Michigan State University

East Lansing, Michigan

Center for Genomic and Evolutionary Studies on Microbial Life at Low Temperature

Principal Investigator:

Michael Thomashow, Michigan State University

Co-Investigators/Collaborators:

Lycely Sepulveda, University of Puerto Rico, Mayaguez
J. Tiedje, Michigan State University
David Lubman, University of Michigan
Sophia Kathariou, North Carolina State University
John McGrath, Michigan State University
Richard Goldstein, University of Michigan
Elbert Branscomb, DOE Joint Genome Institute

Table of Contents

Executive Summary………………………………………………………………….. 2

Research/Training/Management Proposal…………………….………………………… 5
I. Research………………………………..…………………………………….. 5
A. Background Information…..………………………………………… 5
B. Preliminary Results……………………………………………… 16
C. Experimental Plan…………………………………..…………… 29
D. Expertise and Responsibilities of Co-Investigators……..……………. 56
II. Training……………………………………………………….……………. 59
A. Educational Programs………………………………………………. 59
B. Internal Seminars…………………………………….……………… 60
C. Research Exchanges…………………………………...………. 60
D. Symposium on "Genomics and Evolution of Extremophlies"……….. 60
III. Management……………………………………………………...…… 61
A. Center Programs…………………………………………………….. 61
B. Minority Program………………………………………………..…. 62

References…………………………………………………………………………..… 63

Executive Summary

Low temperature is a predominant environmental characteristic of interstellar space, our solar system, including most of the planets and their satellites, and asteroids and meteors. An understanding of the impact of low temperatures on the responses and evolution of biological organisms is, thus, integral to our knowledge of Astrobiology. The research that we propose will explore multiple aspects of microbial adaptation to low temperatures. One major line of investigation will be to conduct structural and functional genomic and proteomic analyses of bacteria that have been isolated from the Arctic and Antarctic permafrost. What genes and proteins enable the permafrost bacteria to inhabit these subfreezing environments? Do they have specific "freezing tolerance" genes and proteins, or "specialized alleles" of commonly found bacterial genes, or both? How is expression of the bacterial genome affected by low temperatures and other conditions that "hitchhiker" bacteria might encounter during travel through space on natural objects or spacecraft? In a second line of investigation, we will directly examine, through "test-tube evolution" experiments, bacterial adaptation to low temperatures. The studies will provide insight into how an organism, with a given complement of genes, can cross niche barriers that are defined by decreasing temperatures. And finally, we will use the information gained to explore the potential development of "signatures" for the presence of life in cold environments including Earth and other bodies such as Mars and Europa.

The proposed studies will provide significant new information relating to multiple goals outlined in the "Astrobiology Roadmap" including: Explore how life evolves on the molecular, organism and ecosystem levels (goal 3); Determine how the terrestrial biosphere has co-evolved with the Earth (goal 4); Establish limits for life in environments that provide analogues for condition on other worlds (goal 5); Determine how to recognize the signature of life on other worlds (goal 7); and Understand the response of terrestrial life to conditions in space or on other planets (goal 10). Moreover, the proposed studies address one of the perceived gaps in the current NAI research program, namely, "**provide understanding of the response of life to the space environment, from gene expression to microbial evolution**." Finally, there are significant potential practical applications of the work ranging from the identification of genes that may be used to confer improved environmental stress tolerance in crop plants to the discovery of enzymes uniquely suited to catalysis at low temperature, a characteristic of importance in numerous biotechnology applications.

The proposed lines of interrelated investigations call for a broad range of expertise. Consequently, we have assembled a group of investigators with diverse backgrounds, training and research interests. The areas of expertise include microbial ecology (Tiedje, Sepulveda), the isolation and characterization of permafrost bacteria (Tiedje), molecular genetics and gene regulation (Kathariou, Thomashow), evolution and population genetics (Lenski, Bennett), cryobiology and mechanisms of freezing tolerance (McGrath, Thomashow), proteomics and protein evolution (Lubman, Goldstein), and structural genomics (Branscomb, Hawkins, Predki). As the proposed lines of research develop, we anticipate that the Center will evolve, bringing in new investigators to provide expertise to attack the next generation of questions. For instance, the structural genomic and gene expression profiling studies will lead to lines of investigation focused on specific genes with the objectives of establishing their roles in cold tolerance; determining their modes of action; and developing hypotheses as to how the genes evolved.

While the current Center investigators are well qualified to conduct these lines of research, it may, perhaps, become evident that adding an investigator with expertise in using X-ray crystallography and/or NMR-spectroscopy to determine protein structure-function relationships would strengthen the research efforts.

We also envision that additional lines of investigation will evolve out of our interactions with researchers in other Centers of the Astrobiology Institute. For instance, among our objectives are to conduct experiments to determine how bacterial gene expression and evolution are affected by conditions that relate to the space environment. Our major focus will be on cold temperatures. However, it seems likely that it would be of interest and importance to other Center members to understand how microorganisms react, in terms of gene expression and evolution, to other environmental conditions of space such as the Martian atmosphere and regolith or perhaps microgravity and solar radiation. We view the defining of mutual research interests and development of collaborative Center research efforts to address fundamental questions in astrobiology to be an underlying goal of our research and educational programs.

Our educational objective is to provide in-depth training in structural and functional genomics, but to encompass that training in a context relevant to astrobiology. In this way our students will have cutting-edge knowledge for the genomics era, but will be able to use that knowledge to explore questions important to understanding how biological organisms may adapt to and evolve in response to extreme environmental conditions present in the universe. We will implement this educational objective by having students major in their existing degree-granting programs, but establish an "area of specialization" in astrobiology. In this way students will not be diluted in their training in their basic discipline but will gain the perspective and knowledge needed to make significant contributions to the field of astrobiology. The specialization approach will also make it possible for students from a variety of departments to enter the program and gain an astrobiology emphasis. The astrobiology specialization will comprise coursework, including a new graduate level course in "Genomics and Evolution of Environmental Stress Tolerance" and short-courses to augment keys areas of student training. One such course that we have in mind initially would focus on the cryosoil environment and history of the Arctic and Antarctic, drawing similarities and dissimilarities to Europa and Mars. Other components of the astrobiology specialization will include an internal seminar series; research exchanges in which graduate students or postdocs go to another institution either within our Center or within the NAI to conduct a collaborative experiment that draws on the unique resources of the partners; and a symposium on "Genomics and Evolution of Extremophiles" (to be held in year three or four) where the research results from our Center will be presented along with findings by prominent invited speakers in related fields. The objective will be to bring together the cutting-edge body of knowledge on the symposium topic for Center members, NAI members and interested peers.

At present, we are not proposing a separate E/PO program. However, what we would like to suggest to the NAI, given our theme of genomics, is our developing a high school level genomics module that incorporates concepts central to astrobiology. Ideally, students would be able to explore actual extremophile genomic data using well constructed queries. The information they learn could be put in the context of "the limits of life." In addition, for the K-8 group, we suggest an update and take-off on the "Microbial Zoo" unit which the CME developed

with MSU's Communications Technology Laboratory. The module, which is available on the Web (http://commtechlab.msu.edu/sites/dlc-me/zoo/) and on CD-ROM (with sound) allows young people to explore environmentally important microbes as if they were in a zoo. Already included in the module is "Space Adventure" (http://commtechlab.msu.edu/sites/dlc-me/zoo/zsmain.html) which includes Martian meteorite images as a center-piece. This is an existing example of our contribution to astrobiology and what we have done for the K-12 sector. The resources necessary for development a full K-14 astrobiology Web site (including the professional educator involvement and design testing on target groups) is beyond the resources available for an individual Center activity, but should not be beyond what the NAI could do. We would have the scientists, communication technologist and educator experience to share if NAI undertook such an effort.

Professor Michael Thomashow will be the PI and Director of the Center and Professor James Tiedje will serve as Associate Director to aid the Director in duties determined by the Director. We have found such a management scheme to work well in our NSF-STC Center for Microbial Ecology; one person has point responsibility, but has a second person fully cognizant of the breadth of Center activities with whom to discuss pertinent issues and implement actions. In addition, there will be an Executive Committee consisting of three co-investigators (membership will rotate). The Executive Committee will establish operating policy, aid in making major decisions, make sure that all perspectives are heard and help ensure that the research being conducted is of the highest possible quality. In addition, we will name a small group of external advisors for the purpose of providing Center management with critique and suggestions on Center activities, progress and direction. After year two, we would specifically ask the group for advice on mid-course adjustments so that these could be implemented before the year three NASA review.

Our proposed Center is strongly supported by the institutions involved and includes cost sharing that will significantly enhance the proposed research effort. The institutional financial commitments include: 1) three graduate student lines in years 1 and 2, and four in years 3-5 plus all faculty and administrative salaries at Michigan State University to be provided by MSU, together totaling approximately $900k; 2) some 30 million base-pairs of DNA sequence provided by the Joint Genome Institute, a value of approximately $100k; and 3) virtually all faculty salary provided by the partner universities at an estimated value of $300k. **By not burdening the Center with faculty salaries, most NASA funds will go directly to researchers in the laboratory, at their computers, or in the field.**

Research/Training/Management Section

I. Research

The overarching research theme of our Center is understanding biological adaptation to environmental conditions that are common in space. Our focus will be on microbial adaptation to low temperature. Cold temperature is a predominant environmental characteristic of interstellar space, our solar system, including most of the planets and their satellites, as well as asteroids, comets and other objects. An understanding of the impact of cold temperatures on the responses and evolution of biological organisms is thus an integral component of an overall knowledge of astrobiology.

In one line of investigation, we will focus on structural and functional genomic analyses of bacteria that are present in permafrost environments on Earth. As discussed below, the permafrost environment provides an excellent model "system" to study microbial adaptation to cold temperatures and offers the unique opportunity to study organisms that have survived freezing for more than a million years. Fundamental objectives of this work include determining how expression of the genome is affected by low temperature and dehydration, two cardinal characteristics of subfreezing environments, and identifying genes and proteins that enable bacteria to inhabit cold environments. As a part of this investigation, we will include a targeted search for, and basic characterization of, two classes of proteins associated with freezing tolerance: antifreeze proteins and cryoprotective polypeptides. In addition, we will conduct "Field Truth" experiments to probe for the presence of hypothesized key cold tolerance genes and proteins *in situ* with an associated objective of developing potential "signatures" indicative of present or past life. Finally, we will conduct "test-tube" evolution experiments with *Escherichia coli* to identify genes that, through natural mutation, can impart increased "fitness" to cold temperatures. Through these studies, we will explore how an organism, with a given complement of genes, may cross niche barriers that are defined by low temperature.

The proposed studies will provide significant new information relating to multiple goals outlined in the "Astrobiology Roadmap." The include: Explore how life evolves on the molecular, organism and ecosystem levels (goal 3); Determine how the terrestrial biosphere has co-evolved with the Earth (goal 4); Establish limits for life in environments that provide analogues for condition on other worlds (goal 5); Determine how to recognize the signature of life on other worlds (goal 7); and Understand the response of terrestrial life to conditions in space or on other planets (goal 10). Moreover, the proposed studies address one of the perceived gaps in the current NAI research program, namely, "provide understanding of the response of life to the space environment, from gene expression to microbial evolution." Finally, there are significant potential practical applications of the work ranging from the identification of genes that may be used to confer improved environmental stress tolerance in crop plants to the discovery of enzymes uniquely suited to catalysis at low temperature, a characteristic of importance in various biotechnology applications.

A. Background Information

1. The permafrost environment and microbial community: a model for astrobiology

The ability of microorganisms to live in deep cold environments continually forces us to redefine the spatial and temporal limits of life in our biosphere and other bodies in space. From an astrobiological point of view, seven of the nine planets of our Solar System, as well as their satellites, comets and asteroids have a cryogenic nature, i.e. the permafrost is a common phenomenon in space. Here on Earth, permafrost is also an important environmental niche. Permafrost, defined as ground (including bedrock, soil and sands) that remains below 0°C for more than two years (Muller, 1943), makes up more than 20% of the land surface of the Earth, including 82% of Alaska, 50% of Russia and Canada, 20% of China, and most of the surface of Antarctica (Harris, 1986; Williams and Smith, 1989; Storad, 1990). Permafrost poses unique challenges to its resident biota:

- Temperatures that have remained below 0°C over geologically significant periods of time.
- Presence of only minute amounts of water in the liquid form, 93-98% of water being in solid (ice) form.
- Inaccessibility of organic matter to the resident microbes, due to the frozen state.
- Minimal diffusion and infiltration of water, organic compounds and minerals.
- Long-term influence of gamma radiation from soil minerals.

These challenges, considered alone and in combination, appear severe. Permafrost, however, is also a stable environment, without sharp fluctuations. A unique outcome of this specialized ecosystem is the preservation of life for the longest documented periods on Earth. Viable bacteria in permafrost soils were first discovered in Russia in connection with investigations of mammoths in Siberia, and reported sporadically thereafter (Becker and Volkmann, 1961; Cameron and Morelli, 1974). Most recently, and almost unnoticed by the western scientific community, Russian scientists under the leadership of D. Gilichinsky reported the existence of high numbers of viable microorganisms (up to 10^5-10^7 cells/g) in buried Siberian permafrost soils of depths varying from a few meters below surface to 400 m (Zvyagintsev et al., 1985; Zvyagintsev et al., 1990; Khlebnikova et al., 1990, Gilichinsky et al., 1988; Gilichinsky et al., 1992). This was surprising, not only because of the constant subzero temperature of the soils, averaging -10°C to -12°C, but also because of the length of time over which the soils were frozen, from a few thousand years up to 2-3 million years. These organisms may well be the only living cells that have survived for a geologically significant period of time. Although viable eukaryotes (yeasts, fungi, mosses) have been isolated from the permafrost, their numbers seem to be low in comparison to those of prokaryotes, especially in samples older than 10,000 years (Gilichinsky et al., 1992).

The permafrost microbial community has been described as "a community of survivors" (Friedman, 1994). Viable bacterial populations in the permafrost may be viewed as the result of a continuous process of selection for those capable of withstanding prolonged exposure to subzero temperature, accompanying low water activity, and other special challenges posed by the permafrost. Furthermore, they are also the outcome of tens of thousands of years of cyclic freeze-thaw selection in the original "active" tundra layers where they were exposed to warmer growth periods in summer and colder surface temperatures in the winter. Enumeration of culturable bacteria has yielded minimal estimates of the number of viable bacteria in the buried

permafrost, and some data suggest that the ratio of readily culturable (hypometabolic) bacteria to those that are in the viable but nonculturable state (deep resting cells) is determined by the length of time for which the buried soils have been frozen (for review, see Vorobyova et al., 1997). Several lines of evidence indicate that ***the microorganisms that have survived in the buried soils are not simply representatives of cold-region surface soil communities.*** The oldest samples have relatively low numbers of culturable bacteria, 10^2-10^4 cells per gram in Pliocene soil of 1 million years or older. Bacteria surviving in these ancient samples were frequently Coryneforms, whereas actinomycetes were rarely isolated from samples below the Holocene stratum (10^4 years). Importantly, non-spore forming bacteria always dominated, independent of depth or age (Gilichinsky et al., 1992). The prevalence of non-spore formers was also found in Antarctic soil (Horowitz et al., 1972), but not in core samples of the Antarctic ice sheet; Abyzov et al. (1987) reported increasing numbers of spore-forming bacteria with increasing depth (down to 320 m) and age of the ice cores (up to 12,000 years). The glacial ice, therefore, appears to provide an environment that supports a fundamentally different, much less dense, microbial community than that supported by the permafrost environment.

If non-spore-forming bacteria are dominant in ancient frozen soil samples, they have either survived in a state of suspended animation (anabiosis) or are metabolically active. There are several indications of metabolic activities *in situ*. Examples include the simultaneous presence of nitrifying bacteria and nitrites, of methanogens and methane, and the presence of fresh ferrous sulfite (Gilichinsky and Wagener, 1994). In addition, it has been estimated that in permafrost, background radiation (about 0.03 rad/yr taking into account the protective effect of frozen soil) would have lethal cumulative effects on the genomes of the endemic bacteria in about 200,000 years. Thus, survival of cells in frozen soils as old as 2-3 million years would seem highly unlikely without the implementation of DNA damage repair processes *in situ* (Friedman, 1994). Indeed, metabolic activity at temperatures of -7 to -10°C has been reported both in lichens (Friedman, 1982; Kappen 1989) and bacteria (Straka and Stokes, 1960; Larkin and Stokes, 1968).

Microbiological studies of permafrost have predominantly involved Arctic samples. Antarctic permafrost imposes even more stringent requirements for microbial survival: longer time periods (up to ~ 10^7 years); lower temperatures (average -20 to -25°C); and lower water activity (1% soil water content with a computed a_w~ 0.82 to 0.78). These are much more challenging conditions than certain privileged Antarctic microenvironments such as the recently described "oases" harboring microbial consortia under the ice cover of Antarctic lakes (Priscu et al., 1998). The survey and characterization of the Antarctic permafrost bacterial community will be of special interest in the context of a growing body of data on the bacterial communities of several other Antarctic ecosystems (e.g. Horowitz et al., 1972; Franzmann et al., 1990; Delille, 1992, 1993; Palmisano and Garrison, 1993; Grossmann and Dieckmann, 1994; Helmke and Weyland, 1995; Bowmann et al., 1997) as well as for comparisons with the data available on the microbiota of Arctic permafrost. Comparisons between Arctic and Antarctic permafrost bacteria should help reveal the adaptive strategies used for survival of the resident microbes in these extreme environments.

The existing bacterial permafrost communities are the result of a long-term natural experiment in cell preservation under specialized, stringent conditions. A pivotal condition,

time, cannot be readily simulated. Nonetheless, ***valuable information on microbial life's potential for long-term preservation and persistence can be obtained from the characterization of these communities in situ, and from the study of the adaptive physiology of representative community members.***

2. Bacterial cold shock and cold acclimation

When bacteria are exposed to sudden temperature downshifts, two major responses are commonly observed, changes in membrane lipid composition and production of novel proteins Jones et al., 1987; Gounot, 1991). Increases in lipid unsaturation and decreases in the length of the fatty acyl chains result in maintaining the membrane in a fluid state (Russell, 1990). Although bacteria that can grow at low temperature appear to have enhanced potential to maintain membrane lipid fluidity (Gounot, 1991), the molecular and physiologic mechanisms underlying the differences in growth temperature minima among microbes remain largely unclear. In discussions of bacterial low temperature physiology, the terms cold shock and cold acclimation refer to the transitory period following sudden temperature downshift and to steady-state growth at low temperature, respectively.

a. Studies on *Escherichia coli*

When *E. coli* cells are subjected to a rapid downshift in temperature from about 37 to 15°C, the synthesis of most proteins stop and growth ceases. A number of "cold shock" proteins, however, are synthesized and gradually (within a few hours) the bacteria resume growth. Significantly, many of the cold shock proteins have roles in gene expression including protein synthesis. The four most abundant cold shock proteins, CspA, CspB, CspG and CspI (Goldstein et al., 1990; Lee et al., 1994; Nakashima et al., 1996; Wang et al., 1999), are members of a gene family comprising nine closely related homologs. Remarkably, CspA constitutes more than 10% of protein synthesized within an hour of cells being shifted to low temperature (Goldstein et al., 1990). Intriguingly, CspA, CspB and CspG are synthesized in large amounts upon cold shock under conditions that normally inhibit protein synthesis, such as amino acid starvation and treatment by inhibitory antibiotics (Etchegaray and Inoye, 1999). Each of the CspA homologs, referred to as "RNA chaperones," bind to RNA without apparent sequence specificity. Recent evidence indicates that CspA and its homologs act as transcription antiterminators. Moreover, they appear to act to induce cold-regulated expression of four other cold-shock proteins, RbfA, IF2, NusA and PNP. Both RbfA and IF2 are involved in the initiation of translation; NusA associates with RNA core polymerase and is involved in both transcriptional termination and antitermination; and PNP (polynucleotide phosphorylase) is a 3' to 5' exonuclease that interacts with Ribonulease E to degrade mRNA.

Other cold shock proteins that have roles in *E. coli* acclimating to low temperature include CsdA, H-NS, trigger factor and Hsc66. CsdA, which is associated with 70S ribosomes at 15°C, has helix-destabilizing activity which is proposed to be important in removing secondary structures in mRNAs at low temperature. While a *csdA* knockout mutant shows no difference in growth from its parent strain at 37°C, it has an increased generation time of 16 hours at 15°C (compared to an 8 hour generation time of the wild type strain). H-NS is an abundant DNA-binding protein that affects DNA supercoiling and impacts the expression of several unrelated

genes, including the *rrn* operons (reviewed in Atlung and Ingmer, 1997). Binding of H-NS to the *rrn* promoter region specifically counteracts activation by FIS protein (Afflerbach et al., 1998). In an *hns* insertional mutant expressing no detectable H-NS protein, growth was significantly inhibited at 12°C and 25°C but only slightly affected at 37°C (Dersch et al., 1994). Finally, trigger factor and Hsc66 appear to be molecular chaperones with roles in acclimation to low temperature. Trigger factor, which has peptidyl prolyl isomerase activity, is thought to promote folding of nascent polypeptides (Kandror and Goldberg, 1997). Following incubation for one week at 4°C, a non-permissive growth temperature for *E. coli*, overexpression of trigger factor increased survival from 15% to 40%. Hsc66 has 42% similarity to the heat shock protein DnaK and can prevent aggregation of proteins *in vitro* (Silberg et al., 1998). An insertion mutation in the gene encoding Hsc66, *hscA*, was not greatly altered in growth following a cold shock although the expression of 5 proteins was altered at low temperature. The characterization of trigger factor and Hsc66 as cold shock proteins suggests that multiple steps in protein folding are negatively affected by low temperature in *E. coli*.

b. Cold shock proteins in other bacteria

The cold shock response has been investigated in several gram-negative and gram-positive bacteria. The most highly conserved feature of the response is the expression of CspA homologs, having been documented in hyperthermophiles, thermophiles, mesophiles and psychrotrophs. (We recognize that there is difference of opinion in the microbiological community concerning the terms most suitable to describe microbes in regard to their ability to grow at low temperature. In our proposal, we use the term **psychrophiles** to refer to microbes that grow at 4°C or lower, but not above 15°C; **psychrotrophs** are microbes which grow at 4°C or lower, but which have optimal temperature of growth above 15°C; and **mesophiles** are microbes which do not grow at 4°C or lower, and have optimal temperature for growth at 35-38°C.) In several species, multigene families have been identified that contain from three to nine CspA homologs. In each case, at least one of these proteins is increased in expression upon temperature downshift (Graumann et al., 1996; Mayr et al., 1996; Michel et al., 1997; Wouters et al., 1998). Interestingly, in the case of the psychrotrophic bacterium *Arhrobacter globiformis* SI55, production of the protein is not limited to the cold shock period immediately following the temperature downshift but continues during growth at low temperature (Berger et al., 1997) suggesting that the protein functions in cold acclimation (in addition to its function during cold shock). Investigation of CspA expression in additional psychrotrophs will reveal whether in these bacteria, unlike the situation in mesophiles, the major cold shock proteins continue to be expressed during steady-state growth in the cold.

