology Gallery - a page that highlights organisms of interest to Astrobiology researchers, and a collection of hands on-activities related to microbes accessible by grade level and content area. The site provides information on how each activity ties into the National Science Content Standards, links to other useful educational sites in microbiology and astrobiology, glossary elements, and instructions on using **micro*scope**'s Lucid and X-ID matrix taxonomic identification guides. Over the course of this program, we will continuously add to these elements, and include new activities developed through our teacher workshops, presentations and collaborative teaching efforts with the MIT/Harvard team. We will develop resources in both English and Spanish about the microbial diversity of Río Tinto that will be contained in or linked to **micro*scope**, providing a valuable tool for minority education.

7.3 EPO PROGRAM MANAGEMENT. Lorraine Olendzenski is a research scientist in microbiology and molecular evolution and participates as a member of our science team and will serve as our EPO lead. She will report directly to the PI and will oversee and carry out all aspects of our on-site teacher enhancement programs, including implementation of the three annual workshops, and acquisition of additional operating funds. She will be responsible for coordinating and implementing the comprehensive program assessment during year three of the Woods Hole Astrobiology team. She will participate in the traveling Collaborative Workshop Team. She will continually liaise between Astrobiology Co-Is, EPO leads from other sites and NAI Central to facilitate the provision of content and other expertise in the review and development of curriculum materials as necessary. Co-I David Patterson, a senior research professor with expertise in protist taxonomy, diversity, and ecology, will oversee all aspects of the development of the **micro*scope** database structure, including expansion of its content areas and the development of generic software tools which expands its capabilities (described in more detail in a separate section, below). As part of her duties, Lorraine Olendzenski will expand the educational portions of micro*scope including activities and supporting materials.

7.4 TRAINING THE NEXT GENERATION OF ASTROBIOLOGY RESEARCHERS: COURSES FOR GRADUATES STUDENTS, POST DOCS AND RESEARCHERS. (M. SOGIN, L. OLENDZENSKI, H. MORRISON, A. TESKE, K. EDWARDS, S. SIEVERT, J. MUSTARD, J. HEAD).

Increasingly, comparative molecular and genomic approaches are being used to address questions of phylogeny, ecology and early evolution related to Astrobiology research goals. The MBL offers intensive laboratory-training courses designed for advanced researchers and graduate students who want expert instruction in molecular biology and bioinformatics techniques. These courses combine presentations from distinguished faculty on current research topics in molecular phylogenetic analyses and genomics with hands-on instruction using the newest equipment, algorithms, software and techniques. Our Astrobiology program participates in two of these:

Advances in Genome Technology and Bioinformatics is a comprehensive, four-week course in genome science that integrates bioinformatics with the latest laboratory techniques for genome sequencing, genome analysis, and high throughput gene expression (DNA microarrays). A distinguished faculty from major universities, bioinformatics centers, The Institute for Genomic Research (TIGR) and the MBL provide instruction that integrates lectures with laboratory exercises both at the computer and in a high technology, high throughput facility. Astrobiology PI Mitchell Sogin is one of the course co-directors, and Co-Is from the MBL Astrobiology team participate as faculty. The major laboratory modules include 1) Genome Sequencing (vector development, library construction, high throughput sequencing technologies, principles of automation using advanced robotic liquid handlers, genome assembly algorithms and closure strategies); 2) Bioinformatics (Gene prediction algorithms, annotation, database construction and searching, phylogenetics and molecular evolution); and 3) Functional Genomics (DNA microarrays, data analysis). Symposia focusing on Environmental and Evolutionary Genomics, Eukaryotic Microbial Genome Projects, Organelle Evolution and other topics are also part of the program. The course takes place during the month of October and is open to twenty-four students.

The Workshop on Molecular Evolution, in its sixteenth year, is recognized as the finest course of its type in the world. The Workshop includes a series of lectures, demonstrations and computer labs that span the field of molecular evolution. A distinguishing feature of the course is a state-of-the-art computer laboratory furnished with Linux and Unix servers and workstations for comparative analysis of molecular data. Software authors and experts in the use of packages such as Clustal W and Clustal X, COMPARE, FASTA, GCG, LAMARC, PAML, PAUP*, and PHYLIP provide demonstration and consultation. This two-week program is open to 60 participants with strong interests in molecular evolution, systematics, and population genetics. A one week extended topics session allows 15 selected students to spend time analyzing their

research data sets. Topics covered in the course include: databases and sequence matching, phylogenetic analysis including Bayesian analysis and maximum likelihood theory, molecular evolution at organismal and higher levels, molecular evolution and development, gene duplication and divergence, gene family organization, evolution of large multigene families, molecular evolution in bioinformatics and comparative genomics.

University Education. Through new courses offered on the graduate and undergraduate level at University of North Carolina, Brown University, and WHOI, our Co-Is are raising the profile of Astrobiology in related disciplines. A new joint MBL/Brown graduate program will allow implementation of additional courses in September, 2004. Ongoing and proposed University courses include:

Extremophilic Microbiology. A. Teske (UNC Chapel Hill) teaches a comprehensive marine microbiology course every fall semester that emphasizes extreme microbial physiologies and ecosystems (hydrothermal vents, the deep marine subsurface) as model systems for astrobiology. He also advises two Ph.D. students and two PostDoctoral students who are working on key gene analyses of deep subsurface sediments from a wide range of geochemical settings, as outlined in our project description on evolution and environmental extremes of sulfur and methane cycles. In this way, we are beginning to train a new generation of Ph.D. students and PostDoctoral students who are growing up with astrobiology as an integral part of their research agenda.