Cold shock and cold acclimation proteins other than the Csp family have also been studied, albeit to a much lesser extent. In *B. subtilis*, the peptidylprolyl isomerase cyclophilin, like *E. coli* trigger factor (which has peptidylprolyl isomerase activity) is a cold shock protein. In *Anabaena* sp. strain PCC 7120, two RNA helicases, CrhC and CrhB, show increased expression following a temperature downshift of 10°C (Chamot et al., 1999). Both of these proteins belong to the DEAD-box family of RNA helicases as does the *E. coli* cold shock gene, *csdA*. Studies with psychrotrophic and psychrophilic bacteria are beginning to provide insight on the involvement of cold-shock and cold acclimation proteins in the ability of the organisms to grow at low temperatures. Most of these studies have involved the psychrotrophic food-borne

pathogens *Yersinia enterocolitica* and *Listeria monocytogenes*. In *Y. enterocolitica*, polynucleotide phosphorylase (involved in mRNA degradation) is a cold shock protein, as it is in *E. coli*. Moreover, PNP is required for growth at 5°C in *Y. enterocolitica* (Goverde et al., 1998); strains carrying transposon insertions in this gene were unable to form colonies at 5°C yet were indistinguishable from the wild type strain at 30°C. In *L. monocytogenes*, twelve proteins are induced following temperature downshift from 37 to 5°C, and four of these proteins continued to be expressed during cold growth, thus being cold acclimation proteins (Bayles et al., 1996). Transposon mutations in several chromosomal loci render *L. monocytogenes* cold-sensitive (Zheng and Kathariou, 1994). The Kathariou lab found that in one such mutant, the insertion was in a locus highly conserved between *L. monocytogenes* and *B. subtilis* and harbored genes involved in DNA repair and cell division. A gene encoding a polypeptide with a RNA-binding domain typical of PNP was also identified at this locus. Other investigators have identified cold-sensitive mutants of *L. monocytogenes* in which membrane fatty acid composition was markedly altered (Annous et al., 1997).

Although no fatty acid desaturase has been identified as a cold shock protein in *E. coli*, cold induced desaturases have been identified in *B. subtilis* and at least two cyanobacteria. In *B. subtilis*, a *des*-*lacZ* fusion was induced 10 to 15-fold following a temperature shift from 37 to 20°C (Aguilar et al., 1998). Levels of unsaturated fatty acids increased at 20°C, but were not required for viability as a *des* null mutant showed growth similar to that of the parent strain at both 37°C and 20°C. Of the four desaturases identified in the cyanobacterium *Synechocystis* sp. PCC 6803, three are clearly induced following a temperature downshift albeit at different rates (Los et al., 1997). In *Synechococcus* sp. strain PCC 7002, three desaturases, DesA, DesB, and DesC, have been identified (Sakamoto and Bryant, 1997). The expression of all three genes was induced within five minutes of a shift from 38 to 22°C. The three genes were also expressed more highly in cells grown continuously at 22°C. Insertion mutants of *desA* and *desB* were unable to grow at 15°C (Sakamoto et al., 1998).

3. Current understanding of freezing tolerance

Poikilothermic organisms (those having body temperatures that fluctuate with the environment) vary markedly in their ability to cope with freezing temperatures. For instance, plants of tropical origin are typically killed by the slightest freeze, while plants from temperate and polar environments can survive freezing at high subzero temperatures (see Thomashow, 1999). Moreover, most temperate and polar plants increase in freezing tolerance in response to low nonfreezing temperatures, a process referred to as "cold acclimation" or "cold hardening." Nonacclimated rye plants are killed upon freezing at about -5°C, but after cold acclimation, can survive freezing down to about -30°C. The ability to cold harden is not limited to plants, but includes other organisms including certain bacteria (Pannoff et al., 1998; Thammavongs et al., 1996). A fundamental goal in low temperature biology research is to determine the mechanisms of freezing tolerance and to identify the genes and proteins that enable poikilothermic organisms to contend with freezing temperatures.

a. Bacterial tolerance to freezing and thawing

Bacteria differ greatly in their ability to tolerate freezing and thawing, for reasons that are poorly understood. Exposure of mesophilic bacteria to low temperatures has been shown to enhance their ability to withstand repeated freeze-thawing ("cryotolerance") (Panoff et al., 1995; Thamavongs et al., 1996). In addition, exposure of lactic acid bacteria to osmotic stress has been recently shown to confer tolerance to freeze-thawing (Panoff et al., 2000). The molecular basis of cryotolerance, conferred either by cold pre-exposure or by osmotic stress, remains to be determined. In contrast to the situation in plants and animals, antifreeze proteins have not been frequently identified in prokaryotes. Bacteria have unique adaptations to low temperature and may have unique and currently unknown mechanisms not only to withstand prolonged exposure to freezing, but also to actually grow at sub-zero temperatures. The remarkable ability of several bacterial species to grow at temperatures as low as -8°C has been described as early as the 1930s (Haines, 1931; Tschistajakow and Botescharowa, 1938a,b).

In *Bacillus subtilis*, where synthesis of three Csp proteins is induced upon cold shock, a minimum of one *csp* gene is required for viability at all temperatures (Graumann et al., 1997). Interestingly, the CspB protein of this organism has been shown to be required for viability of the cells following freezing and thawing (Willimsky et al., 1992), but the role of CspB in freezing tolerance is unknown.

b. Freezing injury and tolerance mechanisms: insights from studies on plants

Many of the basic insights into freezing injury have come from studies with plants (Hughs and Dunn, 1996; Thomashow, 1999). These investigations have revealed that the membrane systems of the cell are the primary site of freezing injury and that freeze-induced membrane damage results primarily from the severe dehydration associated with freezing (Steponkus et al., 1993). As temperatures drop below zero, ice formation is generally initiated in the extracellular spaces due largely to the extracellular fluid having a higher freezing point (lower solute concentration) than the intracellular fluid. Because the chemical potential of ice is less than that of liquid water at a given temperature, the formation of extracellular ice results in a drop in water potential outside the cell. Consequently, there is movement of water down the chemical potential gradient from inside the cell to the extracellular spaces where it freezes. This process continues until an equilibrium in chemical potential is achieved. If the freezing temperature is –10°C, the unfrozen liquid will have an osmolarity in excess of 5 and typically, greater than 90% of the osmotically active water will have moved out of the cell. Formation of intracellular ice is generally thought to be a fatal event.

Although freezing injury is thought to result primarily from membrane lesions caused by cellular dehydration, additional factors may also contribute to freezing-induced cellular damage. There is evidence that freeze-induced production of reactive oxygen species contributes to membrane damage (McKersie and Bowley, 1997) and that extracellular ice can form adhesions with cell walls and membranes and cause cell rupture (Olein and Smith, 1977). In addition, there is evidence that protein denaturation occurs in plants at low temperature (Guy et al., 1998), which could potentially result in cellular damage.

Most of what is know about freezing tolerance mechanisms has come from the study of the cold acclimation response. These studies have shown that key function of cold acclimation is

to stabilize membranes against freezing-injury. Multiple mechanisms appear to be involved in this stabilization. The best documented are changes in lipid composition (Uemura and Steponkus, 1997). However, the accumulation of "compatible solutes," such as proline and simple sugars, that typically occurs with cold acclimation is likely to contribute to the stabilization of membranes as these molecules can protect membranes (as well as proteins) against freeze-induced damage *in vitro* (Strauss and Hauser, 1986; Anchordoguy et al., 1987) and *in vivo* (Nanjo et al., 1999). There is also emerging evidence that certain cold-inducible hydrophilic novel and LEA (late embryogenesis abundant) polypeptides also participate in the stabilization of membranes against freeze-induced injury.

Role of cold-regulated genes encoding unusually hydrophilic polypeptides: COR and LEA proteins. Cold acclimation in plants and other organisms is associated with changes in gene expression (see Thomashow, 1998, 1999). In plants, where the phenomenon has been best studied, it has been shown that the most highly expressed cold-regulated genes encode either newly discovered proteins, such as the Arabidopsis COR6.6, COR15a and COR78 polypeptides, or homologs of LEA proteins, such as Arabidopsis COR47 (Arabidopsis is a small flowering plant that has become an important "model" organism to study basic plant processes) (see Hughs and Dunn, 1996; Thomashow, 1999). Moreover, recent evidence from the Thomashow and Steponkus laboratories (Artus et al., 1996; Steponkus et al., 1998) indicates that at least one of these genes, *COR15a*, encodes a cryoprotective polypeptide.

COR15a encodes a 15 kDa polypeptide that is targeted to the chloroplasts. Upon import into the organelle, COR15a is processed to a 9.4 kDa polypeptide designated COR15am. Constitutive expression of *COR15a* gene in nonacclimated transgenic Arabidopsis plants increases the freezing tolerance of chloroplasts by 1 to 2°C (Artus et al., 1996). This results from the COR15am polypeptide decreasing the propensity of membranes to enter into the hexagonal II phase, a deleterious non-bilayer membrane structure that can form upon freezing-induced dehydration (Steponkus et al., 1998). Moreover, this effect appears to be due to the COR15am polypeptide altering the intrinsic curvature of the membrane monolayers. Studies in mammalian systems have shown that certain polypeptides that form amphipathic α-helices can stabilize membranes against the formation of hexagonal II phase lipids and that this is accomplished by the polypeptides altering the intrinsic curvature of membrane monolayers. A sensitive test of whether a polypeptide has an effect on monolayer curvature is to determine whether the polypeptide causes a shift in the lamellar-to-hexagonal II phase transition temperature. Indeed, Steponkus et al. (1998) have found that the COR15am polypeptide increases the lamellar-to-hexagonal II phase transition temperature of dioleoylphosphatidylethanolamine and promotes formation of the lamellar phase in a lipid mixture composed of the major lipid species that comprise the chloroplast envelope. Whether this effect is dependent on COR15am forming an amphipathic α-helix is not known, but the COR15am is indeed predicted to form an amphipathic α-helix.

It remains to be established whether other COR and LEA genes encode cryoprotective polypeptides that stabilize membranes. In this regard, however, it is significant to note that most of the COR and LEA proteins have a set of distinctive properties in common: they are unusually hydrophilic; many have been shown to remain soluble upon boiling in dilute aqueous buffer; many have relatively simple amino acid compositions, being composed largely of a few amino

acids; many are composed largely of repeated amino acid sequence motifs; and many are predicted to contain regions capable of forming amphipathic α-helices. For example:

- *Arabidopsis COR15a* (Lin and Thomashow, 1992), as mentioned above, encodes a novel 15 kDa polypeptide that is targeted to the chloroplasts. The mature 9.4 kD polypeptide COR15am (Fig. 1) is highly hydrophilic; remains soluble upon boiling; is rich is alanine, lysine, glutamic acid and aspartic acid (they account for 64% of the amino acid residues); and is composed largely of a 13 amino acid sequence that is repeated (imperfectly) four times. Regions of the polypeptide that include the repeated sequences are predicted to form amphipathic α-helices.

- Alfalfa *cas15* (Monroy et al., 1993) encodes a novel 15 kDa polypeptide CAS15 (Fig. 1) that is highly hydrophilic; is rich is glutamate, glycine, histidine and lysine (they account for 68% of the amino acid residues); and nearly a third of the protein is composed of a 10 amino acid sequence that is repeated (imperfectly) four times. Regions of the polypeptide that include the repeated sequences are predicted to form amphipathic α-helices.

- Wheat *wcs120* (Houde et al., 1992) encodes a 39 kDa polypeptide WCS120 (Fig. 1) that is a member of the LEA II group of polypeptides (Close, 1997). It is highly hydrophilic; remains soluble upon boiling; is rich in glycine, histidine and threonine residues (they account for 54% of the amino acid residues); and is composed largely of a lysine-rich sequence, referred to as a K-segment, that is repeated (imperfectly) six times and a glycine-rich sequence, referred to as a ϕ-segment, that is repeated (imperfectly) 11 times. The K-segments are present in all LEA II proteins and are predicted to form amphipathic α-helices (Close, 1997).

- Barley *HVA1* (Hong et al., 1992) encodes a 22 kDa polypeptide HVA1 (Fig. 1) that is member of the LEA III group of proteins (Dure, 1993). It is highly hydrophilic; rich in alanine, lysine and threonine (they account for 53% of the amino acid residues); and is composed largely of an 11 amino acid sequence that is repeated (imperfectly) nine times. The repeated sequences are present in all LEA III proteins and are predicted to form amphipathic α-helices (Dure, 1993).

The intriguing possibility raised is that the similar biochemical properties shared by the novel hydrophilic and LEA polypeptides may reflect a common function of the proteins, namely to stabilize membranes against freezing damage. In this regard, it will be interesting to determine whether any of these proteins effect on the intrinsic curvature of membranes and if so, whether this is due to the regions of the proteins predicted to form amphipathic α-helices.

Cold-regulated genes impart both freezing and dehydration tolerance. As discussed above, freezing tolerance must include tolerance to severe dehydration stress. Thus, one might expect that some of the genes that are important for freezing tolerance would also be important in dehydration tolerance. Indeed, many of the COR, LEA and other cold-responsive genes discovered in plants are also induced upon dehydration stress. Moreover, recent results indicate that the cold-responsive genes impart both freezing and drought tolerance. In particular, the Thomashow (Stockinger et al., 1996; Jaglo-Ottosen et al., 1998; Gilmour et al., 1998) and Shinozaki (Liu et al., 1998; Kasuga et al., 1999) laboratories have identified a family of

Arabidopsis COR15am (novel hydrophilic)

AKGDGNILDDLNEATK **KASDFVTDKTKEA** LADGE
KAKDYVVEKNSET ADTLGKEAE
KAAAYVEEKGKEA AN
KAAEFAEGKAGEA KDATK

Alfalfa CAS15 (novel hydrophilic)

MAGIMNKIGDALHGGGDKKEGEH **KGEQHGHVGG** EHHGEY
KGEQHGFVGG HAGDH
KGEQHGFVGG HGGDY
KGEQHGFGHG DHKEGYHGEEHKEGFADKIKDKIHGEGADGEK
KKKKEKKKHGEGHEHGHDSSSSDSD

Wheat WCS120 (LEA II)

MENQAHIA **GEKKGIMEKIKEKLPGGHGDHKE** TAGTHGHPGTATHGAPA **TGGAYGQQGHAGTT** GTGLHGAHA
GEKKGVMENIKDKLPGGHQDHQQ **TGGTYGQQGHTGTA** THGTPA
TGGTYGQQGHTGTA THGTPA
TGGTYGEQGHTGVT GTGTHGT
GEKKGVMENIKEKLPGGHGDHQQ **TGGTYGQQGHTGTA** THGTPA
GGGTYEQHGHTGMT GTGTHGT
GEKKGVMENIKDKLPGGHGDHQQ **TGGTYGQQGHTGTA** TQGTPA
GGGTYEQHGHTGMT GAGTHST
GEKKGVMENIKEKLPGGHSDHQQ **TGGAYGQQGHTGTR** HMAPL
PAGTYGQHGHAGVI GTETHGTTA
TGGTHGQHGHTGTT GTGTHGSDGI
GEKKSLMDKIKDKLPGQH

Barley HVA1 (LEA III)

MASNQNQGSYHAGETKARTEEKTGQM **MGATKQKAGQT**
TEATKQKAGET
AEATKQKTGET
AEAAKQKAAEA KDKTAQT
AQAAKDKTYET
AQAAKERAAQG KDQTGSALGEK
TEAAKQKAAET
TEAAKQKAAEA
TEAAKQKASDT AQYTKESAVAGKDKTGSVLQQAGE
TVVNAVVGAKDAVANTLGMGGDNT
SATKDATTGATVKDTTTTTRNH

Figure 1. Examples of novel hydrophilic and LEA proteins encoded by genes induced during cold acclimation in higher plants. The amino acid sequences are presented with the repeated motifs aligned (in bold) to highlight this attribute of each protein. Additional details about the proteins are presented in the text.

transcriptional activator proteins, known as CBF1, CBF2 and CBF3 (or DREB1b, DREB1c and DREB1a, respectively), that control the expression of a battery of cold-responsive genes in Arabidopsis including the *COR6.6*, *COR15a*, *COR47* and *COR78* genes mentioned above. Constitutive overexpression of the *CBF* genes in transgenic Arabidopsis plants induces the expression of the *COR* and other cold-responsive genes without a low temperature stimulus and results in increased freezing tolerance at the whole plant level (in contrast, expression of *COR15a* alone has no apparent effect on the freezing survival of whole plants) (Jaglo-Ottosen et al., 1998; Liu et al., 1998). In addition, it results in increased tolerance to stresses that involve dehydration including drought and high salinity (Kasuga et al., 1999).

c. Role of antifreeze proteins in freezing tolerance

Antifreeze proteins (AFPs) have been identified in distantly related organisms (see Ewart et al., 1999). The best studied are from fish and insects, but recently, they have also been described in plants and bacteria. A hallmark characteristic of these proteins is "thermal hysteresis" activity: the proteins decrease the temperature at which ice is formed, but do not affect the melting point of the solution. This effect results from the AFPs binding to the surface of ice nuclei and inhibiting ice crystal growth. In addition, AFPs affect the shape of the ice crystals that form when temperatures drop below the freezing point of the solution and are potent inhibitors of ice recrystallization (they inhibit the coalescing of small ice crystals into large ice crystals).

Significantly, AFPs are highly diverse in amino acid sequence and structure (Ewart et al., 1999). In fish, there are four classes of AFPs: type I are alanine-rich α-helixes (M*r* ca. 3300-5000): type II are cystine-rich globular proteins (M*r* ca. 14000); type III proteins form a compact β-sheet structure (M*r* ca. 6500); and type IV proteins are though to form a four-helix bundle. Two AFPs from insects, spruce budworm and mealworm beetle, are rich in threonine and cytosine and have been shown to form β-helixes. In winter rye plants, six AFPs ranging in molecular mass from 16 to 35 kD have been shown to accumulate in the extracellular fluids of cold acclimated plants (Antikainen and Griffith, 1997). Surprisingly, these AFPs are related to pathogenesis-related (PR) proteins: two of the AFPs are endochitinase-like proteins; two are β-1,3-glucanase-like proteins; and two are thaumatin-like proteins (Antikainen and Griffith, 1997). One of the rye AFPs has been purified and shown to have both endochitinase and antifreeze activity (the protein alters the shape of ice crystals and has a low level of thermal hysteresis activity). In contrast, a purified endochitinase from tobacco, a freezing sensitive plant, was devoid of antifreeze activity. In addition, endochitinase-, glucanase- and thaumatin-like proteins are present in cell extracts of nonacclimated rye plants, but no antifreeze activity can be detected. Thus, antifreeze activity is not an inherent property of these proteins. Taken together, these results indicate that antifreeze activity has evolved independently multiple times.

It is important to note that the different activities of the APFs do not appear to strictly correlate. An AFP from perennial ryegrass has been described that has very low thermal hysteresis activity, but is a very potent inhibitor of ice recrystallization (Sidebottom et al., 2000). In addition, the functions of the AFPs appear to differ among organisms (Antikainen and Griffith, 1997; Ewart et al., 1999). In some cases, the AFPs have a role in frost avoidance. In

polar fish living in waters that are about two degrees below freezing, the AFPs, through their thermal hysteresis activity, contribute to keeping the blood serum in the liquid state. However, in organisms that do not avoid freezing, like cold-acclimated rye plants, AFP function is presumably different. In the case of extracellular AFPs, it is possible that the inhibition of ice recrystallization, or the effects that AFPs have on ice crystal shape, might mitigate against physical damage caused by ice and thereby enhance freezing tolerance. Intracellular AFPs, which do not come into direct contact with ice, might have fundamentally different roles. Indeed, it has been shown that certain AFPs (antifreeze gluycoproteins) can inhibit leakage of ions from liposomes during thermotropic phase transitions (Hays et al., 1996). Thus, some AFPs might have important roles in stabilization membranes at low and freezing temperatures. In this regard, it is interesting that certain of the AFPs, like the COR and LEA proteins described above, are hydrophilic; remain soluble upon boiling; and are composed in large part of simple repeated amino acidic sequences.

B. Preliminary Results

1. Sites of isolating permafrost bacteria

Our Arctic study site is the Kolyma-Indigirka Lowland (152-162° E and 68-72° N), which is a tundra ecosystem in Northeastern Siberia adjacent to the East Siberia Sea. Permafrost occurs throughout the area to depths of ~600-800m. These layers were syngenetically frozen (i.e., sedimentation occurred concurrent with freezing) and thus, the age of the permafrost is equal to the age of the sediment. Evidence suggests that buried permafrost in this region has remained continuously frozen since its period of deposition (Gilichinsky et al., 1995). Permafrost cores of Arctic and Antarctic sediments have been collected by slow rotary drilling, so as not to melt them, and without the use of any muds or solutions. The temperature of the extracted core was never higher than -7°C. Surfaces of the extracted core were trimmed away with a sterile knife and the remaining core was immediately placed in pre-sterilized tins, sealed and kept frozen during storage and transport. When ready to use in the laboratory, the frozen core was placed in a sterile hood and fractured with a sterile knife**. *Only internal segments of the core were used for microbiological studies.*** Previously Khlebnikova et al. (1990) had shown that *Serratia marcescens* could not be detected in internal segments of cores that had purposely received surface contamination with this organism in the field.

Our Antarctic permafrost samples were collected in January 1999 during the Beringia expedition to the Antarctica Dry Valleys (Beacon Valley) by the Gilichinsky team. The same coring methods described above were used. The Antarctic sediments are colder (-18 to -25°C), have less organic content (down to 0.1%), and have been frozen longer than those from the Arctic. Our bore hole went through an ash layer that had been dated at 8.1 million years, so the deeper samples must be at least this old.

2. rDNA analysis of prokaryotic permafrost communities

We have sequenced the small subunit (SSU) rRNA genes of dominant bacteria and Archaea clones amplified from Arctic community DNA extracted from permafrost sediments ranging in (frozen) age from 25,000 years to 2-3 million years. Nearly complete SSU rRNA

gene sequences from bacterial clone libraries revealed two major classes of organisms: (i) clones with sequences very similar to *Arthrobacter* and (ii) clones representing a novel gamma Proteobacteria group. The latter bears resemblance to recently reported Fe(II) oxidizing chemolithotrophs that grow at neutral pH (Emerson and Moyer, 1997). Numerous isolates of *Arthrobacter*-like organisms have been obtained from the permafrost, but no isolates of the latter group have yet been recovered. The Archaea clone libraries are even more interesting (Moyer, Tiedje and Petrova, unpublished). Nearly full-length sequences of the majority of the clones reveal a novel cluster within the *Crenarchaeota* which is clearly distinguished from the other *Crenarchaeota* groups, most significantly from the soil and low temperature marine clusters.

Having completed the SSU rDNA analysis of several bore holes and age layers provides us with considerable information for comparison to proposed Antarctic data. This also provides sequence information for design of probes and primers to quantify the major groups, and to aid in cultivation and perhaps isolation of the novel groups.