Planetary Geoscience. Graduate students in the Brown University Department of Geological Sciences are trained in planetary geosciences. Students that will be supported under this effort will be further trained in astrobiological concepts related to planetary geosciences, with their Ph.D. research focusing on investigations of theRío Tinto region. They will also participate in the analysis of Mars mission data through efforts lead by Co-Is Mustard and Head and thus be able to relate the astrobiological research directly to observations of Mars.

Geomicrobiology. WHOI offers a Joint Graduate Program with MIT that confers Master's Degrees and Doctorates in the areas of Biological, Physical, Chemical and Applied Oceanography, as well as Marine Geology and Geophysics. WHOI Co-Is Stephan Seivert and Katrina Edwards will offer a course "Geomicrobiology on Earth and in the Universe" to Joint Program students. This course will focus on roles of microorganisms in the formation of minerals and signatures, impact of microbes on Earth's geological record, biogeochemical cycling, geology of habitable planets, and potential mineralogical biosignatures for detection of life beyond Earth.

Graduate Courses and Training and at the MBL. Starting in the fall of 2004, the MBL will initiate a Joint Graduate Program with Brown University. Members of our science team currently include researchers from the Brown University Department of Planetary Geology. As part of the proposed graduate program, we will develop upper level survey courses for graduate students focusing on early evolution, planetary geology processes, microbial ecology and evolution and astrobiology. These will be seminar format courses that take place at Brown University campus or the MBL, organized by the Astrobiology PI and CoIs from both Brown University and MBL. The MBL/Brown Graduate Program will accept its first students in September, 2004 and will support a steady-state population of ~40 graduate students whose interests will span basic molecular and cell biology to astrobiology and environmental sciences. Our proposed team will be in a position to explore the creation of an Astrobiology Certificate Program through the University.

Minority Undergraduate Fellowship Program at the Woods Hole Oceanographic Institution. The Woods Hole Oceanographic Institution offers special educational opportunities in oceanography for minority undergraduates enrolled in U.S. colleges or universities. This Minority Fellowship program provides students from minority groups with a first-hand introduction to scientific research. Fellowships are awarded for a ten- to twelve-week period in the summer or for a semester during the academic year, to allow students to pursue an independent research project under the guidance of a member of the WHOI research staff. Each Fellow works with their lab

advisor to design a project that can yield meaningful results in the time allotted, prepares a research paper describing their project and gives an oral presentation of their results. CoIs Stephan Sievert and Katrina Edwards will each accept one qualified fellow per year through this program Fellows will interact with our entire MBL Astrobiology group through the attendance of lab meetings, journal clubs and seminars.

7.5 DEVELOPMENT OF THE MICRO*SCOPE WEB-BASED RESOURCE IN MICROBIAL DIVERSITY AND ASTROBIOLOGY INTO A RESEARCH AND TEACHING TOOL FOR THE ASTROBIOLOGY COMMUNITY (David Patterson).

Tools for gathering information about microbes are few in comparison to those available for larger organisms, yet there are many instances in which this information is crucial for advancing research and improving classroom teaching. The **micro*scope** website fills the gap in available resources and provides students, teachers and researchers with easy access to high quality digital micrographs and information about diverse microbes, both prokaryotic and eukaryotic. A major innovation of the site is the use of uBio software, developed in conjunction with the MBL-WHOI Library, which uses a registry of names as part of the navigational structure and as a filing system upon which to attach information about organisms or other entities. The software allows the recognition of synonyms of names in remote databases and resources. **Micro*scope** is designed to bring together information distributed at other authoritative sites. It uses software to track the taxonomic location of the user and to initiate searches into remotely located databases or other sites generally available on the web, based on the genus being examined. The list of databases and remote sites to be searched can be customized to return results from only certain selected databases.

As an internet-based resource capable of linking distributed knowledge, **micro*scope** has many attributes appropriate to its growth as a knowledge network. Although the content contained in micro*scope focuses on microbes, the underlying organizational structure of micro*scope is flexible and can be adapted to the needs of researchers from a variety of disciplines. The hierarchical classification structure will be expanded so it can be used as an organizing and navigational structure for non-biological data sets. As this system is based on the TNS (taxonomic name server) of uBio, this retains the capacity to reconcile alternative names for the same entity, thereby linking together data about the same thing even if logged under different names. The data model allows for data to be bundled within hierarchical clusters to increase the flexibility of access to data. Linkouts allow access to remotely held data. The result is not a database even though it incorporates databases. The result is a knowledge network capable of integrating data held locally and remotely. It is powerful and flexible, simple to populate, and with relatively little effort can be molded to provide its power to other suites of data.

Our goal is to maintain **micro*scope** as a resource to assist those interested in gathering information about microbial taxa and in understanding better the diversity, complexity and evolution of the microbial world. Through our involvement in the Institute, we will expand its use for archiving image rich data from habitats currently being studied by astrobiologists, including Río Tinto, Guerrero Negro, Baja California, Mexico, and Yellowstone. We will also make the generic structure of **micro*scope** available to interested groups for the creation of similar data management resources to accommodate the diverse different types of data obtained by researchers across the Astrobiology Institute.

7.6 RESEARCH FACILITIES FOR THE GENERAL ASTROBIOLOGY COMMUNITY.

High Throughput DNA Sequencing Facilities at the MBL: DNA sequence data allows astrobiology investigators to survey microbial diversity in the environment, make inferences about the evolution of proteins and physiologies, and reconstruct phylogenies of organisms. Vast quantities of data, including extensive environmental surveys, the sequencing of whole microbial genomes and profiling of protein expression patterns can be obtained quickly using high-throughput sequencing and DNA microarray technology. The Josephine Bay Paul Center at the

MBL maintains and operates an 1800 square-foot laboratory dedicated to state of the art automated DNA sequencing, DNA microarraying, and high-throughput robotics (see **Facilities and Equipment** for details).

Our DNA sequencing capability represents an investment of approximately $1.5 Million in equipment, largely to support our astrobiology research efforts. We designed this facility to include all of the advanced technology currently available within DNA sequencing factories, but at a smaller scale. The operating philosophy is to provide biologists access to the powerful tools of genome science in a cost efficient manner. This capability has radically modified the scale of projects pursued by the Woods Hole astrobiology team and we have carried out DNA based projects for astrobiologists and exobiologists at the University of Washington in Seattle, the University of Colorado, the University of Rhode Island, and the University of Connecticut. We provide access to this facility for NASA supported research studies at our institutional costs for reagents and technician time. The capital equipment acquisition and maintenance costs are not charged back to the users. Our current costs for DNA sequencing are remarkably low at a direct cost of less than $1/reaction. Should demand exceed our capacity, we will seek funds to lease additional equipment for use in astrobiology research projects that would benefit from acquisition of extensive environmental, genomic or other DNA sequence data. The Bay Paul Center also hosts and trains visiting researchers interested in using the equipment for their own projects.

A high-throughput data pipeline connects the sequencing facilities with a well-equipped facility for bioinformatics and computational biology. The core of the center consists of a 104-processor Beowulf cluster, a farm of large-memory 64-bit Alpha servers and several terabytes of fast-access data storage (see **Facilities and Equipment** for details). We maintain the latest versions of both commercial and open source software packages for analysis of genetic sequences and inference of phylogenetic trees. The computer facilities can be accessed by other NAI Teams over fast Internet2 connections, and for visiting researchers we also offer personal workstations and full-time staff support.

The Reflectance Experiment Laboratory (RELAB) at Brown University. The composition of inaccessible planetary surfaces can be derived from remotely obtained reflectance spectra interpreted with the help of spectroscopic laboratory data. The RELAB is supported by NASA as a multi-user spectroscopy facility, and laboratory time is available at no charge to investigators who are in NASA-funded programs. This facility is designed to acquire high precision and accuracy reflectance spectra of materials from 0.3 to 25 μm under variable incidence and emergence geometries and variable temperature and pressure regimes. Data acquired through RELAB form a foundation upon which many remote sensing algorithms have been developed and tested. The primary instrument is a high precision bidirectional reflectance spectrogoniometer that permits the measurement of samples under conditions similar to those for aircraft and space-based systems. The RELAB facility also has a Nicolet Nexus 870 Fourier Transform Infrared (FTIR) spectrometer acquired with funds from the Paduano foundation. This instrument measures biconical reflectance from 1.0 to 200 μm. Microscopic samples can be measured with a recently acquired FTIR microscope.

7.7 INFORMATION TECHNOLOGY AND THE VIRTUAL ASTROBIOLOGY COMMUNITY (B. OLSSON).

The MBL Astrobiology team believes that interdisciplinary collaboration is an integral part of astrobiological research, and that successful team efforts require clear communication, and a good collective knowledge of goals, priorities, and timeframes. We consider a solid information technology infrastructure to be a necessary, albeit not sufficient, prerequisite for building a virtual institute.

We will continue to be an active participant in NAI information technology pilots and projects, and intensify the use of tools already procured by the NAI collaborative research support

team. The WebEx conferencing system will be promoted in projects where it can improve the efficiency and quality of the scientific work, and we are preparing science projects for the incorporation of knowledge management systems planned for NAI pilots in the near future. The Polycom videoconference system was for a time underutilized due to ISDN connectivity problems, and because of internet bandwidth restrictions we were unable to find feasible alternatives. This situation changed dramatically when the Marine Biological Laboratory joined a partnership of local scientific institutions sharing multiple high speed connections to the Internet. The members of this partnership are the Woods Hole Oceanographic Institution (WHOI), the NOAA Northeast Fisheries Science Center, and the USGS Field Center for Coastal and Marine Geology, along with the MBL and a variety of smaller institutions.

The Partnership's Internet services include an OC3 running at 155 Mb/s for Internet2 services and various commercial T1's. The OC3 connects into the Internet2 POP at Northern Crossroads (NOX) in Boston, Massachusetts, which facilitates high performance academic networking in New England. NOX connects to Abilene, the Internet2 backbone, with an OC12 interface running at 622 Mb/s. The speed of this and the commercial T1's, combined, provide a basis for high performance networking and use of advanced network applications at the MBL, including the Polycom system, which has now been permanently switched over to IP-based communication. This has drastically reduced the videoconferencing costs while increasing the bandwidth, and has made our virtual meetings significantly more frequent and productive. A second improvement was last year's addition of new research space to the Bay Paul Center including the construction of a new conference room. We designed this facility to meet the environmental requirements for use of the Polycom by as many as 30 participants. As a result of the improved performance and the availability of a high quality environment, we have begun to meet our original expectations about use of the Polycom system for communication with members of our team at remote locations e.g. A. Teske at the University of North Carolina, colleagues at other NAI sites, etc. In a similar manner, this improved capability will have a major impact on our ongoing collaborative research efforts with the Centro de Astrobiologia. We also will take advantage of the increased connectivity by including desktop video conferencing in our teaching and research activities. Our collaboration with the Brown University Planetary Geology Group is especially exciting in this context, since they pioneered the use of international videoconferencing beginning with broadcasts and interactions with the Soviet Union already in the 1980's.