3. Bacteriologic and molecular characterization of culturable bacteria from Arctic permafrost

Several types of growth media, differing both in nutrient content and in the presence or absence of osmoprotectants, were used to successfully isolate aerobic bacteria, yeasts and fungi from permafrost samples collected in the period 1991-1997. Isolations were all done at 4°C immediately following thawing of the samples, as well as following natural growth of strains in the permafrost soil held at 4°C for 2-18 weeks. Strains from selected samples were also isolated at -2.5 and -4.5°C. Direct bacterial counts in the samples were monitored by staining with epifluorescence using DTAF (Bloem et al., 1995). When surveyed immediately upon thawing, all samples had comparable numbers of total bacterial counts as detected by DTAF staining (average, 10^8 bacteria/g), whereas the number of culturable aerobic bacteria ranged from 0 to 10^2-10^3 CFU/g in samples dating from middle Pleistocene to Pliocene (100,000 years up to 2-3 million years), to ca. 10^5-10^6 in samples dating 25,000-40,000 years. Similarly processed modern tundra soil had 10^5-10^6 CFU/g. Although most isolated cultures were bacterial, several strains of yeasts and fungi were also recovered (Vishnivetskaya et al., 2000).

Over 400 strains, representing as many diverse cultural phenotypes as possible, were isolated and characterized further. With very few exceptions, the isolated bacterial cultures were non-spore forming and grew well in the -2.5 to 20°C range, suggesting that they were psychrotrophic (growing at low temperature but also above 15°C). This was in contrast to the high frequency of psychrophiles (growing at low temperature but not above 15°C) documented in Antarctic sea ice (Bowmann et al., 1997) and in Burton Lake, Antarctica (Franzmann et al., 1990). Growth in the low temperature range, however, was prolific for all strains that we isolated. The bacteria grew readily at -2.5°C, both in liquid and on solid media, and screening of 6 randomly selected strains at -4.5°C revealed that within 7-8 weeks all formed colonies on agar plates. In contrast, an *Arthrobacter* strain isolated from Michigan soil (LTER3) grew at 4°C but was unable to grow below 0°C.

Selected gram-positive and gram-negative strains isolated from different samples, and from different depths of one core sample, were typed by repetitive element Polymerase Chain

Reaction fingerprinting (rep-PCR) (Versalovic et al., 1991). The rep-PCR fingerprints suggested several clusters within both the gram-positive and gram-negative lineages, which were analyzed further by restriction fragment length polymorphisms within PCR-amplified portions of the gene encoding 16S RNA (rDNA). Forty four strains representing distinct molecular fingerprints were characterized by partial 16S rDNA sequencing. These preliminary molecular taxonomy data revealed that the strains were predominantly of 4 groups: Gram-negative strains were in *Psychrobacter* sp. (gamma division of proteobacteria) and in the *Flavobacterium-Bacteroides - Cytophaga* group); Gram-positive bacteria were mostly in the high GC *Arthrobacter* group and in the low GC *Planococcus* group. Gram-positive bacteria tended to predominate (58-85% in the different sediments) and a large proportion were of the high GC groups, especially *Arthrobacter*. About 50% of the strains closely matched (98-100% identity) other lineages in the ribosomal database. Of the 44 strains that were analyzed, 6 were found to have 92-96% identity to their closest relatives in the database, suggesting that they represent novel taxonomic units.

Several themes emerged from the phylogenetic analysis of the bacteria which were isolated from arctic permafrost. It was noted that certain lineages were more frequently retrieved from relatively young (20,000-40,000 yrs) sediments than from the more ancient (200,000 to 2-3 million yrs) permafrost. Thus, *Planococcus* (Fig. 2: Low G+C gram-positives) and *Psychrobacter* (Fig. 2: Proteobacteria) were repeatedly isolated, but only from relatively young sediments. Strains with highest similarity at the level of rDNA sequence (97-99%) to *Arthrobacter psychrolactophilus* were isolated on multiple occasions from different sediments in the relatively young age range, whereas *Arthrobacter* from the most ancient sediments had highest similarity (98-99%) to *A. citreus* and *A. agilis* (Fig. 2: High C+G gram-positives). Certain lineages, such as those with highest similarity (98-99%) to *Rathayibacter, Clavibacter* and *Curtobacterium* (Fig. 2: High G+C gram-positives) and to the gram-negative *Sphingomonas* sp. Ant20 (99-100% identity) (Fig. 2: Proteobacteria) were only isolated from the very ancient (200,000 to 2-3 million yr) sediments. It is worthy of note, although not readily interpretable with currently available data, that the *Rathayibacter-Clavibacter-Curtobacterium* lineages (Fig. 2: High G+C gram-positives) from the ancient sediments were closest to modern bacteria which are associated with plants and agricultural soils. The overall findings, nonetheless, suggest that certain lineages have adaptive features which allow them to compete effectively and predominate in relatively young sediments that contain high numbers of other culturable bacteria but become eliminated as the permafrost ages, whereas others seem to survive preferentially in the very ancient sediments where most other microbes have been eliminated.

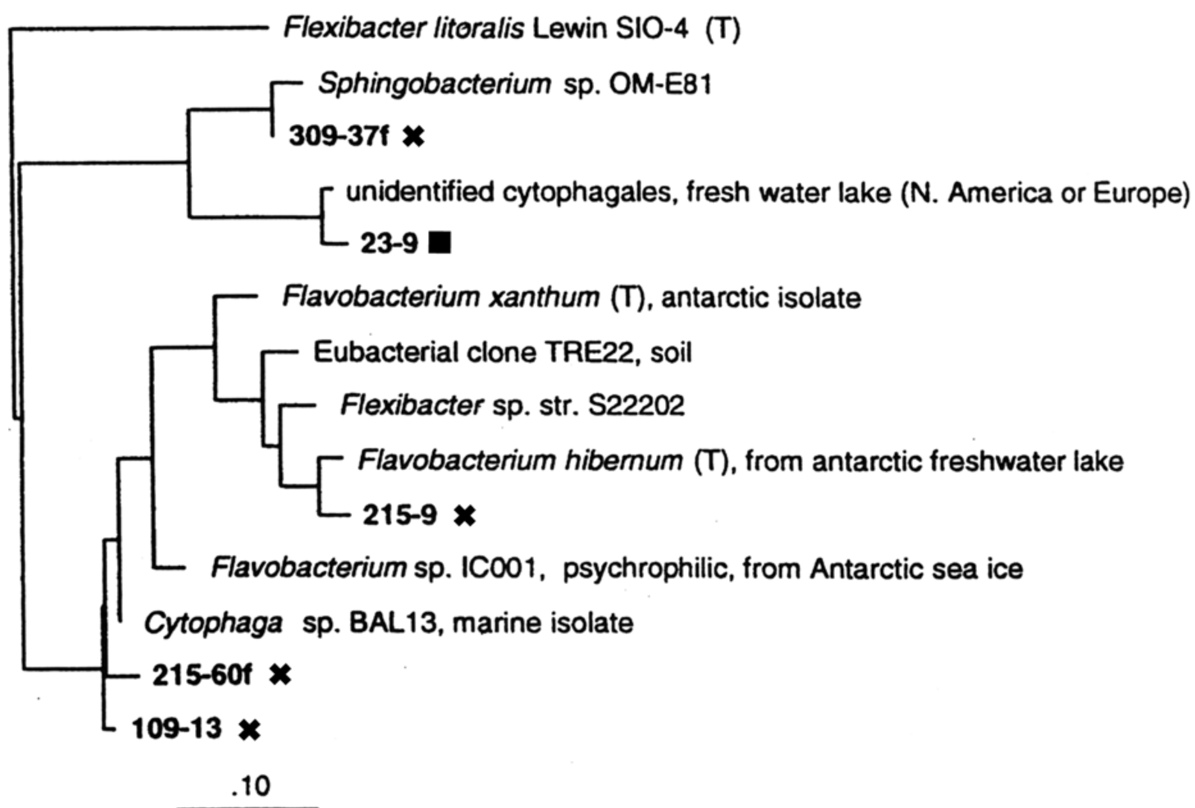
Analysis of the taxonomic data identified one lineage, represented by *Exiguobacterium* strains, which was retrieved from sediments of variable ages (20,000 to 2-3 million yrs) as well as from modern tundra soil (Fig. 2: Low G+C gram-positives). The *Exiguobacterium* lineage appears to represent microbes which can effectively survive under the distinct conditions (including the key impact of time) that characterize permafrost of widely variable ages.

Another interesting finding that emerged from the phylogenetic analysis of the lineages was that 16S rDNA sequences of *Psychrobacter*, *Flavobacterium-Bacteroides-Cytophaga, Sphingomonas* and *Planococcus* strains most closely matched (94-99% identity) with the sequences of strains of the same genera isolated from Antarctic sea ice (Bowmann et al., 1997), Antactic sea brine (Junge et al., 1998), Antarctic soil, and from a freshwater Antarctic lake

Figure 2. Phylogenetic relationships among permafrost isolates.

16S rRNA sequences were compared to public databases (Ribosome Database Project, GenBank) to identify nearest relatives. The phylogenetic trees, generated by maximum likelihood using 337 to 395 homologous nucleotide positions, illustrate permafrost isolates and nearest relatives. Additionally, related type strains (designated "(T)") were included to place relationships within the context of the global bacterial phylogeny. When available, sources of isolation of all relatives are illustrated on the trees. Isolates and clones are color coded based upon the age of their sediments of origin (see code on each tree). Isolates from the current study are indicated by bold type face.

Flavobacteria/Cytophaga



Key:

- **Modern day sample or reference sequence**

✖20-30K years

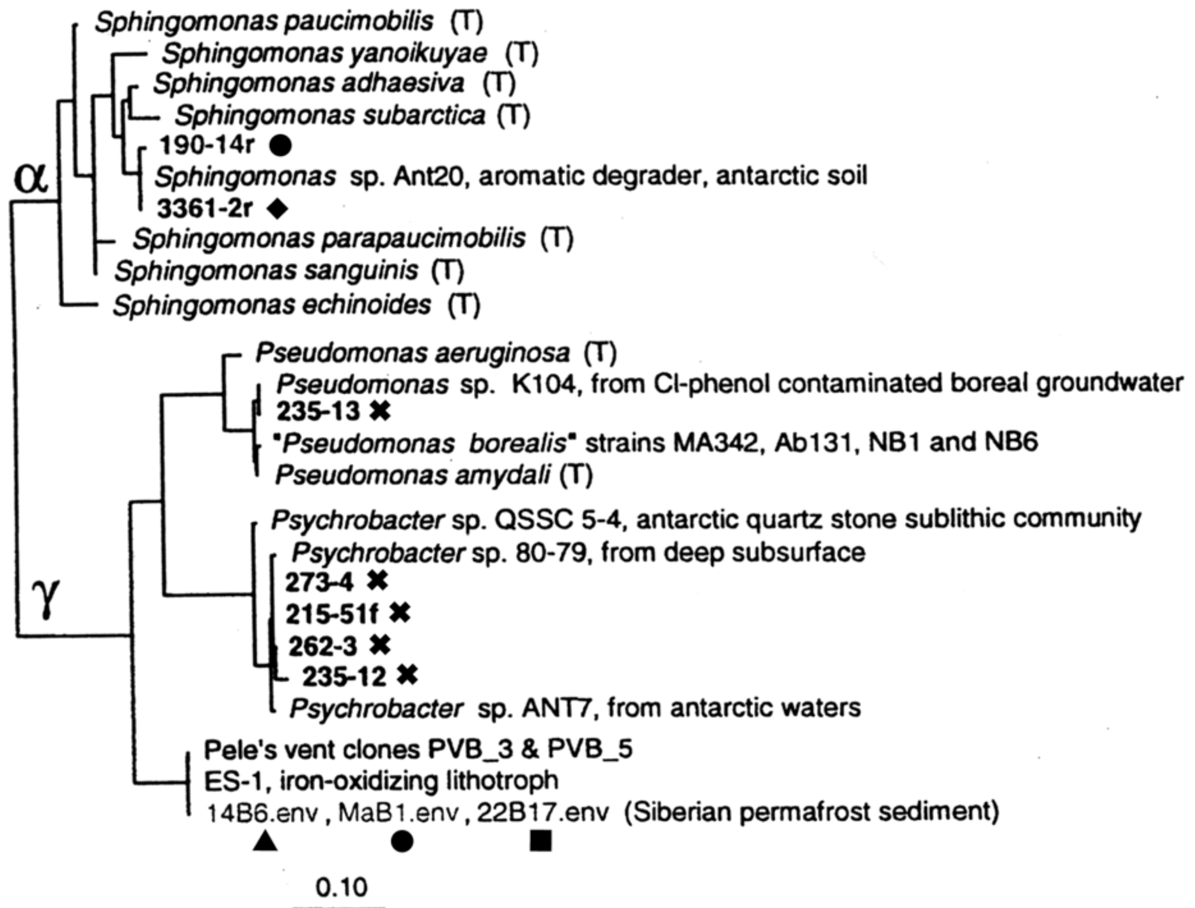
▲60-100K years

●200-600K year s

■0.6-2 million years

◆2-3 million years

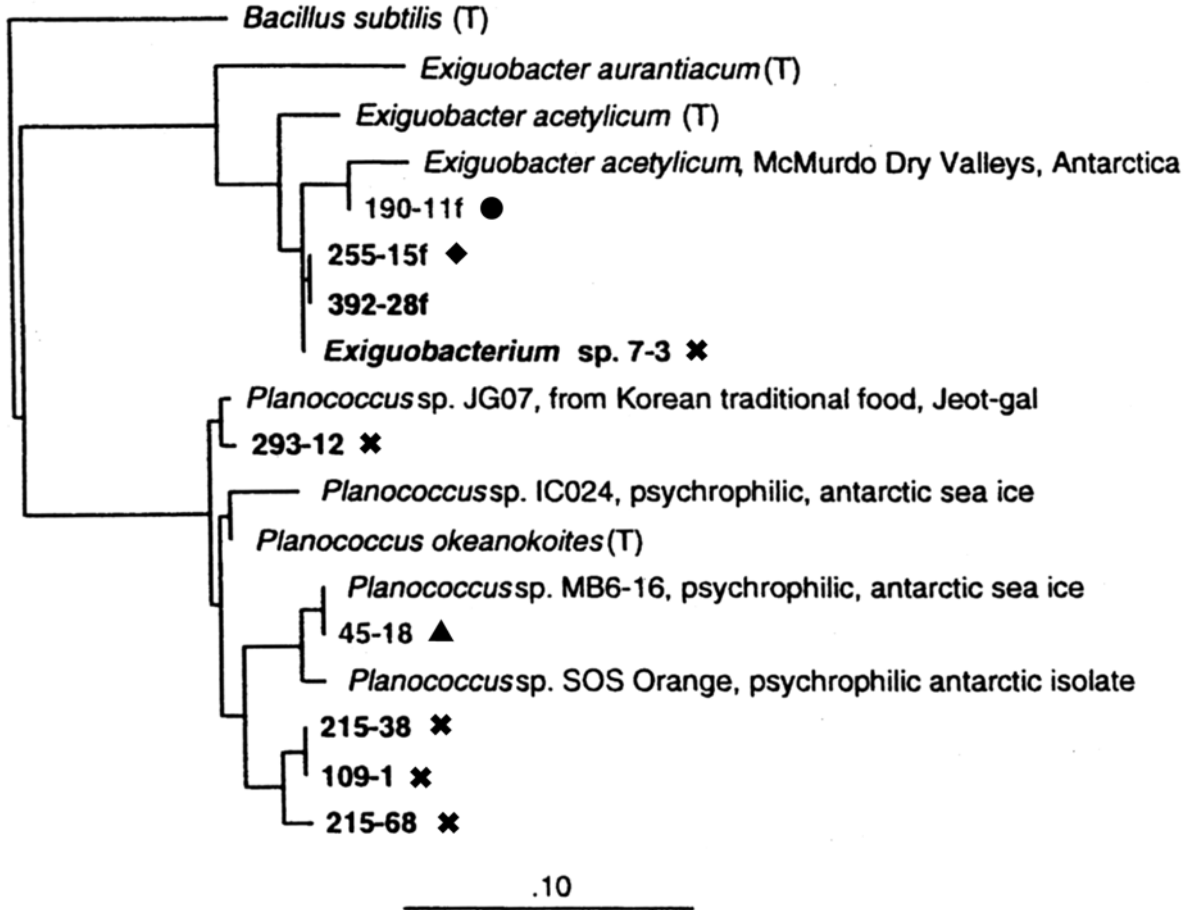
Figure 2. (cont.): Proteobacteria (alpha and gamma subdivisions)



Key:

- **Modern day sample or reference sequence**
- ✖ **20-30K years**
- ▲ **60-100K years**
- ● **200-600K years**
- ■ **0.6-2 million years**
- ◆ **2-3 million years**

Figure 2. (cont.): Low G+C gram-positives



Key:

- **Modern day sample or reference sequence**

✖ **20-30K years**

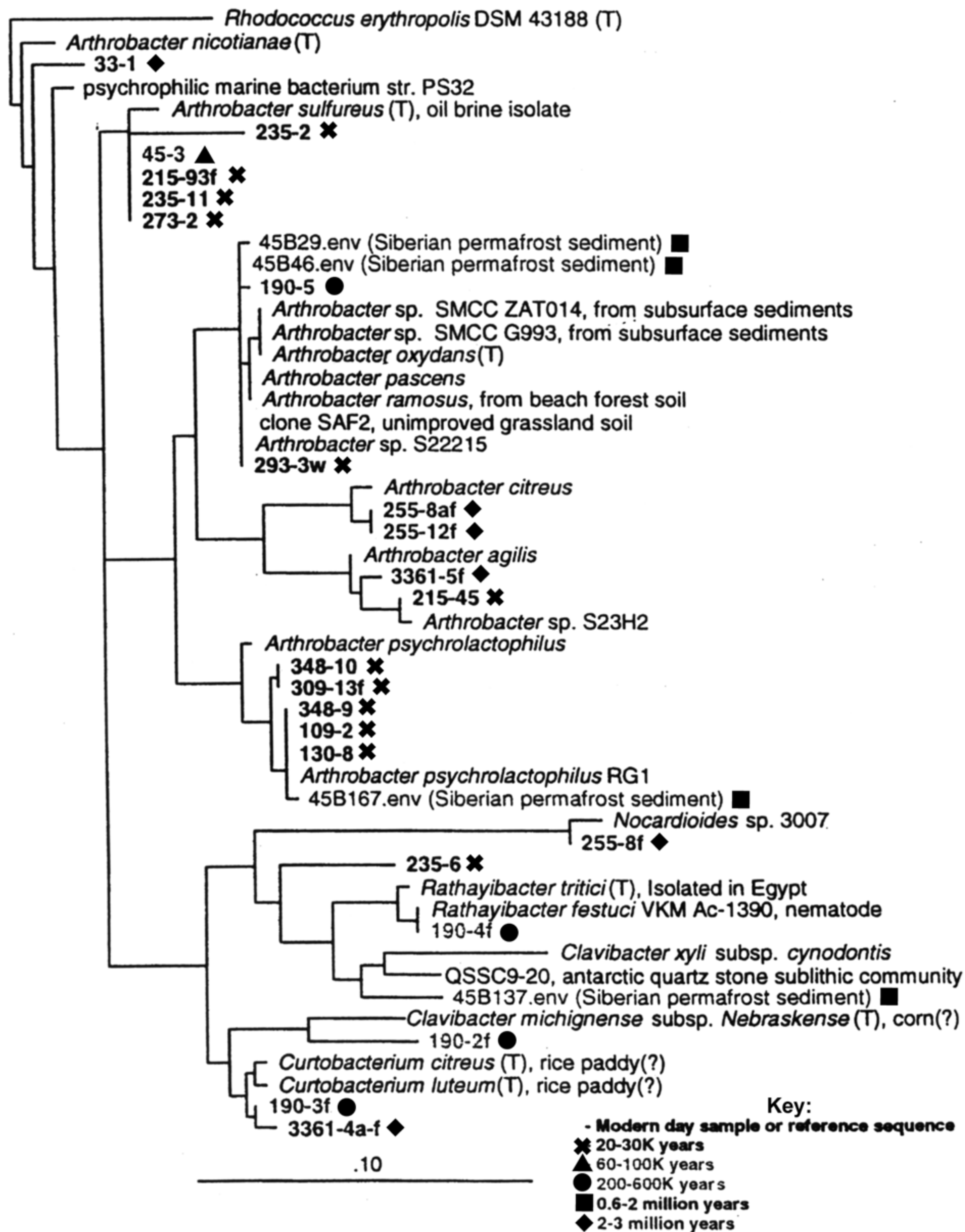
▲ 60-100K years

● 200-600K years

■ **0.6-2 million years**

◆ 2-3 million years

Figure 2. (cont.): High G+C Gram-positives



(McCammon et al., 1998) (Fig. 2). In addition, the closest matches for several of the *Arthrobacter* strains were *Arthrobacter* from Antarctic sea ice brine (Junge et al., 1998) and from deep subsurface sediments (Boivin-Jahns et al., 1995).

These findings suggest that there may be an ecological and/or adaptive link between bacteria from the Arctic permafrost and those from Antarctic habitats. The fact that the Arctic permafrost bacteria were proficient in low temperature growth, but were not limited to it, may suggest that they represent globally distributed, ancient lineages of highly cold-tolerant microorganisms that specialized in the Antarctic sea ice and lakes to become more cold-adapted, and more limited in their temperature range. This has been suggested by other studies of the comparative distribution of psychrophilic bacteria and their non-psychrophilic relatives (Franzmann, 1996).

4. Responses of permafrost bacteria to low temperature

a. Compatible solutes and lipids

We have used HPLC to show that several of our permafrost strains produce compatible solutes when they are grown at low temperature (4°C) compared to growth at 20°C (McGrath, unpublished). In three of the four strains tested estimated intracellular concentrations of solutes increased significantly. In one case, glutamate + proline increased from 200 mM (at 20°C) to 1200 mM (at 4°C). However, in another, the metabolite pool (per unit quantity of cells) decreased in total metabolites when cells were grown at 4°C.