Technology in and by itself does not create a fully efficient virtual organization: Active and productive use of the technology requires demonstrations and training, and group efforts require a trust among the participants that can only be built during personal interactions in a non-technical context. In light of these cultural facts, we plan to place an emphasis on problem-oriented personnel development and regular face-to-face meetings where we can collect feedback and user metrics. Any identified problems will be solved in close collaboration between scientists and technology maintainers within our group. This blended approach will improve both the scientific and the information technology work, strengthen the sense of community, and create a more dynamic and innovative work environment.

While the formal incorporation of information technology has been successful at most NAI lead team sites, we believe there is room for improvement in establishing the *informal* features of our virtual community. This is clearly a concern from a cultural point of view, since most tacit organizational knowledge is gained through informal channels like hallway conversations, mentoring, etc. It is hard to envision a thriving multidisciplinary virtual institute lacking such casual communication. We will proactively address this problem by developing a web communication portal containing user-friendly tools for frequent and spontaneous online communication, both synchronous and asynchronous, to supplement the NAI tools already used for preplanned communications like seminars and focus group meetings. Some requirements for the web communication portal have been given especially high priority:

- Local serving to simplify maintenance and development/testing of customized functions for teams or individual researchers.
- Integration with customizable desktop audio and video "spur-of-the-moment" communication tools.
- Interactive discussion forums for all types of astrobiology dialogues.
- High availability over all bandwidths from all locations.
- Online streaming of stored videos from seminars and workshops.
- Consolidated dynamic and fixed storage of all types of project information (e.g., scientific data, EPO multimedia, software documentation, and streaming seminar videos).
- Well-organized and customizable user interface, to ensure active and productive use.
- Tools for collecting metrics and user feedback, to allow rapid realignment to current needs and fast identification of desired customized functions.
- High user- and group level security, efficient hacker defense, and reliable backup tools.
- Other planned functions include interfaces for distributed computing using XML-RPC, wireless access, and various Java tools for scientific data analysis and visualization.

To ensure cost-effectiveness and an active developer community the web communication portal framework will be built from the open source system PostNuke (http://www.postnuke.com). This content management system uses the freely available MySQL database for content and parameter storage, and provides functionality through modules written in PHP and JavaScript. Functions that cannot be found in open source PHP modules will be written by our own programmers. The communication tools will be designed from Flash Communication MX server modules, which in preliminary tests gave acceptable quality even over modem connections.

The construction of the web communication portal is an ambitious undertaking. However, the bioinformatics and database work routinely performed at the MBL makes us an ideal team for this kind of task, as illustrated by the **micro*scope** web site. A working prototype of the web communication portal containing a subset of the functions mentioned above was demonstrated at the 2003 NAI General Meeting in Tempe, Arizona. We are confident of our ability to create a product that will augment NAI information technology to the same degree that the **micro*scope** site has enriched the institute-wide EPO work.

The introduction of new information tools requires a very good rapport between researchers and technical staff. For that reason, we intend to create an Astrobiology Information Technology tiger team consisting of both IT personnel and researchers at MBL and Centro de Astrobiologia, Spain. This team will tightly integrate information technology and scientific work, and be responsible for the design, implementation, and testing of the web communication portal. The team will also do much exploration of procured NAI information technology in the laboratory and during field work atRío Tinto, Spain, described elsewhere in this proposal. Examples of planned activities include field data collection using wireless devices, frequent desktop video interaction and heavy use of knowledge management tools.

7.8 NAI COMMUNITY PARTICIPATION.

Videoconferencing and Director's Seminar Series. The MBL Astrobiology node regularly attends video presentations made possible through the Polycom system, broadcast from both NAI Central and other sites, and looks forward to continued participation in these opportunities to share scientific results. Team member Linda Amaral Zettler will participate by presenting an upcoming Director's Seminar Series lecture entitled "Life at pH Extremes". Lorraine Olendzenski has

presented the activities and teaching model used in the *Living in the Microbial World* to Institute members via videocon.

Participation in the NAI Focus Groups. Members of the MBL Astrobiology team participate in the EcoGenomics focus group and have been an integral part of a collaborative study on the microbial mat community of Guerrero Negro, Baja California Sud. Our research efforts on expression profiling of the cyanobacterium *Microcoleus chthonoplastes* from Sippewissett Salt Marsh under different light regimes, have served to work out a reliable protocol for further investigations on organisms from Guerrero Negro. This work forms the basis of future comparative projects on organisms from these complex communities under different environmental conditions, including salt concentration and sulfate concentration. We intend to continue our active participation in this Focus Group to obtain data delineated to be of interest to the group as a whole. Jim Head and Graduate Research Assistant G. Wesley Patterson are both active members of the NAI Europa Focus Group.

Collaborations with URI Deep Biosphere group. The URI Astrobiology group focuses on biogeochemical and microbiological analyses of the deep subsurface biosphere. Here we are in an excellent position to collaborate on the microbiology of extreme deep subsurface microbiota, facilitated by proximity and personal ties (A. Teske is also a member of the URI team). Through research cruises on deep subsurface drill ship JOIDES Resolution, the URI group directly probes the deep marine subsurface biosphere, which is out of reach by any other means. So far, extensive sample sets have been collected from hydrothermally active accretionary margins (Nankai Trough, ODP leg 190), from inverted methane-sulfate gradients in the equatorial Pacific and the Peru Margin (leg 201), from methane-hydrate-dominated sediments (Hydrate Ridge offshore Oregon, Leg 204), and deeply buried black shales (Demarara Rise, leg 207). Initial 16S rRNA analyses of Nankai Trough sediments have revealed highly unusual archaeal and bacterial assemblages. In particular, our prokaryotic projects on diversity and evolution of anaerobic and autotrophic metabolisms and their key genes, will benefit from this collaboration.

Collaborations with the Centro de Astrobiología (CAB) in Spain. Our proposed work on the Rio Tinto will strengthen existing collaborative research that is ongoing between the MBL and the CAB. We continue to rely on the laboratory of Dr. Ricardo Amils as a base for processing samples from field studies. This includes chemical analyses of water samples collected as part of our population biology studies, as well as for molecular biology protocols that requires immediate processing. Dr. Amils has also agreed to prescreen Rio Tinto fungal isolates that his lab maintains in a culture collection as part of our investigation into iron oxidation metabolism in eukaryotes. As part of our population biology studies we will coordinate our sampling efforts with those of our Spanish colleagues and freely share data. We will take full advantage of the IT resources available to us to foster good communication and exchange of information. In the past, we have hosted scientists from the CAB in our lab to be trained in phylogenetic methods and inference and intend to continue to serve as a resource to members of the CAB.

Other Collaborations. The Astrobiology community also shares diverse and rare sample materials from extreme environments and microbial ecosystems, such as hydrothermal vent and deep subsurface sediments and water samples. These can be shared between Astrobiology partners on a flexible basis for projects of mutual interest, as exemplified by our successful collaboration with David A. Stahl (UW), on molecular analysis for deeply-branching sulfate reducers in Guaymas Basin hydrothermal sediments.

Relationships to the Space Science Community. PI Mitchell L. Sogin is a member of the Space Sciences Board Executive Committee of the National Research Council. The Space Sciences Board conducts advisory studies and program assessments, facilitates international research coordination, and promotes communications on space science and science policy between the research community, the federal government, and the interested public. He is also a member of the science definition team of the Jupiter Icy Moons Orbiter mission (JIMO). Dr. Norman

Wainwright has been an active participant in Planetary Protection Workshops sponsored by NASA. EPO Lead Lorraine Olendzenski currently collaborates with a team under the direction of NASA Planetary Protection Officer John Rummel to design and deliver training workshops in Planetary Protection for managers and technicians involved in meeting NASA requirements for spacecraft.

Joint Workshops, Seminar Series and Courses with the Center for the Advanced Study of Space Life Sciences (CASSLS) at the MBL. Over the last five years, we have collaborated on a number of research, education and outreach efforts with the Center for the Advanced Study of Space Life Sciences located at the MBL (see letter of support), including workshops and seminar series. Funded through a cooperative agreement between the MBL and the Life Sciences Division of NASA, CASSLS acts as an interface between NASA and the basic sciences community, promoting interactions and discussion in areas of mutual interest. As a direct result of input from PI Mitchell Sogin and our Astrobiology group, CASSLS sponsored the workshop *Outcomes of Genome- Genome Interactions*. This meeting brought together a diverse group of investigators comprised of microbiologists, biogeochemists, ecosystem experts, molecular phylogeneticists and molecular ecologists to foster discussions about how to link biogeochemical measurements with metabolic processes and microbial population structures in natural settings. This type of cross-disciplinary interaction is necessary to advance experimental design and subsequent understanding of how microbes function in consortia and about the responses of structured microbial communities to cyclic and transient environmental cycles. The proceedings of this meeting will be published in The Biological Bulletin in April 2003. Additionally, we work with CASSLS personnel to offer the *Life and Living in Space* Teacher Workshop. NASA Life Sciences also provides support to the *Advances in Genome Technology and Bioinformatics* workshop.

Flight Missions. Jack Mustard is an official collaborator with the French-built OMEGA experiment on ESA's Mars Express mission, which will launch in May 2003. He is also a co-investigator on the Compact Resolution Imaging Spectrometer for Mars (CRISM), which is scheduled to fly on the Mars Reconnaissance Orbiter (MRO) in 2005 and collect high spatial resolution spectroscopic data through 2009. He participated in the planning for the instrumentation on the Mars Odyssey and MRO, and recently participated in the Mars Pathways steering group. He is participating in several efforts to develop flight hardware for upcoming missions.

Jim Head is a co-investigator on the Mars Global Surveyor Mars Orbiter Laser Altimeter experiment, and a co-investigator on the Galileo Mission Solid-State Imaging System, being responsible for planning one half the Europa encounters. He is also a co-investigator on the Mars Scout ARES mission, proposed to be the first airplane flight over the surface at low altitudes. He is a co-investigator on the Mars Express High Resolution Stereo Camera (HRSC) and on the Discovery MESSENGER to Mercury. In the past he has participated in planning and execution of Apollo, Viking Lander, Magellan, Soviet Venera 15/16 and Phobos missions.