In agreement with findings from numerous other bacterial systems (Russell, 1990), examination of the cell lipids of seven permafrost strains using the MIDI system and FAME analysis showed that the lipids were altered when the permafrost microbes were grown at low temperature (4°C) compared to those grown at 20°C (Wagener et al., 1993). The general trend was increased fatty acid unsaturation when grown at low temperature, and no significant change in average fatty acid chain length.

b. Freezing tolerance

Freeze tolerance of bacteria has mostly been studied using rapid, serial freeze-thawing treatments (Pannoff et al., 1998; Thammavongs et al., 1996). The ecology of the permafrost, however, suggested that it would be more relevant to determine tolerance of the bacteria to prolonged exposure to subzero temperatures, or to intermittent thawing followed by prolonged freeze periods. Experiments with several permafrost strains showed marked impact of low temperature growth on freeze tolerance. Bacteria grown at 20°C and 4°C were frozen at -20°C and survival was monitored over 5 months. It was found that during this time survival of 20°C-grown bacteria decreased markedly (in strain ED 23-9 survival dropped from 0.8×10^8 to $<10^4$ CFU/ml in a representative experiment), whereas survival of 4°C-grown bacteria was much less affected (a drop from 0.9×10^8 to 0.5×10^8). Interestingly, growth of the bacteria at 10°C did not confer any protection to prolonged freezing, and growth at temperatures below 4°C (2.5, 0 and -2.5°C) was as protective as growth at 4°C. We found that 4°C growth also conferred significant protection to cells subjected to intermittent thawing followed by prolonged (1 week)

periods of freezing. In contrast, 4°C or 2.5°C growth failed to confer marked protection to long-term freezing and freeze-thawing in a psychrotrophic strain from Michigan soil, LTER3 or another psychrotrophic bacterium, *Listeria monocytogenes*.

c. Detection of respiratory activity in permafrost bacteria

The dye CTC (5-cyano-2,3-ditolyl tetrazolium chloride) can be used as a fluorescent redox probe for the direct visualization of actively respiring bacteria. CTC is a nearly colorless artificial electron acceptor that is readily reduced to insoluble, highly fluorescent, and intracellularly accumulated CTC-formazan by bacterial respiration. Since the fluorescence emission of CTC-formazan is primarily in the red region, actively respiring bacteria can be distinguished from nonrespiring bacteria and nonbiological material that typically emits in the blue to blue-green regions (Rodriguez *et al*., 1992).

We have assessed the use of CTC to determine metabolic activity of bacterial cells from permafrost soils using both laboratory strains and the permafrost samples themselves. We have found that that the detection of metabolic activity using CTC reduction is possible under all of the conditions of interest to us. These include: both gram- positive and gram-negative permafrost strains grown at both low and high temperatures, in the frozen state, with and without high salt concentrations, in cells grown in liquid or on solid medium, in assays done at low or high temperatures, within soil slurries with and without the addition of an additional carbon source. In addition, activity was also detected in the permafrost samples (which were thawed and kept at 4°C during the assay) by direct epifluorescence. Control experiments show that only intact, living, and respiring cells seem to exhibit reducing activity on the CTC dye. Sensitivity and kinetic details remain to be established. We believe that it will be possible to use laser confocal microscopy to determine metabolic activity *in situ* in permafrost soil samples in their original state, i.e. to study a freshly thawed permafrost sample, rather than samples incubated as a soil slurry in a buffered aqueous solution with CTC.

5. Development of methods for proteomic analysis of permafrost bacteria

The Lubman laboratory has been developing methodologies for rapid screening of the protein content of bacterial cells using fast nonporous reversed-phase (NP-RP) HPLC separation of proteins (Kalghatgi and Horvath, 1990; Jenke, 1996; Barder et al., 1997). As each protein fraction elutes, it is collected in the liquid state from the HPLC separation for analysis by mass spectrometry (MS). The use of NP C18 coated silica based packing materials, as compared to the conventional porous materials, is the key issue in this work. The separation of complex protein mixtures from cells can be accomplished in 1/3-1/4× the time required for porous packed column separations with much enhanced protein recovery and resolution. The use of NP packing materials prevents proteins from sticking inside the pores of porous materials, thus allowing much greater speed of separation, a decrease in chemical noise background, and enhanced protein recovery.

In complex protein mixtures, a single NP-RP-HPLC separation may not be sufficient to resolve all the proteins. In present work, Lubman and coworkers have been using two NP columns on-line to separate out complex mixture of proteins. By using a 4.6 × 33 mm column at

60°C followed by a second shorter column, 4.6 × 14 mm, held at 25°C, over a hundred proteins from a Siberian Permafrost bacterial whole cell lysate could readily be separated. One important advantage of the method is that excellent separations can be obtained using two columns held under different conditions and this can be preformed on-line in 15-30 minutes. A second important advantage is that a substantial amount of material can be loaded on the column and high efficiency in recovery achieved. In recent experiments, it has been shown that over 350 μg can be loaded onto the system and the final recovery after separation by the second column is nearly 90 % for water-soluble proteins under 30 kDa. The recovery of protein will be critical for detailed sequencing experiments. The key here is that although some overloading with loss of resolution may occur on the first column, this is compensated for by separation of the eluting peaks from column 1 to column 2. Even though not a genuine 2-D LC separation, the results are very encouraging for the proposed work. In addition, the HPLC separation is highly reproducible on each different run and we are able to quantitate changes in peak profile for expressed proteins with ± 10 %.

The tandem-column NP-RP-HPLC has been used as a rapid method for separation and isolation of proteins in the liquid phase from bacterial whole cell lysates for further analysis by MALDI-TOF MS and CE-ESI-MS. Bacterial strain ED 7-3, isolated from the Siberian Permafrost, was studied in order to examine cold shock response and resulting induction of cold shock proteins (CSPs) at 4°C and room temperature (25°C). The whole cell lysates are rapidly profiled and separated using a tandem-column NP-RP-HPLC method where distinct protein profiles were obtained from each culture. Proteins that were differentially expressed at 4°C were collected and sized for molecular weight by MALDI-TOF MS. The proteins were digested with trypsin to generate peptide maps for identification by MALDI-TOF MS analysis and further confirmed by CE-ESI-MS using database searching procedures (Mortz et al., 1994; Ducret et al., 1998; Chen et al., 1999). As a result, several CSPs and other cold-shock related proteins could be identified.

As an example, Figures 3a and 3b show the protein profiles of the whole cell lysates of ED 7-3 using the tandem-column NP-RP-HPLC method. Employing a relatively short gradient program (about 25 minutes), distinct profiles were obtained for each growth temperature. For the cold (4°C) culture, 7-3-4 (Fig. 3a), the majority of proteins elute between 3 and 17 minutes, while for warm (25°C) culture, 7-3-25 (Fig. 3b), most of the proteins elute after 15 minutes into the gradient. The profiles show that in the culture grown at 4°C, a large number of new proteins were expressed relative to the 25° culture. Some of these protein peaks have been identified (Table 1) including several CSPs such as CSP C, CSI14B, CSI15, and CSI5. Other proteins that are highly expressed belong to the general stress family (GSP170), translational and transcriptional family (IF-1, SENS, and ABRB), degradation regulation enzyme family (DEGQ), and the ribosomal family (S6). Trisephosphate isomerase (TIM) and RNA polymerase sigma-B factor (Sigma-37) are also highly expressed. This RNA-polymerase is also known as general stress protein 84 or GSP84. Many of these proteins are housekeeping proteins and have been previously identified in *B. subtilis* and other bacteria (Jones and Inouye, 1994; Bayles et al., 1996; Berger et al., 1996; Bae et al., 1997; Panoff et al., 1997). They are essential for the proper function of the bacterial cell during the cold shock process (Graumann et al., 1996, 1997, 1999; Kaan et al., 1999).

Figure 3a. NP-RP-HPLC protein profile of the whole cell lysate of 7-3-4 culture (4°C).

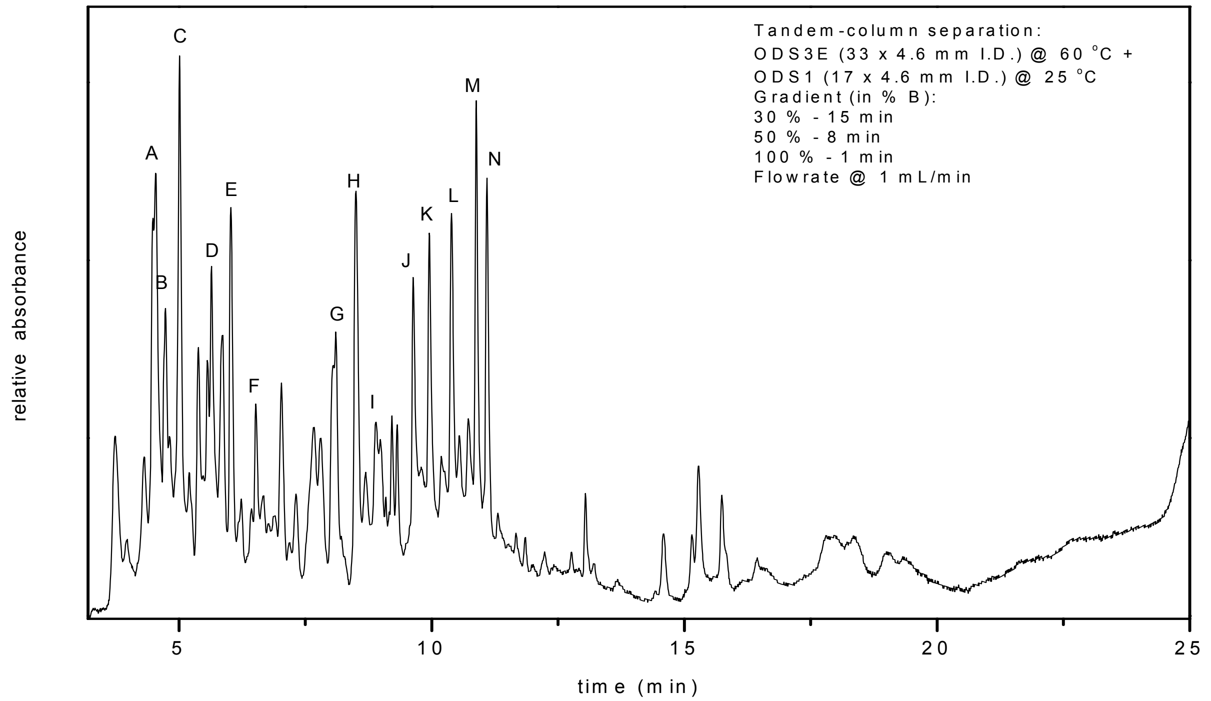
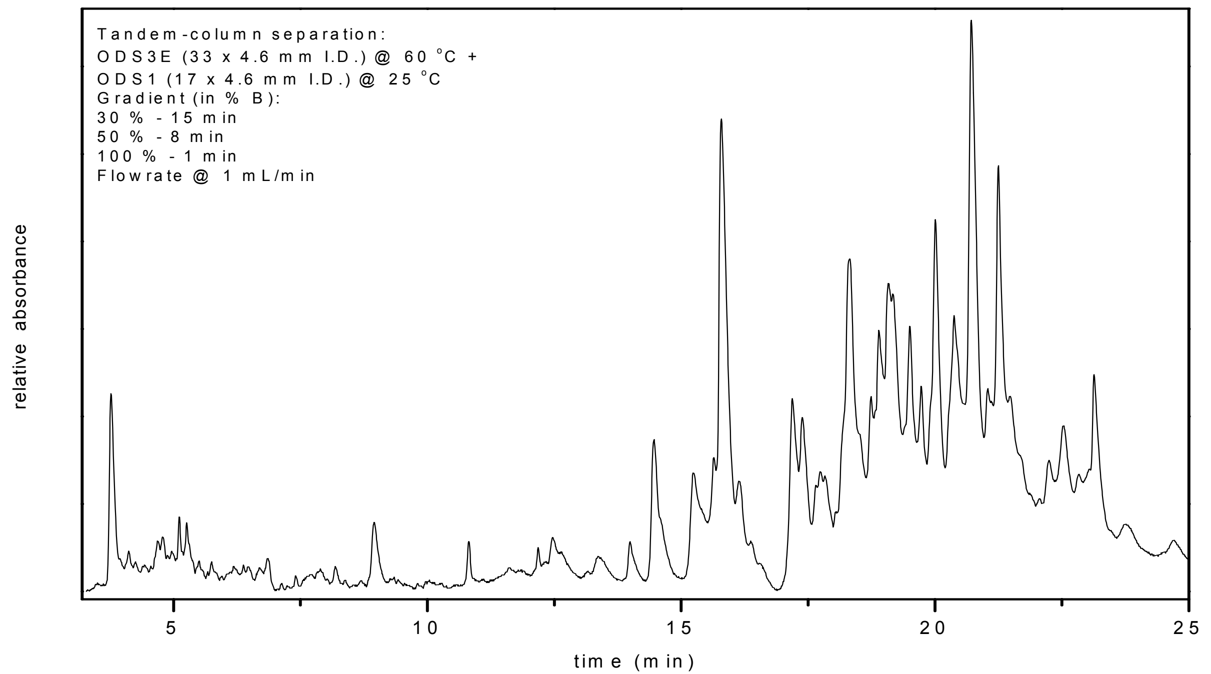


Figure 3b. NP-RP-HPLC protein profile of the whole cell lysate of 7-3-25 culture (25°C).



This tandem-column NP-RP-HPLC method has several advantages for separation of proteins from whole cell lysates in these studies compared to more traditional 1-D or 2-D polyacrylamide gel electrophoresis methods. A major advantage of the NP-RP-HPLC method is the separation and isolation of proteins in the liquid phase. The purified proteins in the liquid phase are readily amenable to analysis by MALDI-TOF MS or CE-ESI-MS as compared to gel methods, which require time consuming extraction, destaining, and purification procedures for analysis of the proteins embedded in the gels. In addition, even after these purification procedures the percent coverage of the protein digests of proteins isolated by NP-RP-HPLC is generally much improved compared to that observed for proteins excised from gels. This will be important for identification and detailed sequencing by MS analysis.

Table 1. Proteins induced during cold shock response in ED 7-3 at 4 ^{o}C . Tryptic peptide maps used for MS-Fit database-search were obtained from both CE-ESI-MS and MALDI-TOF MS

Peak	Protein Name	**kDa / *pI***	**AC**
A	[still to be determined]	–	–
B	[still to be determined]	–	–
C	DNA-binding protein II (HB)(HU)	9884.3 / 8.96	P08821
C	Stage V Sporulation protein M	3017.8 / 10.57	P37817
D	Cold shock protein CSI4B (9 KD Cold shock protein)	1924.3 / 4.52	P81094
D	General stress protein 170 (GSP170)	2158.4 / 8.80	P80242
E	Cold-shock induced protein 15 (CSI15)	2376.8 / 10.83	P54616
E	Transcriptional regulatory protein SENS	7912.5 / 10.00	P21344
F	Translation initiation factor IF-1	8213.6 / 6.83	P20458
G	[still to be determined]	–	–
H	Cold shock proteinC (CSPC)	7255.0 / 4.72	P39158
H	Cold shock protein CSI5 (11 KD Cold shock protein)	1359.7 / 11.17	P81095
I	Transcription state regulatory protein ABRB	10772.7 / 6.29	P08874
J	Degradation enzyme regulation protein DEGQ	5546.5 / 6.16	Q99039
K	30S Ribosomal protein S6 (BS9)	11124.6 / 5.2	P21468
L	[still to be determined] (General stress protein 84) (GSP84)	–	–
M	Trisephosphate isomerase (TIM)	27029.8 / 5.00	P27876
N	RNA polymerase sigma-B factor (Sigma-37) (General stress protein 84) (GSP84)	29901.4 / 5.55	P06574

A second advantage of the NP-RP-HPLC method in these studies is the speed of analysis, whereby a typical tandem-column separation of cell lysate may only take 30 minutes. In comparison, an equivalent 1-D gel separation might require at least 6 hours to obtain reasonable separation, whereas 2-D gel separation would require up to 24 hours. Further, the liquid separation can be totally automated and the data is obtained in digitized form. The gel methods are not readily automated and the gel image must be digitized. Moreover, the NP-RP-HPLC method provides improved separations over 1-D gels. Although 2-D gels provide higher resolution than the NP-RP-HPLC method, the resolution obtained is sufficient for these studies when combined with MS capabilities.

The use of MALDI-TOF MS can provide molecular weight information of the proteins separated in HPLC, while MALDI-TOF MS and CE-ESI-MS can provide mass analysis of peptides generated by enzymatic mapping. By using this information with a database search, we were able to identify many of the bacterial proteins separated by NP-RP-HPLC. Only proteins that appeared in both search results were used as possible candidates for identification. It should be noted that many of the collected fractions from NP-RP-HPLC were shown to contain two or more proteins when sized by MALDI-TOF MS. For these fractions, the MS-Fit database search on the tryptic peptides was performed twice (or more) using the corresponding mass range for each protein. When digesting these fractions, a blank tryptic digest was always performed and then analyzed on MALDI-TOF MS for trypsin autolysis peaks. A continuation of this work will involve performing another digest, possibly CNBr followed by trypsin, to further ascertain the identity of these proteins as well as tandem MS analysis on peptides of interest.

Both MALDI-TOF MS and CE-ESI-MS were used together to confirm the validity of protein identification. CE-ESI-MS detects mainly the lower mass range (up to 1500 Da) of the tryptic peptide fragments, whereas MALDI-TOF MS is more efficient in detecting peptides between 1000 and 3000 Da. Due to interference from MALDI matrix, peptides of 500 Da and under are generally suppressed. The two methods complement each other in terms of total peptide coverage of the total protein map. By combining the peptide maps from both techniques, the total peptide coverage for CSP C, for example, was increased to 85 % (55/66 amino acids matched). Similar results were obtained for other proteins in Table 1.

It should be noted that a database for ED 7-3 is currently not available for MS-Fit searches. In these studies, *B. subtilis* was used to conduct the MS-Fit searches. *B. subtilis* is the species phylogenetically closest to ED 7-3, for which the genome has been completely sequenced and annotated. Even though strain ED 7-3 (*Exiguobacterium sp*) and *B. subtilis* are distinct phylogenetic units, and may not be closely related, the mass spectral tryptic mapping procedure from CE-ESI-MS and MALDI-TOF MS provided reasonable close matches to known proteins of *B. subtilis*. However, the unavailability of an appropriate existing database for other Permafrost bacteria under study in this work, would limit the usefulness of these methods. The ability to sequence these organisms and generate such database is an important part of this study and makes our MS methods powerful tools for identification and sequence verification.

C. Experimental Plan

Objective 1. Permafrost Structural Genomics: determine the complete genome sequence of one permafrost bacterium and obtain substantial coverage of nine additional isolates.

The permafrost bacterial isolates that we will be examining are ones that had been frozen for a minimum of 20 to 30k years up to more than 1 million years. A fundamental question raised is whether these bacteria have unique genes that have enabled them to remain viable in this environment for such long periods of time. As a first step to address this question, we will determine the complete genome sequence of one permafrost organism and obtain substantial sequence coverage of nine additional isolates. These results will provide a broad view of the "informational content" of the genomes such as knowledge of the basic housekeeping genes that are carried by the permafrost bacteria; insights into the metabolic pathways used by the bacteria, some of which might be involved in the synthesis of compatible solutes or other low molecular weight cryoprotectants; and may provide signs of peculiar genome features (sequence or organization) that may be a consequence of the bacteria having been exposed to low background radiation for long periods of time while confined to an environment capable of supporting only low metabolic rates. Of paramount interest, however, will be a gaining of insight into the novel informational content of the genomes. How much of the genome of each isolate is composed of novel genes? Are any of these novel genes highly conserved among the permafrost bacteria even though they represent distantly related species (potentially signaling significant lateral gene transfer)? Are such novel genes present in closely related bacteria that reside in warm environments that have experienced little or no freezing for long periods of time? What is the nature of the polypeptides encoded by the novel genes? Do they have structural and biochemical features that resemble the highly expressed COR and LEA-like polypeptides associated with freezing tolerance in plants? The results from the sequencing efforts will provide fundamental information needed to address these questions. In addition, they will provide the basic resources required for the expression profiling experiments that will be conducted as part of our Functional Genomic analyses, Objective II of our Experimental Plan.

a. Choice of permafrost bacteria to be sequenced.

Table 2 describes the characteristics of a selected group of 10 permafrost strains which appear especially promising candidates for genomic studies. These strains were chosen following a careful consideration of their potential to serve as model systems for special adaptive features of relevance to survival in the permafrost. The selection includes representatives of all lineages that have been isolated from the very ancient sediments as well as lineages predominant or exclusively isolated from relatively young sediments (e.g. *Planococcus, Psychrobacter*). The high-GC gram-positives, which predominate in permafrost, are well-represented. In the case of *Exiguobacterium*, the two listed strains represent isolates from sediments of markedly different age, thus providing special opportunities for comparative genomic analyses. In addition, the selection includes bacteriologic features which are predominant among permafrost strains, and which may have special adaptive physiology roles, such as pigmentation and motility. It is of interest that both of these features have been noted in cultures grown at -2.5°C. It would be of interest to pursue the complete genome sequencing of

strain ED7-3, since this strain has been already included in the proteomic studies, and pursue partial genomic sequencing (80% coverage) of the other nine strains.

Table 2. Characteristics of selected strains from Siberian permafrost of special interest for genomic studies.

Strain	Age of sediment (years)	Group	Bacteriologic features	Closest relative in database
255-12f[1]	2-3 million	High-GC gram-positive	Pigmented (yellow), rods, motile, growth at –2.5°C	*Arthrobacter citreus*
215-93f[1]	20-30K	High-GC gram-positive	Pigmented (yellow), rods, motile, growth at –2.5°C	*Arthrobacter sulfureus*
190-4f[1]	200-600K	High-GC gram-positive	Pigmented (orange), oval rods, growth at –2.5°C	*Rathayibacter tritici*
3361-4a-f[1]	2-3 million	High-GC gram-positive	Pigmented (orange), cocci, motile, growth at –2.5°C	*Curtobacterium* sp.
ED7-3[2]	20-30K	Low-GC gram-positive	Pigmented (orange), rods, motile, growth at –2.5°C	*Exiguobacterium sp.*
255-15f[2]	2-3 million	Low-GC gram-positive	Pigmented (orange), oval rods, growth at –2.5°C	*Exiguobacterium sp.*
45-18[2]	20-30K	Low-GC gram-positive	Pigmented (orange), cocci, growth at –2.5°C	*Planococcus sp.*
3361-2r[3]	2-3 million	Gram negative α proteobacteria	Pigmented (orange), rods, motile, growth at –2.5°C	*Sphingomonas sp*. Ant20
23-9	20-30K	Gram negative Favobacterium-Cytophaga	Pigmented (yellow), rods, growth at –2.5°C	*Flavobacterium sp.*
262-3[3]	20-30K	Gram negative γ proteobacteria	Rods, growth at –2.5°C	*Psychrobacter sp.*

[1]Phylogenetic placement shown in Fig. 2: High GC gram-positives
[2]Phylogenetic placement shown in Fig. 2: Low GC gram-positives
[3]Phylogenetic placement shown in Fig. 2: Proteobacteria

Further characterization will determine whether these strains satisfy additional criteria of relevance for genomic and genetic studies including: 1) the genome should be 2-3 Mb or less (i..e., an average size for bacteria, or smaller; in this regard it is important to note that we have recently determined that the *Psychrobacter* isolate 262-3 listed in Table 2 has a genome size of 2 Mb); 2) the isolate should be more freezing tolerant that typical mesophiles such as *E. coli*, and preferably, would increase in freezing tolerance upon low temperature treatment (as noted above, this characteristic has been demonstrated for several of the permafrost strains); 3) there should be the potential for developing a genetic system for the organism (e.g., it should be sensitive to standard antibiotics used in transformation selection); and 4) for at least some isolates, we will want to have available closely related bacteria from an environment that has not been exposed to

freezing to conduct comparative genomic and adaptive physiology studies. We anticipate that the availability of the *Exiguobacterium* isolates will be of substantial value in this regard. Currently our collection includes a "temporal freeze series" of closely related *Exiguobacterium* strains derived from permafrost layers that have been frozen for 20-30K, 200-600K, and 2-3 million years, as well as from modern tundra soil.