Links to Other Agencies. Jim Head participates in the National Science Foundation Antarctic Research Program in the Dry Valleys, investigating Mars analogs. He recently completed a six-week deployment and is presently preparing the results for publications.

8.0 FACILITIES

The Marine Biological Laboratory at Woods Hole MA

The Josephine Bay Paul Center at the MBL has approximately 10,000 square feet of recently renovated laboratory space for molecular biology. This space includes rooms for microscopy, tissue culturing, cold rooms, dark rooms, and a climate controlled room for housing larger computer systems. We have a dedicated 1800 square-foot laboratory for automated DNA sequencing, DNA microarraying, and high-throughput robotics. A separate computer room houses our Hewlet Packard and Dell File servers.

Computational capabilities: The Josephine Bay Paul Center maintains extensive high-performance computing capabilities for genomic, phylogenetic and bioinformatic analyses as well as web server presence. We have nine Alpha systems including a Compaq EV-6 DS10 server with over a terabyte of memory storage, a dual Compaq EV-6 DS20 server with 4Gb RAM, and a newly acquired large memory (8 gigabytes) Hewlett-Packard EV-68 ES-45 server. For high performance parallel processing we operate a Beowulf Linux cluster containing 108 AMD Athlon XP and Intel P4 processors. Scattered around the laboratory we have dozens of Intel P4 servers and workstations running various versions of Linux and Windows, Macintosh G4 workstations running Apple OSX, and Sun SPARC workstations running Solaris 9. The centralized data storage consists of a 2.5 terabyte Apple Xserve RAID connected over fiber-channel to a dual-processor 1.33GHz G4 Apple Xserve that provides connectivity to Linux, Unix, Windows and Apple systems over a 100BaseT switched network. Data integrity is ensured by a dedicated Intel P4 backup server controlling a 1.6 terabyte RAID and a 100Gb Ultrium LTO tape-drive. All systems handling large data sets will soon be connected directly to the RAID device, to form a fiber-web SAN that will further increase the analysis throughput. High-availability web presence is ensured by four dedicated Intel P4 systems running Microsoft IIS 5.0 under Windows 2000 and Apache under Linux, and which serves mySQL and Filemaker databases, dynamic PHP sites like micro*scope and a web communication portal, Flash Communication MX servers, and various interactive Perl scripts.

We actively maintain locally installed genomics and bioinformatics packages such as Phred, Phrap, SEALS, DARWIN, Accelrys GCG and the Arachne whole-genome shotgun assembler. We also maintain local Genbank, Swissprot and Pfam databases, phylogenetic analysis programs including PAUP, PHYLIP, PUZZLE, MEGA, etc. and a host of supporting software for examination, visualization and presentation of data. A full time administrator with special expertise in astrobiological research and scientific programming, and a database specialist highly knowledgeable in PHP, SQL and design of interactive database-driven web sites manage the systems and network.

Office: Dr. Sogin has an ~150 square foot office and our administrative assistants have 160 square feet of office. The institution provides purchasing, personnel, and grants administration services. The Josephine Bay Paul Center provides office and laboratory space for seven other senior scientific investigators. Our offices and laboratories lie adjacent to the MBL library, which subscribes to a wide variety of journals and provides electronic access to many databases.

Other: General equipment in our laboratories include a very well accessorized Zeiss Axioskop 2 microscope equipped for fluorescence microscopy and digital imaging, modern preparative and table-top ultracentrifuges plus rotors, three high speed Sorvall centrifuges plus rotors, a Beckman DU 70 UV spectrophotometer, SpeedVac concentrators a Virtis freeze dryer/vacuum pump system, centrifuges plus rotors for microplates, multiple bench top water baths, gyratory shaking

water baths, heating blocks, analytical balances, a pH meter, several incubators and microplate shakers for 96-well cultures, six large -80°C freezers, several -20°C freezers and refrigerators, fraction collectors, column monitoring equipment, film processor, and cell disruption equipment. Equipment for recombinant DNA work and molecular biology include an Eppendorf electroporation apparatus, ten gradient thermocyclers, a rotary hybridization oven, a dozen high and low voltage power supplies, miscellaneous horizontal gel boxes, trans-illuminators, an Ultra-Lum gel documentation system, and a variety of manual and electric pipetters, an Amersham-Pharmacia FPLC and a Storm 860 Phosphorimager/Fluorimager.

Major Equipment: For high-throughput sequencing, the Bay Paul Center operates an Applied Biosystems 3730 XL capillary sequencing system, a Beckman CEQ2000 capillary system, and seven LI-COR Long ReadIR, two-channel DNA sequencing systems. We have a GeneMachines RevPrepOrbit II for automated, production of ~3600 sequencing DNA templates/day, a Tecan Miniprep75 robot for the preparation of DNA sequencing reactions, a GeneMachines Mantis colony picker, an OmniGrid DNA arrayer, an Axon array reader, and a BioRad CHEF-Mapper pulsed-field gel electrophoresis system for separation of chromosomes and high-molecular weight DNA. For cycle-sequencing, we have several 96-well and 384-well dual-block AppliedBiosystems 9700 thermal cyclers. All of this equipment is located in the **W.M. Keck Ecological and Genetics Facility**, which is housed and managed by the Josephine Bay Paul Center at the MBL.

Microscopy: The MBL maintains a complete facility for state of the art electron microscopy and new higher performance confocal microscopy. Equipment includes a Zeiss 10CA transmission electron microscope and a JEOL 840 scanning electron microscope (with cryo stage and EDS), supported by preparative and ancillary equipment such as ultramicrotomes, sputter coaters, etc. and a Balzers freeze-etch apparatus

Conference Facilities: The MBL has a large modern dormitory (the Swope Center) in the village of Woods Hole, with close proximity to banking, shopping, restaurants, and bus lines connecting to major airports in Providence, RI and Boston, MA. The Swope Center also provides access to computers, online email, the WHOI/MBL library, plus catered dinners. MBL also operates the J. Erik Jonsson Center of the National Academy of Sciences (http://www.mbl.edu/housing/jonsson/) with facilities for meetings and conferences, located just outside of the village of Woods Hole.

Woods Hole Oceanographic Institution

The Laboratories of Katrina J. Edwards consist of three rooms, ~1400 ft^2 total in the Department of Marine Chemistry and Geochemistry. Facilities and equipment exist for standard microbiological, molecular biological, aqueous geochemical, and light, fluorescence, and atomic force microscopy and imaging. Within Edwards' department, facilities exist for a variety of chemical and geochemical analysis of both solid material and aqueous solutions (organic & inorganic).

Laboratory Equipment: Laminar flow bio-hood & chemical fume hood, Barnstead Water distillation & purification (UV/UF), Shimatzu UV-1601 Spectrophotometer & printer, Autoclave, Beckman High-speed refrigerated centrifuge , Controlled temperature & benchtop shakers, Benchtop microfuges, Benchtop pH meter, Mettler Analytical balance, Unisense clark-type microelectrodes (for O_2, H_2S, pH), PA2000 picoammeter, MM33 micromanipulator (apparatus also suitable for field usage), Stirring hot plates, -20 & -80°C freezers, & 4°C refrigerators, Fisher Water bath & dry-block incubators, Fisher & Hibid Hybridization ovens, Sonicator, Vortexes, Perkin-Elmer PCR machines, Biorad iCycler PCR machine, Transilluminator, photodocumentation, & electrophoresis power supplies, Invitrogen E-gel mini-

electrophoresis apparatus, Pressure-filtration equipment, Bead beaters, Pipetters, Multiple computers, printers, software

Microscopic analysis*:* ZEISS Axiovert ST100 TV research grade microscope, Digital Instruments Atomic Force Bioscope (AFM). This microscope can be used in conjunction with the inverted ZEISS, or independently as a free-standing AFM, 2 Vibration isolation tables, HBO100 mercury-arc lamp, Filters for UV, Rhodamine, FITC, OG, CY3, and A-Pol (for rDIC) fluorescence, HAMAMATSU Orca CCD video camera (B/W) & Polaroid color camera, Nikon D1 Digital camera (for microscope use or with independent lens), OPENLAB software modules for the Macintosh computer for image capture and analysis, Nikon, Photoshop, NIH-image, and Digital Micrograph software for image capture and analysis, 10X, 20X, 40X (+ rDIC prism), & 63X (+ rDIC prism; oil) objectives

Isotope Ratio Analysis (MC-ICP-MS): The Woods Hole Oceanographic Institution Inductively Coupled Plasma Facility is an analytical service lab that provides trace element and isotopic analyses for a variety of clients within and outside WHOI. The instrumentation is based on the argon inductively coupled plasma. Currently the ICP Facility operates a high resolution single collector ICP-MS, the Element1, and a high resolution multicollector ICP-MS, the Thermo Finnigan Neptune MC-ICP-MS. The argon plasma source provides a high ionization potential and the ability to atomize aqueous as well as gas sample sources. This makes available approximately 75% of the elements in the periodic table for analysis. The instrumentation housed in the ICP Facility is the best of this type of instrument currently available. The Element1 is a general purpose ICPMS that can sequentially scan masses from 4 to 240 atomic mass units. This instrument is generally used for trace element concentration measurements with detection limits ranging into single parts per quadrillion (10^{-15} or 0.001 parts per trillion) to parts per billion (10^{-9}) range. The Element1 is also used for isotope ratio analyses. The Neptune is expressly designed to make high precision isotope ratio analyses for many different isotope systems. Due to its ion optic design the Neptune can operate in high-resolution mode eliminating common interferences for elements such as sulfur and iron. The facility is staffed with highly qualified technical assistants who are aware of the proposed astrobiology project and are eager to assist with the analytical and methodological development described in the text.

The laboratories of Eric A. Webb in the WHOI Department of Biology are well equipped, modern laboratories of approximately 1500 ft^2 that are set up for microbiology and molecular biology. These ancillary lab facilities will be used for some biochemical analysis (see letter). They include three separate laboratories; one specialized for culture and physiological studies, one equipped for molecular biology and one for the use of radioisotopes. Available equipment includes: Biorad electrophoresis equipment, Perkin Elmer PCR machine, Biorad iCycler quantitative PCR machine, Eppendorf gradient Master cycler and thermomixer R, column chromatography equipment, Shimadzu 1601 UV Spectrophotometer, Beckman and Packard Liquid Scintillation Counters, Beckman High Speed Refrigerated Centrifuges, Beckman Preparative Ultracentrifuge, Pecival incubators