2. Genome sequencing

DNA sequencing will be carried out at the DOE Joint Genome Institute using a whole-genome shotgun approach. For complete genome sequencing, the majority of sequence will be generated from ~2 kb and ~10 kb insert libraries (6x and 3x sequence coverage respectively), along with a small amount of fosmid end-sequence (5x genome coverage). The JGI sequencing operation is a high-throughput facility employing 84 capillary sequencers (MegaBACE 1000). At the current pass rate (~85%), readlength (550bp) and throughput (>1 million lanes per month), it is possible to draft sequence a 3 Mb genome in approximately 1 day of sequencer time (1-2 weeks prep time). Sequencing from both ends of plasmid DNA generates paired-end sequences critical for providing order and orientation information for the contigs resulting from sequence assembly. This, combined with a new whole-genome assembly program currently in development at the JGI, is expected to produce near-finished quality sequence from the shotgun data. The few remaining gaps and low quality sequence regions can be addressed manually.

The resulting genome sequence will be run through our semi-automated microbial annotation pipeline. Gene models are selected using a combination of 3 gene-finding programs (Generation, Glimmer, Critica), and annotated based on BLAST match, as well as CoGs and metabolic pathway assignments. The various RNA genes as well as sequence repeats are identified separately. Final assignments and genome annotation will be done by manual curation by both JGI personnel and outside experts. Of particular interest will be predicted ORFs encoding polypeptides similar to described antifreeze proteins, as well as ORFs similar to the freezing tolerance associated COR and LEA-like proteins of plants.

DNA sequencing of 9 additional microbes will be performed at 2-3x coverage using ~2kb insert libraries; a process estimated to require ~2-3 days of sequence capacity. This scanning should be sufficient to identify the majority of genes in these microbes. Plasmid DNA from all ~2 kb libraries will be made available for ORF amplification using a common set of primers.

3. Gene conservation.

Computational comparisons of DNA sequences obtained for the permafrost bacteria will enable us to determine whether there are genes, other than ones involved in "housekeeping" functions, that are conserved between the permafrost bacteria. Such genes would be of great interest, especially if found among distantly related isolates (as determined from the 16S rRNA gene sequences) as they might have roles in freezing tolerance and might be examples of lateral gene transfer. We will also, however, determine whether the genes from a given "sequenced" permafrost isolate are present in closely related isolates obtained from either the permafrost or "warm" environments. Genes present in the permafrost isolates and missing from the "warm" isolates would again be candidate "freezing tolerance" genes.

We will conduct this comparative analysis by "DNA-DNA" microarray analysis. Microarray slides will be prepared from PCR products generated from the genomes of the sequenced permafrost isolates. With the organism for which we have determined the entire genome sequence, we will design oligonucleotides to amplify PCR products for all of the open reading frames that are predicted to encode proteins. Given what has been found in other bacteria for which there is complete genome sequence, we anticipate there being about 3000 open reading frames (assuming a 2.5 Mb genome). For the permafrost genomes that have not been completely sequenced, we will choose about 2000 plasmids carrying open reading frames (we will include all "novel" sequences) and amplify PCR products using a set of primers that surround the cloning site in the plasmid (the same set of primers will be able to be used for all of the cloned inserts). Because the yield of PCR product can vary quite a bit for a given reaction (more than 10 fold), we will quantify the PCR products and adjust all DNA concentrations to a fixed amount before preparation of arrays. Preparing thousands of plasmids, PCR reactions and adjusting DNA solutions will be a major time and resource investment. Fortunately, the MSU Genomics Technology Support Facility has BioMek 2000 and Quiagen 3000 robots which can be programmed to carry out many of the repetitive manipulations. Moreover, Dr. Tom Newman, who is the director of the facility, has extensive experience in handling, cataloging and storing more than 35,000 clones and DNA preparations from previous projects, including a pioneering Arabidopsis EST project conducted at MSU (Newman et al., 1994).

The microarray slides will be hybridized with Cy-labeled genomic DNA isolated from the "reference" (i.e. organism from which the PCR products were produced) and "test" (other isolates from permafrost and warm environments) organisms. Standard protocols will be used (e.g., http://cmgm.stanford.edu/pbrown/protocols/4_genomic.html). In brief, total DNA from the homologous and test organisms will be labeled with either FoluroLink Cy5-dCTP or Cy3-dCTP (Amersham Pharmacia) by random-priming (e.g., using the BioPrime labeling kit of Gibco/BRL). The probes will be purified, combined, denatured and then hybridize to the microarray slides as described (http://cmgm.stanford.edu/pbrown/protocols/5_hyb_yeast.html). Hybridized arrays will be scanned with a GenePix 4000 laser scanner (Axon) using laser lights of wavelength at 532 and 635 nm to excite the Cy3 dye and Cy5 dye, respectively. Fluorescent images will be captured as multi-image tagged image file format (TIFF) and analyzed with GenePix Pro 3.0 software (Axon) and software obtained from the National Human Genome Research Institute (NHGRI). This publicly available software consists of a suite of programs that performs spot location; background calculation; signal intensity measurements for each channel; and calculations of the ratio of intensities. The software normalizes the signals and provides a statistical estimate of the significance of the ratio values. This data will be entered into a database constructed using the schema provided by the NHGRI. The NHGRI has also developed tools that allow sophisticated analysis of large data sets. Using Websql from Sybase and Java applets from NHGRI, the database containing the accumulated array data can be queried using a web browser.

Although microarray analysis is becoming incorporated into more and more research programs, it is not a trivial undertaking. However, the Tiedje, Lenski and Bennett laboratories have experience conducting such analyses and our colleagues at MSU have been at the forefront of structural and function genomic analyses. Researchers at MSU, including Thomashow, were

among the first to conduct a large scale Arabidopsis EST project (Newman et al., 1994). Moreover, faculty at MSU are investigators in the Arabidopsis Functional Genome Consortium (AFGC), an NSF-funded plant genome project (http://afgc.stanford.edu/). The specific task that MSU has in this multi-institutional research effort is to conduct microarray analysis experiments in collaboration of individual labs in the Arabidopsis research community. In short, we will have access to requisite equipment (robotics, slide printers, array readers) and technical expertise to carry out the comparative DNA-DNA experiments described here and the RNA profiling experiments described below.

Objective 2. Permafrost Functional Genomics: gene expression profiling, genetic analysis and targeted protein exploration to identify genes and proteins that enable bacteria to inhabit the permafrost environment.

As in space, low temperature and low water activity are cardinal characteristics of the permafrost environments. Thus, a fundamental "first stage" goal of our research program will be to detail how the genomes of the permafrost bacteria respond to these conditions at a gene expression level. How much of each permafrost bacterial genome is devoted to encoding cold- and dehydration-responsive genes? What is the nature of the genes that are induced in response to low temperature and dehydration stress? Are bacteria in a frozen environment capable of altering gene expression? These experiments will provide a broad picture about the numbers and types of genes that are specifically induced at low temperature and low water activities. In addition, they will set the stage for detailed genetic analysis of genes and proteins that enable bacteria to inhabit the permafrost environment. And finally, we will conduct a targeted search for candidate proteins with potential roles in freezing tolerance. In particular, we will explore whether the permafrost bacteria synthesize antifreeze proteins and potential cryoprotective COR and LEA-like hydrophilic polypeptides.

1. RNA microarray profiling

RNA microarray analysis is one approach that we will used to determine how gene expression in the permafrost bacteria is effected by temperature and water availability. DNA microarrays will be prepared as described above for the "DNA-DNA" hybridization experiments. RNA will be isolated from bacterial samples, Cy-3- and Cy5-labeled probes prepared and microarrays hybridized using protocols that have been proven effective with other bacteria (Richmond et al., 1999; Talaat et al., 2000; de Saizieu et al., 2000; http://cmgm.stanford.edu/pbrown/protocols), though it is likely that modifications will be needed to optimized the procedures for our specific isolates. The bacterial isolates will be grown in liquid culture at temperatures ranging from the maximum growth rate of the isolate down to 0°C (e.g. 20, 10, and 0°C). In addition, we will expose cells to subzero temperatures. High subzero temperatures, such as -2°C, will not result in freezing due to the solute concentration of the medium, and thus, will simply represent a continuum of temperature range. However, at lower subzero temperatures, freezing will occur which, as described above, will add a second environmental "stress," namely cellular dehydration. At -5°C, for instance, the nonfrozen liquid will be at approximately 2.5 osmolar (the amount of unfrozen water at -5°C will depend on the starting osmolarity of the medium.)

Our primary interest is to determine what genes are continually expressed during active growth at a given temperature. In these experiments, bacterial cultures will be maintained at the temperature of interest for many generations and gene expression will be determined in cells from exponential phase. This strategy will hold for all temperatures where the cultures remain liquid. However, the situation becomes more problematic with the cultures are in the frozen state. In this case, we will take cells that have been grown at relatively high (e.g. 20°C) and low (e.g. 0°C) temperatures and transfer the cultures to subfreezing temperature of -5, -10 and -20°C. RNA samples will be prepared from the starting cultures before freezing and harvested from the frozen cultures after 1 week, 1 month, 1 year and then once a year thereafter. The major question will be whether the bacteria are capable of altering gene expression in a frozen environment and if so, to determine what portions of the genome are affected. In our experiments, we will also expose bacteria to osmotic "stress" to determine how they respond to decreased water activity at a given temperature. The osmolarity of the cultures will be adjusted by the addition of molecules that the organisms cannot metabolize such as sodium chloride.

Managing the data associated with the proposed microarray experiments will require a well-planned bioinformatics and laboratory information management (LIMS) support system. This is due, in part, to the fact that the volume of data produced by these projects is very large. As the research community gains more experience with microarray technology, the preferred algorithms for analysis of microarray data continue to evolve. In addition, our data management solution needs to be able to serve geographically distributed collaborating researchers (web based). These features are not specific to microbial microarray projects. However, microbial projects present some specific additional challenges. For example, unlike EST-based efforts, we are not guaranteed that each probe targets a single gene, or even a single transcription unit, and our concept of what each probe targets will likely change over the course of the project. It will be important for the information management system to automatically link all array hybridization results to our most current target picture. Other unique features include the project's multi-organism focus and resulting need to track potential for cross-hybridization.

Tiedje and co-workers are currently developing the data management system for another microbial microarray project and intend to use this system with modification on this project. The system is being built on top of the Oracle RDBMS and uses a layered architecture conforming to the Sun Enterprise Java specification. It will consist of one unit that will provide tracking of the probes and other reagents used to construct the microarrays and perform the hybridizations, including organism, gene target, and probe synthesis aspects. A second unit will store post-hybridization scan and analysis data in a software package neutral format. The system will accept and provide these data through a web interface to support remote collaborators. We have completed design and implementation of the 20 relational tables making up the probe tracking unit. We are currently implementing the Java - web layers to handle the requests (queries) required for this unit. In addition, we have developed detailed designs (schemas) for the second unit and expect to begin implementation as soon as the first unit is operational. We believe that this information system can be easily modified for the work proposed here. We anticipate that most of the work will involve gathering and implementation of project-specific use cases and presentation logic, with only relatively minor changes necessary in the lower levels of our layered architecture. We will need to implement an additional unit to periodically search for and store updates to our understanding of our probes' targets.

2. Proteomic profiling

As described above, we will use 2-column nonporous (NP) RP-HPLC to profile the highly expressed protein from selected Permafrost bacteria. This profiling will be performed on bacteria cultured at various temperatures used in the RNA profiling experiments. The major proteins that show differential expression between the temperatures will be collected in the liquid phase and sized for molecular weight (MW) by MALDI-TOF MS. The proteins will then be digested by trypsin and CNBr to obtain peptide maps, which will be used to identify the proteins. An important aspect of this work is that the database will now become available for these bacteria due to DNA sequencing efforts of the group. MALDI-TOF MS and CE-ESI-MS can then be used to perform detailed sequencing of the proteins from multiple digests. In addition, MS/MS is available in our CE-ESI-MS experiments so that even small changes in protein structure can be detected and post-translational modifications (PTM) detected against the database.

Once the expressed proteins that change as a function of temperature are identified, gene knockout experiments can be performed in collaboration with the Kathariou laboratory. These gene knockout experiments can be used to determine which of the differentially expressed proteins at low temperature are significant in terms of cold tolerance. Changes in the expressed of the protein profile as genes are knocked out will be monitored using 2-column NP-RP-HPLC. In addition to knocking out individual proteins, one might expect changes in protein-protein interactions and the resulting cascades downstream. These changes can be monitored using our liquid separation techniques. The proteins that change in the subsequent protein profile can be collected and identified using MALDI-TOF MS. Furthermore, these proteins in the altered profiles can be sized by MALDI-TOF MS for MW and detailed sequencing performed by CE-ESI-MS/MS. One might expect that as protein interactions are altered, one would observe various modifications on the proteins expressed.

An important aspect of these studies will involve studying the evolution of the proteins that impart cold tolerance. Once key proteins are identified, they can be collected and isolated from permafrost bacterial strains ranging from 10^6 years to 6×10^4 years. CE-ESI-MS/MS can be used to perform detailed sequence analysis on these proteins to identify changes in sequence or PTM that have occurred through evolution. These analysis will be combined with functional analyses with a view to understand the evolution of specific activities. Integral to these studies will be application of the computer modeling methodologies developed in the Goldstein laboratory for studying evolutionary changes in protein sequence and function (see below).

An important feature of the liquid separation methodologies being developed in the Lubman laboratory is that they can be used to prepare significant amounts of protein for further analysis. Once proteins are identified as being significant to life at low temperature, then relatively large quantities can be collected using the nonporous separations. Using 8 mm diameter columns specially prepared for us by Eichrom Technologies, we can load > 1 mg of sample on the 2-column separation. Considering the number of highly expressed proteins observed, as much as 30-50 μg of each major protein can be collected. This can be performed in 10-20 minutes. Further isolation of the proteins can be performed on an additional NP column

with different properties. The isolated proteins can be studied for a number of properties including hydrophobicity/hydrophilicity, structure, effects of proteins on membrane layers, effects on dehydration properties of cells, antifreeze activity, etc., by McGrath, Thomashow and collaborators.

In initial work, the 1-D/2-column experiment will be sufficient for identifying the highly expressed proteins related to cold tolerance. However, as the bacterial systems and key proteins of interest are identified, more detail on the total protein expression might be required. This will be performed using a preparative 2-D liquid separation that has been used to successfully map large numbers of protein cells in our lab (Wall et al., 2000). The method uses a liquid IEF-based separation in a Rotofor device, which produces 20 pH fractions. The loadability of this device can be in the tens of mg range so that large amounts of protein can be prepared. In addition, relatively low abundance proteins can be detected using this method. Each *pI* fraction is then separated by the NP-RP-HPLC to produce a total 2-D map of the protein content. This map covers a pH range from pH 2.5 to 11, therefore, it is more comprehensive than most 2-D gels. It can resolve the lower MW proteins (< 15 kDa) not readily resolved by gels, which are so important in the proposed work. Furthermore, by the use of solubilization methods including 9 M urea and n-octyl (sugar)pyranoside (a nonionic detergent), large proteins > 100 kDa including protein complexes, can often be separated and detected. The method still uses a total liquid format and is readily interfaced to our MS analysis. It will be used to obtain more detailed information on the less highly expressed proteins during the cold shock processes and the various protein cascades that result. Although these detailed 2-D separations require more effort than the 1-D liquid separation, they still can be prepared in 5-7 hours, provide pure proteins in the liquid phase for future study and provide a 2-D map for study of the less abundant protein especially in the lower mass range required in this work.

3. Genetic analysis

The overall aim of the genetic analyses is to identify genes that have key roles in cold tolerance. We will address this aim using three experimental approaches.

a. Experimental approach 1: mutagenesis of candidate genes identified from the RNA and proteomic profiling studies

The experiments described above will point to candidate cold tolerance genes. These would include genes that are induced in response to low temperature and/or dehydration as well as genes that encode potential antifreeze proteins or polypeptides that resemble the COR and LEA hydrophilic polypeptides of plants. To evaluate the functional role of these genes, mutants will be constructed in which expression of the gene is inactivated. The choice of the mutagenesis scheme to be used will depend on the strain. In the case of low-GC gram-positive bacteria (e.g. strain ED7-3 and Planococcus strains), suicide vectors utilized for insertional inactivation of genes in *B. subtilis* will be employed. For instance, the temperature-sensitive shuttle vector pKSV7 (Smith and Youngman, 1992), and its conjugative derivative, pCON-1, can be introduced to the permafrost strain by electroporation and conjugation, respectively. In the latter case, conjugation will be between an antibiotic-resistant variant of the permafrost strain (we have successfully generated streptomycin-resistant derivatives of numerous permafrost

strains in the laboratory) and *E. coli* S17-1, which can serve as conjugative donor of the plasmid pCON-1. The selection marker on the vector (chloramphenicol) can be used with several of the permafrost strains, which are chloramphenicol-sensitive. These integration mutagenesis schemes will require the cloning of an internal portion of the gene into the vector, and selection for integrants under conditions which prevent replication of the plasmid in the permafrost bacteria. The available temperature-sensitive plasmids such as pKSV7 and pCON-1 will be utilized first. If these replicons prove suboptimal (e.g. because the restrictive temperature for replication is lethal for the permafrost bacteria) alternative suicide plasmids which can be efficiently introduced into the permafrost strain but cannot replicate will be utilized.

Our intent is to start the genetic analysis studies with strain ED7-3, since proteomic analysis of this organism is already under way and at least partial genomic sequences will be available. Another pre-disposing factor for the choice of this strain is the potential to utilize genetic approaches already established for several other low-GC gram-positive bacteria (e.g. *B. subtilis*, *Listeria*, lactic acid bacteria). In addition, strain ED7-3 is a member of a cluster of genetically closely related strains which have been isolated from permafrost sediments of variable ages (20,000-40,000 years, 200,000-600,000 years and 2-3 million years) as well as from modern tundra soil. Depending on genomic conservation among the strains, it may be possible to utilize PCR (polymerase chain reaction) to amplify the homologous genes from the different members of the cluster, thus identifying possible changes in nucleotide sequences that may be of relevance to the age of the bacteria in the sediments. The role of genes of special interest in ED7-3 can be investigated in similar strains from variable ages, utilizing mutagenesis approaches used for ED7-3.

Once the mutants are constructed, they will be used in functional studies to determine the impact of the mutation in phenotypes of interest, primarily low temperature growth and survival upon exposure to sub-zero temperature and freeze-thawing. In addition, the mutant strains will be characterized at the proteomic level in the Lubman laboratory to determine the impact of the mutation on the levels of expression of different proteins. Mutants of special interest will also be characterized at the level of RNA expression utilizing microarrays as described above. Such analyses will determine whether a single protein and/or RNA species is impacted by the mutation, or whether a number of proteins and RNAs are affected. The latter, for instance, may be expected if the inactivated gene encodes a regulatory protein essential for transcription of several others. An additional reason concerns the possibility of polar mutations. Depending on the genomic organization of the gene of interest, the insertional mutation may have polar effects on downstream genes that may be co-transcribed with the gene that harbors the insertion. In such cases, and if the mutant appears to have interesting phenotypes, additional work will be done to construct mutants bearing an internal gene cassette that does not have polar effects or in-frame deletions of the gene of interest. The genetic tools for such constructs are available for *B. subtilis* and in-frame deletion constructions have been used extensively in the Kathariou laboratory for mutant constructions in *L. monocytogenes*. Since cassette constructs and deletions are substantially more labor-intensive than insertional inactivations, they will be pursued only if (1) the gene is indeed a member of an operon, and is followed by additional downstream genes and (2) the insertionally inactivated mutant does indeed have a phenotype of interest. To determine whether the gene is a member of an operon, we will rely on both the analysis of genomic organization and sequence features in the region (e.g. transcriptional orientation,

presence of terminators, length of intergenic regions) and on experimental data utilizing Reverse-Transcription PCR (RT-PCR).

An important factor in these experiments is the accessibility of nucleotide sequences for site-specific construction of mutants. It is anticipated that the complete genome of one permafrost strain and partial (at least 80% coverage) genome sequences of nine additional strains will be available during this project. Although the relevant genes are expected to be readily identified among the sequences of the completely-sequenced genome, in the case of other strains it is conceivable that the genes of interest may be absent from the sequences (i.e. may be included in the ca. 20% of the genome which will not be covered by the sequencing project), for reasons having to do with peculiarities in nucleotide sequence and under-representation of the corresponding sequences within the cloned libraries. If an especially interesting gene appears to be lacking from the known sequences, special effort will be made to clone the gene by PCR using degenerate primers constructed on the basis of the available protein sequence data. Otherwise, our attention will be focused on those genes the sequences of which have been revealed in the course of the genome sequencing projects.

Another situation to consider is the potential inability to obtain mutations is specific genes. Certain genes may have functions essential for bacterial survival, and their inactivation is expected to be lethal. Thus, inability to obtain a specific mutant (when the mutagenesis scheme is operational for the organism of interest) suggests that the protein has an essential function. The mutagenesis schemes will be pursued with a number of genes of the genome of ED7-3 early in the course of the project, to ascertain the feasibility of the insertional inactivation and/or deletion protocols. If a specific mutant cannot be obtained, additional evidence for its essential role can be obtained utilizing a recombinant strain which harbors the gene of interest cloned on a plasmid, or integrated elsewhere in the chromosome, as has been described in other systems (Graumann et al., 1997).

b. Experimental approach 2: transposon mutagenesis of permafrost bacteria

In addition to site-specific mutagenesis, transposon mutagenesis of selected strains will be done to generate mutants in selected phenotypes of interest. Preliminary data of ours suggest that the conjugative transposon Tn916 (tetracycline resistance) (Clewell et al., 1995) and its erythromycin-resistance derivative, Tn916 E can be used to generate mutants of strain ED7-3 and a similar strain, ED5138, from Siberian permafrost (20,000-40,000 yrs). Southern blot data suggested that single copies of these transposons were inserted in various locations on the chromosome of different mutants, indicating that such mutagenesis will be suitable for genetic analysis. Alternative transposon systems, such as Tn917 (Tomich et al., 1980), can be used to enhance the range of target sites on the genome. The Kathariou and Thomashow laboratories have had experience with transposon mutagenesis systems of a number of different bacteria.