Dr. Stephan Sievert maintains a fully equipped, 980 sq. ft. microbiology and molecular biology laboratory in the Department of Biology, Redfield Laboratories of WHOI, with special capability in the culture of anaerobic prokaryotes. Available equipment includes spectrophotometers, autoclaves, hybridzation oven, scintillation counters, high speed centrifuge and microfuges, a Zeiss Axioplan 2 epifluorescence research microscope with digital imaging capabilities, an inverted microscope with micromanipulator, pH meters, incubators, Coy anaerobic hood, culture rooms, and Hungate apparatus for anaerobic culture. Equipment for PCR (PE GeneAmp PCR 9700 thermocycler), quantitative PCR (Biorad iCycler), and Denaturing Gradient Gel Electorphoresis (DGGE) (Biorad Dcode) is available as well as capability for Fluorescence In Situ Hybridization (FISH) and DNA sequencing. For initial preparation of clones for high

throughput sequencing, an automated Mantis colony picker is housed within the WHOI Department of Biology. Sequencing is accomplished using the facilities at the Josephine Bay Paul Center, MBL.

Brown University

Earth and Planetary Image Processing System: A network of computers to support image processing and data analysis is housed in the laboratories of the Planetary Geology group at Brown University. These include a network of Sun workstations (1 Blade 2000, 5 Blade 150's, 5 Ultra 10 Creator #D, 1 UltraSparc 30 Creator 3D) and a Sun V880 server with 8 processors, 32 Gbytes of RAM and over 1.5 terrabytes of storage. Mission data analysis is accomplished using the VICAR and ISIS software developed by JPL, EASI/PACE software package, ENVI for hyperspectral analyses, and the full ARC/INFO software package including ArcView. Additional software has been and is continuing to be developed at Brown to implement new analysis techniques. A number of Apple and Microsoft operating system microcomputers are also available.

RELAB: The Reflectance Experiment Laboratory (RELAB) is a multiuser facility housed at Brown University for the acquiring high precision and accuracy reflectance properties of materials from 0.3 to 25 μm. Funded under a joint agreement between NASA and Brown, RELAB accepts samples from NASA funded investigators. Data acquired through RELAB form a foundation upon which many remote sensing algorithms have been developed and tested. The primary instrument is a high precision bidirectional reflectance spectrometergoniometer that permits the measurement of samples under conditions similar to those for aircraft and space-based systems. Measurements can be made under variable incidence and emergence geometries and variable temperature and pressure regimes. Details of the instrument capabilities are contained in the user's manual available on request. The RELAB facility also has a Nicolet Nexus 870 FTIR spectrometer acquired with funds from a much appreciated grant from the Paduano foundation. This instrument measures biconical reflectance from 1.0 to 200 μm. Recent acquisition of FTIR microscope allows the measurement of microscopic samples.

Field Spectrometers: Two Analytical Spectral Devices field spectrometers (a PSII and a FieldSpec FR) are available for use. The PSII has a 512channel detector that covers the wavelength range from 0.3 to 1.05 μm, while the FieldSpec FR covers the wavelength range from 0.35 to 2.5 μm. Both instruments are portable and simple to use, and acquire high quality data. In addition, we have an Oriel 100 watt quartzhalogen radiometric lamp and power supply for laboratory measurements.

General Facilities: Facilities to determine sample composition are available at the Department of Geology at Brown which include: a microprobe, X-ray diffraction, X-ray florescence, and a wet chemistry laboratory to determine Fe^{2+}/Fe^{3+} ratios, equipment for sedimentological analysis (grain size and geophysical properties), and uniaxial stress analysis. The geochemistry laboratories are well equipped for the application of chemical and particularly isotopic techniques. Equipment includes: automated mass spectrometers capable of analyzing both light and heavy elements, including O, C, Rb-Sr, REE and U-Th-Pb; a clean laboratory for the preparation of samples for isotopic analysis, an XRF spectrometer for major and minor element analysis; all the necessary chemical and vacuum equipment to support mass spectrometer operations; and the ancillary equipment used in mineral separation, radiochemical storage, etc., a Cameca electron microprobe analyzer.

University of North Carolina, Chapel Hill

Dr. Andreas Teske's laboratory (started in Sept. 2002) currently contains molecular biology equipment for PCR, Denaturing gradient Gel Electrophoresis (DGGE), and digital gel image analysis and documentation. Spectrophotometers, electroporators, and incubation

chambers for transformed cells are available and updates to real-time PCR are planned. A new deionizing/UV water purification system and -80°C freezer dedicated to sample storage are in place. Current computers(Macintosh G4 and G3) allow medium-volume sequence alignment and phylogenetic analysis; the computer facilities will be updated with a Linux workstation in order to run the high-throughput phylogeny package ARB. Automated 96-well plate high throughput sequencing facilities are available in the sequencing center of the University, although the laboratory generally plans to collaborate with MBL on sequencing. The Department of Marine Sciences provides a permanently installed Polycom Viewstation SP videoconferencing system (with 50-inch plasma display screen).The molecular biology equipment will be complemented with the setup for anaerobic microbial cultivations, including anaerobic hood and Hungate gassing station for anaerobic culture during the year 2003.

The current laboratory space (ca. 700 square feet) and office space (ca. 200 square feet) accommodates at present two Ph.D. students and two Postdoctoral. The University of North Carolina at Chapel Hill has started an ambitious building program for a new Science Complex, and will provide new, personally customized laboratory and office space in for the Teske group in winter 2004/2005 (Phase I of the Science complex).