The transposon mutants will be screened for several phenotypes of interest, including ability to grow at different low temperatures (e.g. 4, 0, and -2.5°C) and to survive prolonged freezing and repeated freeze-thawing. In addition, mutants deficient in pigmentation will be isolated to evaluate the impact of pigmentation on low temperature survival and freeze tolerance. Mutants which are non-motile will be isolated as the screening for such mutants is easy, and can

provide information on the efficiency of the mutagenesis scheme. In addition, the possible impact of motility on low temperature survival will be evaluated. It is of interest that in several permafrost bacteria appear to be motile when grown at low temperatures. Transposon mutants with phenotypes of interest (e.g. cold-sensitive, freeze-sensitive etc) will be evaluated at the proteomic level in the Lubman laboratory as well as in terms of RNA expression profiles using microarrays, to evaluate the impact of the mutation on expression of distinct protein and RNAs, as described above. Transposon mutations are expected to have polar effects on transcription of downstream genes in the same operon, and mutants of interest may have to be characterized further with additional mutagenesis schemes, if the insertion is indeed found to be in a location that would allow such polar effects.

To identify the inactivated gene in a transposon mutant, genomic DNA of the mutants will be digested with an enzyme that does not cut into the transposon, ligated to a vector of known sequence (e.g. pUC18) and used as template in single-specific primer PCR (Shyamala and Ferro-Luzzi Ames, 1989) employing a primer at the terminus of the transposon and a vector primer. Such PCR products will contain one of the transposon-flanking fragments. The sequence of this fragment will be analyzed against the available genomic sequences of the bacterium. Homologies with other genes and gene products in the databases may provide clues as to the function of the gene. If the gene is not present in the available genomic sequences of the permafrost bacterium, the entire gene can be obtained by inverse-PCR. The utilization of SSP-PCR and additional approaches to obtain the sequences of transposon-targeted genes has been used extensively in the Kathariou laboratory.

To confirm that the mutant phenotype is due to the transposon insertion, genetic complementation will be employed, introducing the wild type gene cloned into a plasmid that can replicate in the permafrost bacteria. Our intent is to pursue transposon mutagenesis of ED7-3 first, and follow with similar work of high GC- gram-positive bacteria from the permafrost. Of special interest would be some of the *Arthrobacter* strains, which are prevalent in permafrost sediments of various ages.

On the basis of available preliminary data we anticipate that mutagenesis of ED7-3 and other low GC-gram-positive bacteria will be accomplished without undue difficulty. If transposons other than Tn916 and Tn917 are desired, transposon systems based on Insertion sequences (IS) identified during the genome sequencing of the bacteria, as constructed for other low GC-bacteria (Maguin et al., 1996). Transposon mutagenesis of Arthrobacter may prove more challenging, as transposon delivery systems for this class of high GC-gram-positives have not been established yet. We will examine the suitability of Tn916 and Tn917 for mutagenesis of these organisms, and will include a number of different permafrost strains, to maximize the likelihood of at least some them serving as transposon recipients. Tn*5564*, successfully used in *Corynebacterium glutamicum* (Tauch et al., 1998), and Tn*5367*, utilized with *Mycobacterium paratuberculosis* will also be employed. It is expected that novel genetic tools suitable for use in the high GC-gram-positive bacteria will be also becoming available, and these will be included in the investigations as well.

Genetic complementation in trans, with the wild type gene on a plasmid, may be difficult under the low-temperature conditions of the assays. It is possible, for instance, that replication of

the plasmids may be sub-optimal at 4°C or other low temperatures (the replicons of available plasmids are derived from mesophilic bacteria). Proper complementation under such conditions may require integration of the gene in a chromosomal location. Such integration constructs can be made for mutants of special phenotypic interest.

c. Experimental approach 3: isolation and characterization of genetic variants in phenotypes of interest

In the course of our work with several permafrost strains, including ED 7-3 we encountered genetic variants in certain phenotypes, such as pigmentation, filamentation of cells, and production of polysaccharide. The variants were identified on the basis of their distinctive phenotypes and were stable in the laboratory. Such variants likely represent spontaneous mutants of the progenitor culture, and can be used to evaluate the impact of the mutation on additional phenotypes of interest, such as low temperature growth and freeze tolerance. Variants with especially interesting phenotypes in these regards will be analyzed at the proteomic level to determine the impact of the mutation(s) on expression of different proteins. Such analysis will reveal possible changes in the size of the protein conferred by the mutation. In addition, the mutant protein is likely to yield an altered peptide map.

It is recognized that the variants may harbor mutations in more than one locus, thus hindering the association between the original phenotype (e.g. pigmentation) and the functional effects (e.g. freeze tolerance). If the proteomic analysis shows that a specific protein is affected, the corresponding gene can be mutated separately to confirm the function of the protein and create isogenic mutants. This additional confirmatory work will be done only if the variant appears to have unusually interesting phenotypes in terms of cold and freeze tolerance.

4. Targeted exploration for candidate freezing tolerance proteins

a. Antifreeze proteins

As described earlier, antifreeze proteins have been described in diverse organisms including fish, insects, plants and bacteria. These studies have revealed that proteins with very different amino acid sequences and structure (secondary and tertiary) are capable of displaying the activities associated with antifreeze proteins—namely thermal hysteresis, inhibition of ice recrystallization and altering the shape of ice crystals—indicating that antifreeze activity has evolved multiple independent times. Many fundamental issues, however, remain to be addressed regarding these fascinating proteins. The precise structural features of antifreeze proteins that impart each of the associated activities is not clear; the roles that the proteins play in organisms that tolerate freezing, as opposed to avoid freezing, remains to be determined; and the diversity and evolution of proteins with antifreeze proteins has only begun to be explored. Indeed, to our knowledge, only one antifreeze protein from a bacterium has been identified, and in that case, the protein appears to have a novel activity, that of ice-nucleation (Xu et al., 1998).). Although the *B. subtilis* CspB has been shown to be essential for freeze tolerance in this organism (Willimsky et al., 1992) it has not yet been shown that the protein actually acts in an anti-freeze capacity.

Given that the permafrost bacteria have been frozen from tens of thousands to millions of years, it does not seem unreasonable to think that they are potentially a good source to "mine" for proteins with antifreeze activity. Thus, we (McGrath and Thomashow laboratories) plan to conduct a survey of all the permafrost isolates that we have obtained for antifreeze activity. The isolates will be grown at low temperature and the cells and supernatant fluids will be collected and assayed for antifreeze activities. Supernatant fluids will be tested directly and after being concentrated by lyophilization and suspension in buffer. Cells will be disrupted, fractionated into soluble and insoluble fractions and then tested for activity. The activities to be examined are the three that are best characterized: thermal hysteresis, inhibition of ice recrystallization, and effect on ice crystal shape. Thermal hysteresis activity will be measured by the capillary freezing-melting point techniques as described by DeVries (1986) and the nanoliter osmometer method described by Chakrabartty and Hew (1991). Inhibition of ice recrystallization will be measured using cryomicroscopy as described by Chao et al. (1996) and effect on ice crystal growth will be monitored as described by Griffith et al. (1997) and Sidebottom et al (2000).

If antifreeze activities are observed, we will purify the proteins, obtain a partial amino acid sequences and isolate and sequence the corresponding genes. From the sequence information, we will be able to determine whether the proteins are related in amino acid sequence or share common structural features that have been observed in other antifreeze proteins such as repeated amino acid motifs. This information will provide the foundation for future detailed analysis to understand better structure-function relationships among antifreeze proteins and the evolution of antifreeze activities (see section on computational analysis of protein evolution below). As a part of this analysis, we will test for two additional activities that have been observed with antifreeze proteins including the ability to nucleate ice formation (Xu et al., 1998) and to stabilize membranes (specifically inhibit the leakage of small molecules from liposomes during thermotropic phase transitions) (Hays et al., 1996). Of course, we would also want to make mutants of the isolates that are no longer able to produce the antifreeze protein and asses what effects this has on growth at low temperature and the ability of the isolates to tolerate freezing. Whether this can be done in the short run will depend on how easily the specific isolates can be genetically manipulated (see section on genetic analysis above). Other issues will also be of interest to address. Do increased levels of thermal hysteresis activity and ice recrystallization inhibition correlate with enhanced resistance to freeze/thaw damage of these microbes, suggesting that they are important determinants of freeze protection? Since there are lower water activities and lower temperatures in the Antarctic compared to the Arctic, are the thermal hysteresis activities and ice recrystallization inhibition capacities more pronounced in the former? What are the limits of low water activity and low temperature for Arctic and/or Antarctic microbes that express AFPs- i.e. can they survive more extreme conditions that may exist at extraterrestial sites. Are the thermal hysteresis activities or ice recrystallization inhibition capacities of AFPs produced by Arctic or Antarctic microbes unusual in any way compared to known behavior? For example, do they produce unusually potent activity or inhibition? What environmental conditions (e.g. low temperature) trigger the appearance of thermal hysteresis activity and ice recrystallization inhibition?

b. COR and LEA-like hydrophilic polypeptides

As noted earlier, it is widely observed in plants that during cold acclimation, the most highly expressed cold-inducible genes encode extremely hydrophilic polypeptides that are thought to be involved in freezing tolerance, though their specific roles remain largely unknown. Our goal here is to determine whether the synthesis of extremely hydrophilic polypeptides occurs upon exposure to low temperature in the permafrost bacteria and if so, to explore their roles in life at low temperature. In short, is the production of such polypeptides a conserved feature of organisms that must contend with freezing temperatures?

One way in which we will address this issue is to analysis the genomic sequence that we obtain for ORFs that encode cold-regulated hydrophilic polypeptides that resemble those found in plants and, in the microarray analysis experiments, determine whether they are induced in response to low temperature. In addition, we will screen directly for the synthesis of "boiling-soluble" polypeptides in cultures of permafrost bacteria that are grown at low temperature. In these cases, culture supernatants and cell extracts will be boiled, the coagulated proteins will be removed by centrafugation, and the soluble polypeptides will be fractionated in the Lubman laboratory using the NP columns described above.

If our results indicate that the synthesis of novel COR and LEA-like hydrophilic polypeptides occurs in the permafrost bacteria, we will want to determine whether they have importance roles in life at low temperature. An important approach will be to make mutants that do not express the proteins. As with the antifreeze protein experiments, however, our ability to perform these experiments will be subject to how easily we can genetically manipulate the given permafrost isolate. We are also aware that if a given isolate produces multiple COR and LEA-like polypeptides in response to low temperature, that we may not detect any effect on cold tolerance by "knocking out" one gene as the proteins may have redundant functions (a similar situation exists with the *Csp* genes in *E. coli*). In this case, we would have to construct strains with multiple mutations to asses their importance to life at low temperature. As a part of our initial analyses, the McGrath and Thomashow laboratories will also conduct *in vitro* experiments to see if the proteins have activities that have been associated with other cyroprotective molecules such as sucrose and proline, namely the stabilization of proteins, membranes and cells against freezing and desiccation damage (Csonka, 1981; Anchordoguy et al., 1987; Carpenter and Crowe, 1988; Artus et al., 1996; Hayes et al., 1996; Steponkus et al., 1998). Given some of the overt similarities between many of the novel, COR and LEA-like hydrophilic polypeptides and certain antifreeze proteins, we include in our analyses whether the proteins have activities displayed by antifreeze proteins, namely thermal hysteresis, inhibition of ice recrystallization and effect on ice crystal shape.

Objective 3. Evolutionary Adaptation to Low Temperature: experimental evolution studies on improving "cold fitness" and modeling evolutionary relationships of "cold tolerance" proteins.

1. Experimental evolution studies on cold fitness

Historically, most studies of organic evolution have been phylogenetic and comparative in their approach (Harvey and Pagel 1991; Avise 1994; Lenski 1995a; Larson and Losos 1996; Bennett and Lenski 1999). In recent years, experimental studies of evolutionary change using

bacteria and other microorganisms have flourished while addressing a wide range of questions from dynamics of adaptation and bacteria-virus interactions to molecular evolution and social behaviors (Chao et al. 1977, 1983; Dykhuizen 1978, 1990; Hartl and Dykhuizen 1979; Hall 1983; Lenski and Levin 1985; Adams and Oeller 1986; Bouma and Lenski 1988; Lenski 1988, 1995a, 1995b, 2000; Chao 1990; Lenski et al. 1991; Bull and Molineux 1992; Hillis et al. 1992; Cohan et al. 1994; Lenski and Travisano 1994; Rosenzweig et al. 1994; Elena et al. 1996; Bull et al. 1997; 2000; Schrag et al. 1997; Sniegowski et al. 1997; Nakatsu et al. 1998; Rainey and Travisano 1998; Turner at al. 1998; Velicer et al. 1998, 2000; Appenzeller 1999; Burch and Chao 1999, 2000; De Visser et al. 1999; Papadopoulos et al. 1999; Turner and Chao 1999; Vulic et al. 1999; Wichman et al. 1999; Bohannan and Lenski 2000; Cooper and Lenski 2000; Schneider et al. 2000). Bacteria have many features that facilitate an experimental approach to evolution, including their rapid generation times, large population sizes and ease of culture. Albert Bennett and Richard Lenski have been among the pioneers in using this approach, including their collaborative work on thermal adaptation in *E. coli* (Bennett et al. 1990, 1992; Bennett and Lenski 1993, 1996, 1997, 1999; Lenski and Bennett 1993; Leroi et al. 1994a, 1994b; Travisano et al. 1995; Mongold et al. 1996, 1999).

Experiments to study evolutionary (genetic) adaptation to very low temperature pose a particular challenge, namely that the absolute rate of adaptation by natural selection depends on genotypic differences in the net rate of cell growth (cell division minus cell death), and these rates are extremely slow at very low temperature. Nonetheless, we believe that we can make significant progress toward understanding the process and mechanisms of evolutionary adaptation to low temperatures by a series of four experimental studies, which are described below. The first of these concerns a further analysis of an already existing evolutionary lineage, which demonstrates the feasibility of this approach. The next three studies are proposed new experiments in which new lineages of bacteria will be evolved and subsequently analyzed. Following a description of these four evolution experiments, we then describe two types of genomic analyses that we will perform to understand the genetic bases of adaptation in all four of these experimental systems. Related studies exploring the evolution of cold fitness traits in permafrost isolates would be considered once they are more thoroughly characterized at a genetic and physiological level.

a. Experiment 1: evolutionary adaptation to 20°C, a moderately low temperature

This experiment started with a clone from a population of *E. coli* that had previously evolved in a minimal glucose-supplemented medium at 37°C for 2,000 generations (Lenski et al. 1991). From this founding clone, six new populations were propagated in the same medium at 20°C for 2,000 more generations (Mongold et al. 1996). 20°C was chosen as the selective temperature because it was within 1°C of the minimum temperature at which the ancestral population could replicate itself under the defined culture conditions. After 2,000 generations at 20°C, new clones were sampled from each of the populations and placed in competition with their 37°C-selected ancestor across a wide range of temperatures from 20-40°C. Details of competition experiments can be found elsewhere (Lenski et al. 1991; Mongold et al. 1996), but the important points are these: (1) The ancestral strain was stored as a clone in glycerol at -80°C during the time that the 20°C-selected populations were evolving. (2) All evolved clones were

also subsequently stored at -80°C. (3) Prior to competitions at each assay temperature, the evolved and ancestral clones were separately acclimated for several generations to the assay temperature; hence, any difference between them has a heritable genetic basis. (4) During the competition between the evolved 20°C clones and their ancestor, the competitors were distinguished on the basis of a selectively neutral genetic marker.

Figure 4 shows the average fitness of the 20°C-selected clones relative to their 37°C-selected ancestor across a range of temperatures from 20 to 40°C. Fitness is calculated here simply as the ratio of the number of doublings achieved by the two competitors during their competition for limiting glucose at that temperature (Lenski et al. 1991; Mongold et al. 1996). Evidently, there has been significant genetic adaptation to the lower temperature, such that the 20°C-selected clones now grow about 10% faster than their ancestor at 20°C. The 20°C-selected clones retain their competitive advantage at 27 and 32°C, whereas they are marginally less fit than their ancestor at 37°C and much less fit at 40°C.

In this experiment, populations were propagated by serial transfer with daily dilution of 1:100 into fresh medium. In order to persist against this rate of dilution, a population must achieve 6.64 (= $\log_2 100$) doublings per day. The 37°C-selected ancestral strain is just able to persist at 19.5°C, whereas it is progressively diluted to extinction at 18°C (Bennett and Lenski 1993; Mongold et al. 1996). As a consequence of their genetic adaptation to 20°C, all six of the 20°C-selected lines can persist indefinitely at 18°C against a 100-fold daily dilution, but they cannot persist at 17°C with this same dilution factor (Mongold et al. 1996). In essence, the 2,000 generations of evolution at 20°C (which required 300 days in real time) allowed us to push downward by 1.5°C the thermal limit for persistence against a 1:100 daily dilution. In the next experiment, we will use the 20°C-selected clones as starting material for continued selection at progressively lowered temperatures. Also, we will use these 20°C-selected lines as the first material for genomic analysis of the genetic basis of adaptation to low temperature. These genetic analyses can be undertaken while the three additional evolution experiments are in progress.

b. Experiment 2: evolutionary adaptation to progressively lower temperatures

In this proposed experiment, we will take the six 20°C-selected lines and use of each of them to found four new populations (24 populations in all). The new populations will then be propagated under identical conditions, including the 100-fold daily dilution into fresh medium, except they will be maintained at 18°C, where they just barely persist. At 500-generation intervals, we will assay the lower thermal limit for persistence (as defined in the preceding section) of each 18°C- evolved line, including as controls their proximate 20°C-selected progenitors and their common 37°C-selected ancestor. At such time as 50% or more of the 18°C-selected lines have evolved the capacity to persist (against a 100-fold daily dilution) at a temperature at 17°C or lower, we will then shift the incubation temperature to that new lower limit. In the event that some lines have not evolved this capacity, we will replicate other lines that did so in order to maintain a total of 24 populations at the new temperature. We will then propagate the 24 lines at this new lower temperature, assay their lower limits for persistence at 500-generation intervals, and shift them to a still lower temperature once a majority have evolved the capacity to persist at their former low temperature.

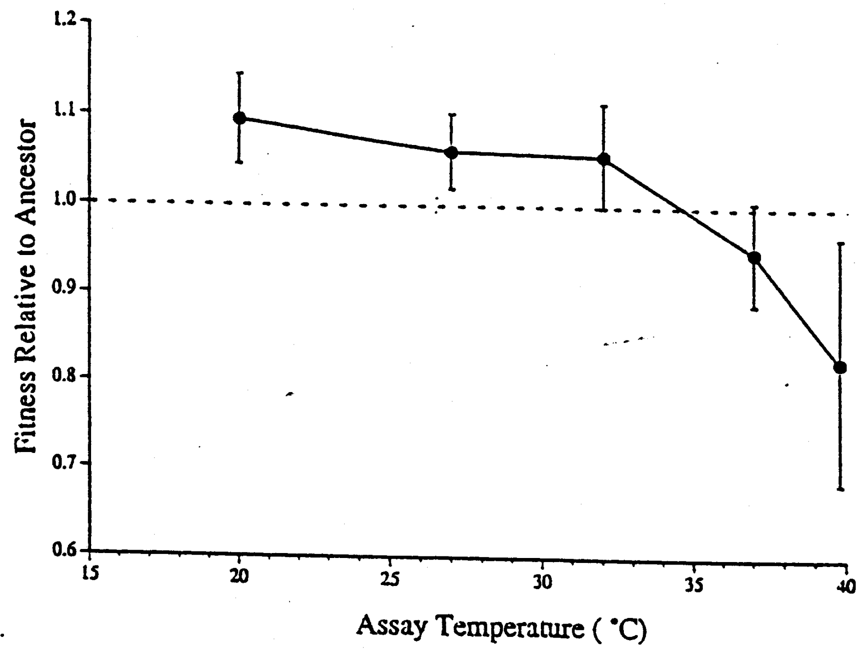


Figure 4. Average fitness of 20°C-selected lines in competition with their 37°C-selected ancestor across a range of assay temperatures. Error bars indicate the 95% confidence intervals. A value >1 indicates the 20°C-selected lines are more fit; a value <1 indicates the 37°C-selected ancestor is more fit. (From Mongold et al., 1996.)

We will continue this selection scheme indefinitely through the first 4 years of this proposal. This should allow more than 10,000 generations of evolutionary adaptation to progressively lower temperatures (given 6.6 doublings per day to offset the 100-fold dilution). Note that the Lenski lab has successfully maintained *E. coli* lines for more than a decade – more than 25,000 generations – as part of another long running experiment (Lenski et al. 1991; Lenski and Travisano 1994; Papadopoulos et al. 1999; Cooper et al. 2000; see also feature article in *Science* by Appenzeller, 1999, posted at http://www.msu.edu/user/lenski/).

All the evolved clones with progressively reduced thermal limits for persistence will be stored in glycerol at -80°C, where they and their progenitors (which serve as controls) will be available for the genomic analyses described later.

c. Experiment 3: evolutionary adaptation following a sudden shift to 4°C

Experiment 2 is predicated on the assumption that evolutionary adaptation to low temperatures is most likely to be achieved in a "stepping-stone" fashion, that is by a long series of incremental reductions in temperature. This approach is demonstrably successful (see results of Experiment 1) but it is also time consuming; for example, the first 1.5°C shift required 2,000 generations.

Experiment 3 will be run concurrently with Experiment 2. In Experiment 3, however, we will use a strategy in which we look for adaptation to a sudden and more extreme shift in incubation temperature. We have previously used this "sudden shift" approach to find several mutant clones that can grow at temperatures 1-2°C above those that allowed growth by their progenitor, which we did not obtain by more gradual changes in temperature (Bennett and Lenski 1993; Mongold et al. 1999). Specifically, for the present project, we will initiate 72 experimental populations, including six started from each of the six 20°C-selected lines and 36 started from their 37°C-selected ancestor. These 72 populations will be grown to stationary phase at 20°C, and then inoculated into fresh medium at 4°C, which is several degrees below where "wildtype" *E. coli* is reported to sustain growth in minimal medium (Ng et al. 1962; Shaw et al. 1971; Ingraham and Marr 1996). Nonetheless, there may be some mutations that permit sustained growth at this temperature, even if the resulting growth is very slow. By incubating these populations for weeks and even months at 4°C, any such mutant should increase in the population, eventually resulting in a dense, turbid culture. We will monitor these populations indefinitely, isolating clones from populations that achieve increased numbers of cells over time.

Assuming that one or more populations respond, we must confirm that: (a) the clone is not a contaminant; and (b) the difference between the 4°C-selected clone and its ancestor is a heritable, genetic difference, as opposed to some very slow and strictly phenotypic acclimation response. The former is readily determined by a series of phenotypic and genetic markers that uniquely characterize the strains used in all the Lenski and Bennett evolution experiments (Lenski et al. 1991; Bennett et al. 1992; Papadopoulos et al. 1999). The latter is best achieved by performing a series of paired assays in which the putative 4°C-selected mutant and its immediate progenitor are removed from storage at -80°C, re-acclimated to 20°C, then re-transferred to 4°C. If the difference is inherited, then the mutant should consistently achieve high density at 4°C,

whereas the progenitor should not. Alternatively, no consistent difference between the putative mutant and its progenitor is predicted if the original growth at 4°C represented a slow acclimation response.

If 4°C-adapted mutants are found and confirmed, they will be subjected to the genetic analyses described below. And if we find multiple 4°C-adapted mutants, then the above experimental design with 72 populations will also allow us to test the hypothesis that prior adaptation to 20°C (as opposed to 37°C) genetically predisposed the populations to be able to make the shift more often and/or faster at 4°C. Similarly, among the six independently evolved 20°C-selected lines, we can assess whether certain ones were better able than others to make the evolutionary transition to 4°C. Thus, we have an experimental system that allows us to examine not only the process of evolutionary adaptation, but also the potential role of historical contingency in promoting or constraining adaptation (Gould 1989; Travisano et al. 1995; Huey et al. 2000).

d. Experiment 4: evolutionary adaptation to freeze/thaw cycles

As discussed elsewhere in this proposal, the capacity to survive and even grow at a sub-zero temperature is likely to present a bacterium with a different set of physiological challenges from growing in liquid medium at a low temperature, including hyperosmotic stress. Therefore, as a bridge between the two realms, we propose in Experiment 4 to examine the evolutionary adaptation of *E. coli* to repeated cycles of freezing and thawing.

At first glance, it might seem improbable that a mesophile like *E. coli* could evolve – over an experimental time scale – new mechanisms to withstand freezing, and we recognize that this proposed experiment is a long-shot in that regard. On the other hand, many compatible solutes that allow an organism to withstand freezing (such as proline, trehalose, and glycerol) are in fact produced by *E. coli* as routine parts of its existing metabolism. Indeed, production of trehalose is induced in wild-type *E. coli* by high external osmolality, and some mutants of *Salmonella* over-produce proline and can thus grow at higher osmolality than their wild-type counterparts (Csonka 1981; Strom et al. 1986; Ingraham and Marr 1996). It is quite possible, therefore, that existing metabolic pathways might be co-opted to allow cells to accumulate an adequate quantity of a cryoprotective metabolite.

We will pursue two experimental approaches to evolve populations under a freeze-thaw regime. In the first one, we will allow populations to grow to stationary phase (where cells are generally more robust to stress than during exponential growth), then subject them to varying duration of freezing without providing any extrinsic cryoprotective substance. Survivors will then be transferred to fresh medium, allowed to grow to stationary phase, and again subjected to sudden freezing. If this protocol is unsuccessful – for example, there may be few or no survivors even for the briefest periods of freezing – then we will repeat the experiment, except by varying the concentration of glycerol added to the medium while keeping the duration of freezing constant at 24 h. We know, from past experience, that our *E. coli* strains survive in high numbers for more than a decade at -80°C in medium containing 15% glycerol. In essence, we will identify the glycerol concentration that just permits survival and then gradually reduce the

concentration below that limit to see whether we can identify mutant genotypes with reduced requirements.

As with Experiment 3, we can run this experiment with various 20°C- and 37°C-selected lines, to determine whether they are differentially predisposed to adapt to the freeze/thaw challenge. As with all of the experiments, any cryotolerant mutants that are isolated will then be subjected to the genomic analyses described below.

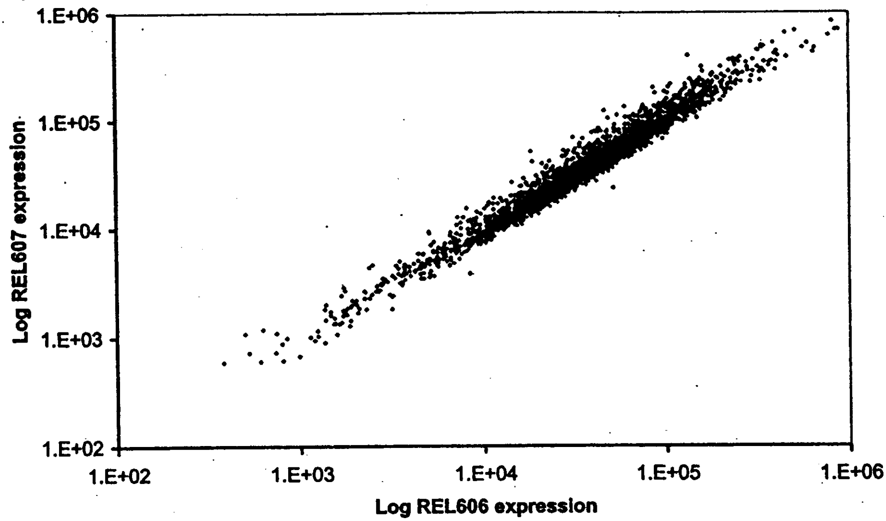
e. Genetic analysis

Genomic mRNA expression profiles. We will examine the mRNA expression profiles of selected lines from each of the four evolution experiments described above, along with the profiles of their progenitors as paired controls. To do so, we will use the commercially available Panorama™ *E. coli* gene expression arrays from Sigma-Genosys Biotechnologies. Each expression analysis module will ideally consist of four arrays, with mRNA isolated from the following cell populations: (i) the ancestral strain grown under the conditions used during its own evolution; (ii) the selected mutant grown under the ancestral conditions; (iii) the selected mutant grown under conditions corresponding to its own evolutionary derivation; and (iv) the ancestral strain grown under the conditions used to obtain the derived mutant. (In practice, the fourth category may sometimes be impossible to include because ancestral cells will be nonviable under the corresponding selective regime.) Standard protocols will be used to isolate the mRNA; synthesize ^{33}P-labeled cDNA using primers that exclude rRNA and tRNA; hybridize the labeled cDNA probes to the array; and scan the arrays using a PhosphoImager to record the intensity of hybridization (Tao et al. 1999).

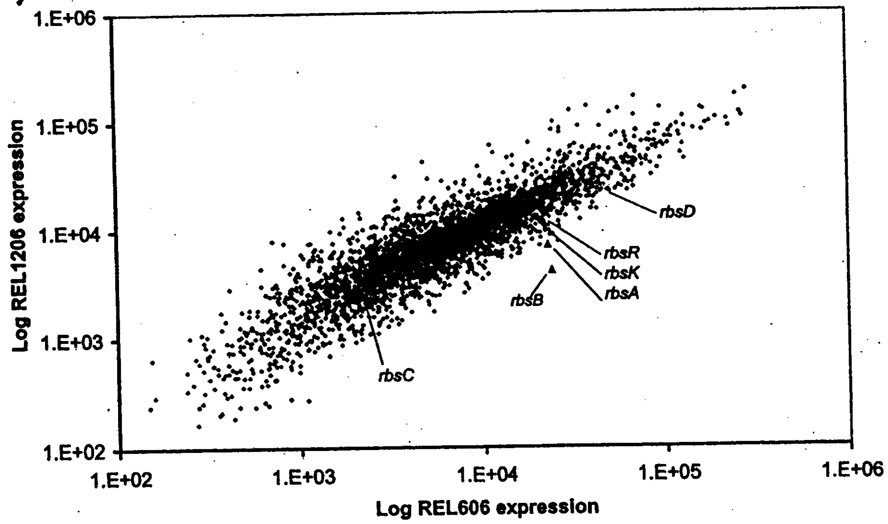
Both the Bennett and Lenski labs have successfully employed these gene expression arrays. Figure 5 shows the results of a preliminary microarray experiment using three clones from the long-term evolution experiment from the Lenski lab (Lenski et al. 1991; Lenski and Travisano 1994; Cooper and Lenski 2000). In Panel A, the relative intensities of all mRNA are contrasted for the ancestral clone (REL606) and another clone that differs only by a single neutral mutation that is used as a marker in competition experiments (REL607). Note the extremely tight correspondence of the expression profiles for these two genotypes, which show the high reproducibility of these profiles. Panel B shows the expression profile of a selected genotype (REL1206) isolated after 2,000 generations of evolution at 37C against its ancestor (REL606). Notice the greater variation between these two strains in their expression profiles; more than 100 genes differ in their level of expression by at least 3-fold, whereas fewer than 10 did when the two ancestral variants were compared. Among the genes that show much lower level of hybridization in the evolved mutant are those in the *rbs* operon, which encodes the ability to grow on ribose. This catabolic function has in fact been lost by this evolved genotype; the underlying mutational event has been characterized; and subsequent genetic manipulation and competitions have demonstrated the selective benefit of this mutation (Cooper 2000; Cooper and Lenski 2000).

Transposon tagging of selected mutations. A second approach that we can use to identify the mutations that were selected under any of the four selective regimes is transposon tagging. This approach is considerably more labor-intensive than genomic expression profiles,

(A)



(B)



and hence can be applied to only the most interesting evolved lines. But it allows the unambiguous identification of the exact mutational event responsible for a given improvement under one of the selective regimes.

To identify a beneficial mutation, we apply the familiar genetic approach of transposon-tagging in a novel way to what seems like a daunting problem — how to find one or a few beneficial mutations, among 5 x 10^6 bp and in nearly isogenic lines, whose only known phenotypic effect is improvement in growth rate or survival under a particular thermal regime. Using a mini-Tn*10* transposon that carries an antibiotic-resistance marker and is engineered to insert randomly into the genome (Kleckner et al. 1991), we create ~1500 sub-clones from an evolved line of interest, each one carrying an inserted marker. We mix these sub-clones, infect them with transducing phage P1, use the lysate to infect the ancestor, and select for resistance. Most ancestral cells that acquire the resistance marker will not carry a beneficial mutation, but among the 1500 marker insertions, a few will be near a beneficial mutation; when these markers are transduced into the ancestor, the beneficial mutation will often be co-transduced. We then screen the transduced ancestors for those that have a fitness advantage under the relevant thermal regime (corresponding to their selective history). In those cases where the transductant has improved fitness, we now have a well-defined marker that is linked to the trait of interest. We can then quantify the linkage by measuring the co-transduction frequency between the marker and improved fitness; we use these data to estimate the physical distance between the marker and the beneficial mutation; and we sequence regions on either side of the marker at the appropriate distance, thereby finding the beneficial mutation of interest.

The Lenski lab has successfully applied this approach to one of the 37°C-selected long-term lines, which allowed them to: (i) identify a 1-bp insertion in an upstream regulatory region of two co-transcribed genes involved in cell wall synthesis; (ii) demonstrate that this mutation confers a fitness advantage of 8% or so on the ancestral background at 37°C; and (iii) show that this mutation swept through that population between generation 500 (at which time 0 of 50 clones sequenced carry the mutation) and 1500 (when 29 of 30 clones carry it). Moreover, after having transduced a beneficial mutation of interest into the ancestral background, one can use genomic mRNA expression profiles of the ancestor, the evolved line, and the transductant bearing a known beneficial mutation to determine quantitatively which aspects of the resulting phenotypic expression profile can be attributed to that one mutation as opposed to others that might have been fixed in the population during the evolution experiment.

2. Modeling evolution of cold tolerance proteins

Central goals of the experiments described in Objectives 1 and 2 above are to identify proteins with roles in cold tolerance and to determine their modes of action. As these proteins are defined, we will also want to understand and model their evolution. How did the proteins evolve their capacities? What are the intermediary steps in the evolutionary process? How do we define the evolutionary changes in terms of the physical chemical characteristics of the proteins and their constituent amino acids? How do we identify when, in the evolutionary process, changes in selective pressure occurred? And can we use these models to provide important information about how the proteins function and to identify other putative “cold tolerance” proteins?

Every evolutionary analysis starts with a model of how sequences change. The standard approach towards modeling natural site substitutions in proteins is with a substitution matrix, a 20 X 20 array that represents the probability of any given amino acid changing to any other in a given length of evolutionary time. These substitution models directly encode the selective pressure acting on the proteins by representing which substitutions are likely or unlikely. Most methods for deriving these matrices use the approach developed by Dayhoff, based on an analysis of corresponding amino acids in pairs of closely-related homologous proteins (Dayhoff and Eck, 1968). Given such an evolutionary model and a set of putative homologs, it is possible to recreate the most probable phylogenetic relationship of the proteins using standard computer software. However, these standard substitution models are based on a number of highly problematic assumptions particularly poorly suited for the work described in this proposal. Generally researchers use matrices derived on a broad range of proteins, the vast majority of which are from mesophilic organisms. It is not clear how well suited these models are for representing the evolutionary changes of proteins from cold adapted bacteria. It is difficult to prepare more specialized models for the permafrost bacteria due to the large number of adjustable parameters and the limited amount of available data. Most importantly, the models assume that the rates of substitutions are the same in all locations in all proteins for all time, effectively assuming that all regions of the proteins are under constant homogeneous selective pressure. Obviously this assumption is inconsistent with exploring variations and changes in the selective pressure resulting from adaptation to cold environments.

For the past few years, the Goldstein lab has been working to create models of the evolutionary process that overcome the limitations of the current model. Initial work involved constructing substitution matrices specific for different types of local protein structure (Koshi and Goldstein, 1995). While this approach includes some aspects of structural heterogeneity in the evolutionary model, it is still based on the assumption that *all* positions in a particular local environment have the same fitness requirements and thus similar substitution rates. More recently, a set of more general probabilistic models have been developed (Koshi et al., 1997; Koshi and Goldstein, 1998) and used in phylogenetic analyses (Koshi et al., 1999). In this approach, rather than creating one "one-size-fits-all" model for all locations in the protein for all time, site heterogeneity is explicitly included in the model. Different *types* of locations, termed "*site classes*," are considered, each with its distinct fitness function and corresponding substitution matrix. While in principle these different site classes might correspond to locations with different secondary structures or functional significance, the nature of these site classes do not need to be define *a priori*. Nor does one need to assign different locations in the protein to specific site classes. Rather, there are adjustable probabilities that any location can be described by each site class. These probabilities are optimized simultaneously with the adjustment of the underlying fitness functions corresponding to the different site classes in order to best fit the observed sequences. Thus, one can *a posteriori* assign locations, at least probabilistically, to the various site classes and interpret the corresponding substitution matrices in terms of the selective pressure acting on each of those locations (Koshi and Goldstein, 1997). This approach achieved impressive gains in the ability to model the evolutionary process, compared with the more traditional single-model method.

In order to use this approach with limited data-sets (such as proteins from specific sets of organisms) we must greatly reduce the number of adjustable parameters. Rather than express the substitution rate as a function of the identity of the amino acids (*e.g.* the rate of alanine being substituted for glycine), the rates are expressed as functions of the corresponding physical-chemical properties of these residues (*e.g.* the rate of an amino acid with a given hydrophobicity, size, and charge being substituted for another amino acid of different hydrophobicity, size, and charge). Besides making the models more computationally tractable, it also allows one to "*read off*" the selective pressure at each location in the protein (Koshi and Goldstein, in press). It has been demonstrated how these methods can be applied to specific groups of proteins, such as hypervariable and constant regions of antibodies (Koshi and Goldstein, 1998) and to thermophilic proteins (manuscript in preparation).

A straightforward use of these models is in phylogenetic analysis. Analysis of protein sequences (in comparison with DNA sequences) allows us to "see" further back in evolutionary time, due to the higher degree of conservation at the amino acid level. In addition, the ability to model protein evolution on the amino acid level allows us to more accurately compute protein phylogenetics, in order that the resulting evolutionary trees can be better compared with the organism phylogenetics. In this way, we hope to determine the origin of the cold tolerance proteins that we identify in our structural and functional genomic studies, regardless of whether they arose through simple adaptation of pre-existing proteins, through gene-duplication, or through horizontal transfer. In addition, we will use the models to recreate ancestral protein sequences and examine specific changes in the sequences during particular periods of the evolutionary history of the protein (Koshi and Goldstein, 1996). This analysis will include expressing the proposed ancestral protein sequences in the laboratory and testing their specific biochemical properties. For instance, if we find that there are antifreeze proteins in the permafrost bacteria, we would model their evolution and test the intermediates for the activities associated with these proteins. Such experiments not only have the potential to reveal how a given antifreeze activity evolved, but has the potential to shed new light on the interrelationships between protein structure and the various activities observed in antifreeze proteins, a fundamental issue regarding these proteins. In these experiments, the "ancestral genes" will be created using standard molecular genetic/biology approaches, cloned into appropriate expression vectors and expressed at high levels in the host organism. The Thomashow lab has extensive experience making recombinant gene constructs and conducting protein overexpression experiments (e.g., Thomashow et al, 1986; Gilmour et al., 1996; Stockinger et al., 1997; Jaglo-Ottosen et al., 1998).

Most current evolutionary models assume constant selective pressure during the entire modeled evolutionary process. If we are interested in studying how proteins respond to *changes* in selective pressure resulting from the need to adapt to cold conditions, we need to have evolutionary models that include the possibility of such changes, especially if we want to explicitly include both psychrotrophs and non-psychrotrophs in the analysis. In this way we can directly separate sequence modifications due to cold adaptation from other random changes, as well as identify when in the evolutionary process the cold adaptation occurred. The models currently being used in the Goldstein laboratory assume that the selective pressure is constant during the evolutionary process. Our "site class" approach is, however, ideally suited to relaxing this constraint. We are now making the simple modifications necessary in order to explicitly

consider different substitution models at different periods of evolutionary time, providing information regarding these changes in selective pressure. Comparisons of protein and DNA evolution can also allow us to identify adaptive evolution from the background of random substitutions, using such conventional methods as comparing synonymous and non-synonymous DNA substitution rates.

There is already a set of proteins that have been identified to have antifreeze activity. Insight into how these proteins function can be gained by comparing their sequences and evolutionary patterns with related proteins (homologs) in other organisms. We will initially use standard methods of homology detection such as FASTA (Pearson, 1991) and PSI-BLAST (Altschul et al., 1997). Additional homologs may be detected by creating more specific statistical models for these particular proteins, including developing substitution models optimized for homolog detection for this particular set (Kann and Goldstein, in press) and by using more advanced methods such as Hidden Markov Models (HMMs) (Krogh et al., 1994). It may be possible to utilize the evolutionary approaches described above for homology detection; for instance, we can possibly improve our sensitivity by searching for homologs to our reconstructed ancestral sequences. As described in the background section of the proposal, proteins having very different amino acid sequence and structures can have similar antifreeze activity. In addition to finding homologous proteins, we are also interested in searching the psychrotrophic bacteria sequences for *analogous* proteins, proteins with similar antifreeze activity but with no evolutionary, sequence, or structural relationship. While these proteins may have no evolutionary relationship with the known proteins, it is likely that there would be similarities and regularities in the selective pressure acting on all of these various proteins. Using the methods described above, it may be possible to characterize the types of selective pressure acting on antifreeze proteins, and to use these characterizations to identify putative members of this set.

Finally, the Goldstein laboratory has developed a number of different approaches towards incorporating evolutionary information in the prediction of protein properties (Thompson and Goldstein, 1996a; Thompson and Goldstein, 1996b; Koshi and Goldstein, 1997; Thompson and Goldstein, 1997). While up until now they have concentrated on the prediction of local structure, the methods are robust and can be extended to prediction of other local attributes.

Objective IV. "Field Truth" Studies: assessing the importance of identified cold adaptive traits to success in the field and development of potential signatures for extraterrestrial life.

The first goal of these studies will be to bring the laboratory findings from above full circle, namely to test whether the cold adaptive traits that we identify have significant roles in the field. The second goal is to determine whether these traits can be converted to signatures of life useful for detecting extraterrestrial life. Toward these ends, we propose to address the following five questions.

1. Are cold adaptive physiological traits present in permafrost bacteria found in related bacteria from warm environments?

Our rationale is that the organisms we have isolated carry traits that have been selected for growth and survival in the permafrost habitat, and that these traits will not be as frequent in isolates from constantly warm soil environments. To test this hypothesis, we propose to compare various physiological traits of the permafrost strains with nearest relatives isolated from Puerto Rico. Puerto Rico occupies a tropical oceanic position at latitude 23.5°N, just south of the Tropic of Cancer. Persistent northeast trade winds maintain a uniform climate during the year with average temperatures ranging from 24 to 27°C. The island was formed by volcanic activity during the Lowest Cretaceous, approximately 120 million years ago, although the oldest rocks located in Southwestern Puerto Rico were created during the Jurassic era, approximately 195 million years ago (Jolly et al., 1998). **We have chosen to use tropical isolates for comparisons rather than temperate isolates to avoid any seasonal selection that would occur in temperate winters.**

The environmental temperatures prevailing in Puerto Rico and the permafrost are radically different. However, there are potential parallels regarding water availability. Precipitation is heavy in eastern Puerto Rico because the trade winds arrive at this side of the island first, leaving behind an average rainfall of more than 510 cm per year. Puerto Rico has an inner mountain range, the Cordillera Central, which covers the island from east to west, and determines the rain pattern observed in different parts of the island. Most of the residual rain remains in northern Puerto Rico with an average of 150 cm rain per year. Leeward of the Cordillera, rainfall is much less, averaging 0.9 cm per annum on the south coast and less than 0.8 cm for southwestern Puerto Rico (htpp://www.fs.fed.us/colorimagemap/images/ m411.html). Southwestern Puerto Rico has an average temperature of 30°C and the evaporation rate is 2.5 times faster than the annual precipitation rate, making this side of the island one of the most arid ecosystems observed in the tropics (Jolly et al., 1988). Thus, it is possible that genetic mechanisms important for dehydration tolerance in permafrost bacteria may have counterparts in bacterial isolates from southwestern Puerto Rico.

Our tropical sampling sites will be located in different sites in Puerto Rico. Some samples will be taken from the hot and moist soils that the northern side of Puerto Rico (no cold or desiccation stress) while other samples will be taken from the southern part of the island where hot and dry conditions prevail (no cold stress). In order to avoid biases introduced by sample-specific conditions, three samples of each soil type will be analyzed, each consisting of a composite of several representative samples. Bacteria will be isolated from these soil composites following the protocols used to isolate bacteria from the permafrost sites (Vishnivetskaya et al., 2000). Two incubation temperatures will be used, 27 and 4°C. The 16S rDNA gene of these organisms, as well as from DNA isolated directly from the soil samples, will be amplified by PCR using primers designed to amplify the 16S rDNA gene of bacteria isolated from Siberian and Antarctic sites. A phylogenetic tree will be constructed between the Puerto Rican and the polar isolates in order to determine the relatedness of these bacteria. Phylogenetically close bacteria isolated from Puerto Rican and polar soils will be compared in terms of their metabolic capabilities relating to cold and dehydration tolerance including: growth rates at 20, 10, 0 and -2.5°C; long-term freeze survival and freeze-thaw survival; synthesis of compatible solutes and alterations in membrane composition in response to low temperature and low water activity.

2. Are the taxa that appear uniquely cold distributed in fact found principally in polar habitats?

It is striking that four of the six major taxa isolated from the Siberian permafrost have strains from the Antarctic or other cold habitats as their closest known phylogenetic relatives. This suggests the selection for more than a few traits to realize this pattern. This correlation has not been thoroughly tested however since the postulated polar-unique taxa have not been directly searched for in warm environments. We will attempt to isolate from Puerto Rican soils taxa that are postulated as uniquely polar by using the enrichment conditions that were used for obtaining the permafrost isolates. Since this method may not be highly selective, we will also develop primers for the 16S rDNA of the permafrost clades and use these to for a more sensitive detection of these taxa in tropical soil DNA. Since several of the polar groups form a monophyletic lineage, it should be possible to identify clade specific primers. Any PCR products recovered will be cloned and sequenced to determine whether the clone belongs to the clade. If two primers can be successfully identified, we will try to identify a third probe/primer region so that we can quantify rDNA for these clades using quantitative (realtime PCR). We have successfully developed, verified and used this method for quantifying another (*Rhodococcus)* clade (Rodrigues, Aiello and Tiedje, unpublished). With this method we can also quantify how dominant the Siberian clades are in permafrost DNA, and thus have a better indication of the success of these groups in polar regions.

3. Are the genetic traits identified in the above objectives as likely important for cold adaptation found in the polar isolates but absent or rare in tropical strains?

If the traits are uniquely selected during the permafrost organisms' long term polar life, then these traits would not be expected to be maintained in the tropical bacteria. These traits will be screened for by using genes identified from the above studies to probe the DNA of the tropical isolates. If hybridization is found, we will identify primers from the genome sequence to attempt to amplify segments for sequencing to see if the nucleotide sequence is in fact similar. We can also screen for the particular expressed protein using the HPLC MALDI-TOF mass spectrometry protocal that will have been worked out under Objective 2 for these proteins. We will initially focus this test on the Arthrobacter group since we expect to be able to isolate tropical relatives of the Siberian strains. If question #2 above does not yield any tropical relatives of the postulated polar-unique taxa, then we will complete this test with other tropical isolates we recover.

4. Do the genetic traits identified in the above objective actually confer improved survival or growth under polar conditions?

This is an important test of the information gleaned from the genomic and proteomic studies. We will determine whether mutants with the target traits deleted have reduced survival or growth in simulated polar conditions, in other space-like conditions, and after several years when returned to their Siberian permafrost site. Similarly, we will add the trait to wild-type strains that lack the trait to see if their survival or growth is improved under the polar conditions. Under this objective we hope to cooperate with other NAI members who have facilities to simulate other planetary or space environments. We will return the mutant strains to the site in

microcosms perpared with sterile soil or in microcosms with genetically marked strains. Survival and growth will be measured by plant counts.

5. Can the cryo-adaptive traits be detected in permafrost samples?

If the above tests have identified traits important to microbial survival in permafrost, then we should be able to directly test for those traits in polar samples. We would initially attempt to devise specific primers for conserved regions of such genes since that is currently the easiest, most sensitive and specific means to detect particular genetic traits in environmental samples. Furthermore, the trait can be quantified, in principle, to one gene copy per DNA sample. But, the test does not have to be limited to PCR and DNA. Other chemical biomarkers that could be assayed are specific proteins, membrane components or electron carriers. If some taxa are proven polar-unique, there are likely a variety of chemical differences distinctive for these groups that could become potential assays. We will initially compare the permafrost and tropical sites to determine the presence (and perhaps quantity) of these cryo-traits in the two environments.

If cryo-markers (or, at a later stage, other markers for extraterrestrial features), are identified and validated as above, we will explore whether they can be automated and miniaturized for extraterrestrial life detection. We would intend to do this in cooperation with another MAI group who has expertise in approaches to reducing the idea to practice. We would bring to the team the biological ideas and approaches but the planetary chemistry, experience with feasible approaches for non-earth measurements, device engineering would need to be provided by others. Dr. McGrath, a member of our Center, is a mechanical engineer so he can at least provide us with beginning considerations on what might be feasible.

Our ultimate goal is to mine genomic and proteomic information to better understand how microbial life on earth adapts to cold temperatures using the ancient permafrost microbes as the model, and to mine that information for approaches to search for and help characterize any life in extraterrestrial environments.

D. Expertise and responsibilities of Center Co-investigators

1. Michigan State University

Michael Thomashow will be responsible for integrating the structural and function genomics studies and will directly supervise the gene conservation and gene expression microarray analyses. He will also assist McGrath is the proteomic studies targeted at identifying candidate freezing tolerance proteins (antifreeze proteins and potential cryoprotective polypeptides). Dr. Thomashow is molecular geneticist who has worked on both plant and bacterial systems and has extensive experience working on the cold acclimation response in plants with emphasis on the regulation and function of genes induced in response to low temperature.

James Tiedje will be responsible for the permafrost culture collection, the physiological characterization of these strains, development and use of probes to detect cryo-adapted strains

and genes, and for coordination of the work on Field Truth. Tiedje will also assist Thomashow in array expression studies. Tiedje has expertise in microbial ecology, including use of molecular and array technologies for understanding microbial communities and their activities.

Richard Lenski will supervise one of the three new evolution experiments described in this proposal (adaptation following a sudden shift to 4°C). He will also supervise some of the genetic analyses of the derived lines from the evolution experiments, including transposon tagging of selected mutations and genomic mRNA expression profiles. All of these techniques are currently being performed successfully in his laboratory. Lenski will collaborate closely with Bennett (University of California, Irvine) in all aspects of the design and analysis of the evolution experiments. Lenski is a leader in the field of evolutionary biology and is a pioneer of experimental evolution studies.

John McGrath will supervise, in close collaboration with Thomashow, the proteomic studies targeted at identifying and determining the specific activities of antifreeze proteins and potential cryoprotective polypeptides. In addition, he will lead the efforts in applying HPLC methods to analyze compatible solutes. McGrath is recognized for his published work on thermodynamic and mathematical modeling of the responses of cells to freezing and thawing which provides a means of predicting the state of cells in the frozen state. He has been a leader in the development and application of instrumentation aimed at exposing cells to controlled low water activity and low temperature states and is currently President of the International Society for Cryobiology.

2. University of Michigan

David Lubman will lead the proteomic profiling studies and closely collaborate with Thomashow, McGrath and Kathariou on the purification of specific cold tolerance proteins. Lubman has developed a number of separation methods for screening proteins that include liquid phase separations of proteins, chromatographic separation, isoelectric focussing of proteins and various ion-exchange methods for separating proteins. In addition, he is involved in developing sequencing methodologies using a variety of capillary separation methods including capillary electrophoresis, capillary chromatography and capillary electrochromatography for separation of protein digests. He has also been involved in the development of new mass spec techniques for the characterization of proteins including ESI and MALDI and the development of new hybrid mass spectrometers for on-line sequencing.

Richard Goldstein will develop and implement the bioinformatics techniques for analyzing the protein sequence information and model the evolution of cold tolerance proteins. This work will involve delineating the phylogenetic relationships between the sequenced psychrotrophic proteins as well as those of related species, and using this information to understand the origin and evolution of the proteins responsible for cold adaptation. He will also use standard and novel statistical approaches to identify proteins that are homologous or analogous to “antifreeze” proteins. Goldstein has developed a number of novel methods for understanding proteins in their evolutionary context, as well as using evolutionary information to understand the properties of specific proteins. He also has experience in developing models of evolution that explicitly consider protein structure and amino acid properties.

3. University of California, Irvine

Albert Bennett will supervise two of the three newly proposed evolution experiments with laboratory populations of the bacterium *E. coli* (adaptation to progressively lower temperatures, and adaptation to freeze/thaw cycles). He will also supervise the mapping of the thermal niches of these derived lineages and their fitness changes. He will be responsible for conducting the genomic analyses of these and previously derived lineages using high-density arrays. All of these techniques are currently being undertaken successfully in his laboratory. Bennett will work closely with Lenski (Michigan State University) in all aspects of the design and analysis of the evolution experiments. Bennett is a leader in the field of ecology and evolutionary biology.

4. North Carolina State University

Sophia Kathariou will supervise the genetic experiments directed at identifying genes with key roles in cold adaptation. Karhariou is an expert in microbial genetics and for a number of years has been studying low temperature gene expression and cold-essential genes in the food-associated psychrotropic bacterium (and pathogen) Listeria monocytogenes. Kathariou spent a sabbatical at MSU in 1997-1998 and worked with McGrath, Tiedje and Gilichinsky on the isolation and characterization of cold tolerance of the Arctic permafrost isolates.

5. University of Puerto Rico-Mayaguez

Lycely Sepulveda is the new microbial genetics faculty member at UPR-Mayaguez. She has expertise in cloning sequencing and characterizing functional genes in environmentally important organisms, and is well trained in microbial ecology. She will be responsible for isolating tropical organisms for comparison to the permafrost strains, their 16S rDNA sequencing, measurement of some of the cryo adaptive traits, and to coordinate with Tiedje the DNA-based methods for organism and trait detection.

6. Joint Genome Institute (JGI)

The JGI is one of the five major international high-throughput genome centers that participated in completing the human genome. They have also completed or are currently completing several microbial genomes, e.g. *Nitrosomonas, Prochlorococcus, Rhodobacter.* The JGI has high capacity and high quality product at cost-effective rates. In this project they are responsible for delivering approximately 90 million base pairs of sequence, assembly as appropriate for the level of coverage proposed, and the standard annotation. We are pleased that their science leaders (Brauscomb, Hawkins and Predki) have joined the team since their interest in these novel organisms, advances in technology expected during this project, and suggestions and interpretations for the array and proteomics phase, will be important to the research group.

II. Training

Our educational objective is to provide in-depth training in structural and functional genomics, including both an overall understanding and knowledge of experimental approaches, but to encompass that training in a context relevant to astrobiology. In this way our students will have cutting-edge knowledge for the genomics era, but will be able to use that knowledge to explore questions important for understanding the possible strategies of extraterrestrial life. We will implement this strategy by having students major in their existing (strong) degree-granting programs, but to establish an "area of specialization" in astrobiology. In this way students will not be diluted in their training in their basic discipline but will gain the perspective and knowledge needed to make significant contributions to the field of astrobiology. The specialization approaches also makes it possible for students from a variety of departments to enter the program and gain an astrobiology emphasis.

A. Educational program

The educational program for the "astrobiology specialization" will consist of the following:

1. A new graduate level course at MSU in "Genomics of Environmental Stress Tolerance." The course will cover the latest developments in our understanding of structural and functional genomics in prokaryotes, eukaryotes and archea; the experimental approaches and methodologies used in such studies including fundamental bioinformatics; and applications of the science to understanding the systems organisms have evolved to tolerate environmental stresses such as extremes in temperature, dehydration and desiccation, and UV-light and other forms of radiation. The course will be taught by Thomashow and other faculty at MSU and will include guest lectures from other Center investigators. The lecture outlines, references and other materials related to the course will be made available on the Web for students at the other Centers, members of the Astrobiology Institute and interested public

2. Short-courses to augment keys areas of student training. We have two courses in mind initially. One by Dr. Gilichinsky during his visits to MSU that is focused on the cryosoil environment and history of the Arctic and Antarctic and drawing similarities and dissimilarities to Europa and Mars. The second would be in informatics and data analysis relevant to genomic studies. This would help our students and post-docs get an early start on more in-depth research.

3. Access NAI internet courses of relevance to our theme. There are many institutions in NAI that have existing excellence in relevant planetary, space and geochemical sciences. It makes sense for our students to access that excellence instead of trying to offer this portion of the cross-training at our Center. In return, our Center will offer its expertise in genomics and evolution to other Centers that lack this expertise. We would expect our students with this specialization to take one such course provided that an appropriate quality course exists

4. Other relevant courses. To provide the breadth of knowledge for the specialization, MSU students would be required to take at least 2 of the following courses: Diversity of Prokaryotes (MIC 827); Advanced Microbial Ecology (MIC 829); Molecular Evolution: Principles and Techniques (ZOL 855); Quantitative Methods in Ecology and Evolution (ZOL

851); Evolution (ZOL 445); and Environmental Geochemistry (GLG 421). Students not at MSU would take courses of similar subject matter. These courses provide knowledge of the extant microbial physiologies, microbial interactions, molecular and deduced evolution and the geochemical environment of the lithosphere, all topics important to astrobiology.

Because our Center is distributed among institutions, we will use the following operating principles for students outside MSU. Those students and postdocs will participate in the Center's annual science retreat. Second, they will enroll in courses with content similar to the above at their institutions. Students at RAS and at UPR will spend one summer or similar time at MSU for a phase of their research. Students at UM, because of proximity, will participate in some of the MSU components depending on their relevance to their program.

B. Internal seminars

Three times each semester all Center researchers will hear half-hour seminar presentations on the progress and plans of several of the graduate students and post-docs. With this schedule, each researcher would present once per year. This mechanisms keeps everyone informed of the entire research spectrum, and provides both an incentive and check-point for progress and quality. One session per year would be at UM. Since MSU and UM are 1 hour drive apart, this activity is feasible for members at both institutions.

C. Research exchanges

We have budgeted funds for graduate students or postdocs to go to another institution either within our Center or within the NAI to conduct a collaborative experiment that draws on the unique resources of the partners. We can envision that some of our mutant strains with postulated special survival properties could be tested in space/planetary-like conditions available at other NAI Centers.

D. Symposium on "Genomics and Evolution of Extremophiles"

In year three or four, we would host an open symposium in which our results are summarized along with presentations by relevant invited speakers. The purpose is to bring together the cutting-edge body of knowledge for Center members, NAI members and interested peers. Our intent is to stimulate new ideas for the next phase our experimentation both with the Center and with other NAI members. At this stage in our research we should have several genomes well analyzed, a good idea of at least postulated evolutionary strategies, and hopefully some genes and their proteins identified that have key roles in life at low temperature.

The above program we believe provides for the appropriate cross-training and provides the astrobiology context so that the power of genomics, proteomics and evolutionary mechanisms can be brought to bear on astrobiology.

III. Management

Professor Michael Thomashow will be the PI and Director of our Center. Professor James Tiedje will serve as Associate Director and aid the Director in duties determined by the Director. We have found such a management scheme to work well in our NSF-STC Center for Microbial Ecology; one person has point responsibility, but has a second person fully cognizant of the breadth of Center activities with whom to discuss pertinent issues and implement actions.

An Executive Committee of three will be named from among the Center co-investigators that will establish operating policy and aid in making major decisions. The membership of the Executive Committee will rotate. The Executive Committee will make sure that all perspectives are heard and help ensure that the science conducted is cutting-edge in nature.

A significant challenge for any large integrated research effort is to ensure that members know each other, their expertise and their project goals, and to achieve this without burdensome requirements that sap enthusiasm. We propose to achieve this by our educational and training program (above) and the following activities.

A. Center Programs

1. Annual Retreat. We will have an annual retreat for all Center members, i.e. graduate students, post-docs and faculty, in which all work is presented as posters or oral talks and working groups meet to discuss plans in subtopics areas. Especially in the first years members need to get to know others on a personal basis so that anyone, even new students, know who has what expertise and feel comfortable in directly accessing that expertise. The Center for Microbial Ecology was able to achieve this level of familiarity and much success and efficiency derived when students perceived opportunities or solutions via others and acted on them directly. In a dispersed Center, it is money well spent to bring all members together once per year (at least in initial years) to review and plan the science.

2. External Seminar Program. We will have a seminar program coordinated and co-sponsored with the Center for Microbial Ecology and relevant departments. The speakers will be asked to stay long enough to meet with students and post-docs and interested faculty. In the first year we will emphasize inviting key scientists at other NAI institutions so that we can gain a better idea of possible research or educational collaborations.

3. Web-based genomics database. As detailed above, we propose to enhance our Web-based microbial informatics database to store and make accessible to center scientists the genome, proteome and expression data that is the core of this project. Our design serves researchers well at any partner institution, and the information is immediately available to all other members. We envision this tool to be at the heart of our research interactions.

4. External Advisors. We will name a small group of external advisors for the purpose of providing Center management with critique and suggestions on Center activities, progress and direction. We would likely invite these advisors to our annual retreat. After year two, we would

particularly ask them for advice on mid-course adjustments so that these could be implemented before the third year NASA review.

Our operating philosophy, which is consistent with the cooperative agreement mode of funding, will be to adjust our project funding in subsequent years based on our recent discoveries, relevant external advice, effectiveness of members and the needs of the NAI. We would use information from the retreat and planning sessions, the external advisors and the NAI management in making these decisions. Because of this operating principle the subsequent year budgets in this proposal are subject to change.

B. Minority program

The Center includes the University of Puerto Rico (UPR), Mayaguez, as one of the partner institutions for three reasons. First, one important test of genes hypothesized as important to cold adaptation is whether they are uniquely selected in polar regions or whether they also exist in ancient tropical environments. Hence the Center research need matches the involvement of UPR. Secondly, UPR students, because of their tropical ecosystem awareness, are attracted to other environmental research problems. Hence, our UPR collaborator believes that recruiting minority students to astrobiology research should be successful. Finally CME has built a strong relationship over a 15 year period with the three UPR campuses that teach microbiology. We draw on our linkages and mutual program experience in what is proposed here. We believe our past program has been successful. More the 30 UPR undergrads have done summer research internships at the CME and most have gone on to science careers. Four have received PhDs (one NIH and one NSF Fellow) under the CME program and four more are at various stages of their PhD program. We believe the key to minority progress, at least our sciences, lies in establishing and maintaining relationships with minority institutions; building a reputation where the minority students knows previous students have been successful; and listening to the minority institutions programmatic ideas since they know their students well. We, therefore, propose the following:

a) Summer internships at the majority institutions for undergraduate and graduate students to experience research in the former case and in the latter to learn particular skills for the thesis research.
b) Focus on a genomics/astrobiology theme for the molecular microbial ecology workshops offered annually by CME and UPR faculty to UPR microbiology majors on the three campuses. This is a two day intensive hands on laboratory during their spring vacation and rotated among the campuses. These have been oversubscribed and extremely popular. We propose that our Astrobiology Center takes over organizing two of these in the five year period.
c) Develop an astrobiology module for undergraduate microbial diversity course at UPR. In this manner advanced students who already have a budding interest in the diversity (including extremophiles) of microbes would be introduced to the special conditions and opportunities of astrobiology. This same module could also be used in MSU senior level microbial ecology course to interest students in the field. Since this course is spring term, students could also be attracted to joining Center research projects in the summer.

References

Abyzov, S.S., N. Ye Bobin, B.B. Kudryashov. 1987. Anabiosis of microorganisms in the Antarctic ice sheet. *In*: Ye Beker M. (ed) Anabiosis of cells. Academy of Sciences of Latvian SSR. A. Kirkhenshteyn Institute of Microbiology. Riga, pp: 43-54.

Adams, J. and P.W. Oeller. 1986. Structure of evolving populations of *Saccharomyces cerevisiae*: adaptive changes are frequently associated with sequence alterations involving mobile elements belonging to the Ty family. Proc. Natl. Acad. Sci. USA 83:7124-7127.

Afflerbach, H., O. Schröder, and R. Wagner. 1998. Effects of the *Escherichia coli* DNA-binding protein H-NS on rRNA synthesis *in vivo*. Mol. Microbiol. 28:641-653.

Aguilar, P.S., J.E. Cronan, JR., and D. de Mendoza. 1998. A *Bacillus subtilis* gene induced by cold shock encodes a membrane phospholipid desaturase. J. Bacteriol. 180:2194-2200.

Anchordoguy, T.J., A.S. Rudolph, J.F. Carpenter, and J.H. Crowe. 1987. Modes of interaction of cryoprotectants with membrane phospholipids during freezing. Cryobiology 24:324-331.

Annous, B.A., L.A. Becker, D.O. Bayles, D.P. Labeda, and B.J. Wilkinson. 1997. Critical role of anteiso-C15:0 fatty acid in the growth of *Listeria monocytogenes* at low temperatures. Appl. Environ. Microbiol. 63:3887-3894.

Antikainen, M. and M. Griffith. 1997. Antifreeze protein accumulation in freezing-tolerance cereals. Physiol. Plant 99:423-432.

Appenzeller, T. 1999. Test tube evolution catches time in a bottle. Science 284:2108-2110.

Artus, N.N., M. Uemura, P.L. Steponkus, S.J. Gilmour, C.T. Lin, and M.F. Thomashow. 1996. Constitutive expression of the cold-regulated *Arabidopsis thaliana COR15a* gene affects both chloroplast and protoplast freezing tolerance. Proc. Natl. Acad. Sci. USA 93:13404-13409.

Atlung, T. and H. Ingmer. 1997. H-NS: a modulator of environmentally regulated gene expression. Mol. Microbiol. 24:7-17.

Avise, J.C. 1994. Molecular Markers, Natural History and Evolution. Chapman & Hall, New York.

Bae, W., P.G. Jones, and M. Inouye. 1997. CspA, the major cold shock protein of *Escherichia coli*, negatively regulates its own gene expression. J. Bacteriol. 179:7081.

Baker, S.S., K.S. Wilhelm, and M.F. Thomashow. 1994. The 5'-region of *Arabidopsis thaliana cor15a* has *cis*-acting elements that confer cold-, drought- and ABA-regulated gene expression. Plant Mol. Biol. 24:701-713.

Barder, T.J., P.J. Wohlman, C. Thrall, and P.D. DuBois. 1997. Fast chromatography and nonporous silica. LC GC-Mag. 15:918.

Bayles, D.O., B.A. Annous, and B.J. Wilkinson. 1996. Cold stress proteins induced in *Listeria monocytogenes* in response to temperature downshock and growth at low temperatures. Appl. Environ. Microbiol. 62:1116-1119.

Becker, R.E., C.M. Volkmann. 1961. A preliminary report on the bacteriology of permafrost in the Fairbanks area. Proc. Alaskan Sci. Conf. 12:188.

Bennett, A.F., K.M. Dao, and R.E. Lenski. 1990. Rapid evolution in response to high temperature selection. Nature 346:79-81.

Bennett, A.F., R.E. Lenski, and J.E. Mittler. 1992. Evolutionary adaptation to temperature. I. Fitness responses of *Escherichia coli* to changes in its thermal environment. Evolution 46:16-30.

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

Bennett, A.F. and R.E. Lenski. 1996. Evolutionary adaptation to temperature. V. Adaptive mechanisms and correlated responses in experimental lines of *Escherichia coli*. Evolution 50:493-503.

Bennett, A.F. and R.E. Lenski. 1997. Evolutionary adaptation to temperature. VI. Phenotypic acclimation and its evolution in *Escherichia coli*. Evolution 51:36-44.

Bennett, A.F. and R.E. Lenski. 1999. Experimental evolution and its role in evolutionary physiology. American Zoologist 39:346-362.

Berger, F., N. Morellet, F. Menu, and P. Potier. 1996. Cold shock and cold acclimation proteins in the psychrotrophic bacterium *Arthrobacter globiformis* SI55. J. Bacteriol. 178:2999.

Berger, F., P. Normand, and P. Potier. 1997. *CapA*, a *cspA*-like gene that encodes a cold acclimation protein in the psychrotrophic bacterium *Arthrobacter globiformis* SI55. J. Bacteriol. 179:5670-5676.

Bloem, J., M. Veninga, and J. Shepherd. 1995. Fully automatic determination of soil bacterium numbers, cell volumes, and frequencies of dividing cells by confocal laser scanning microscopy and image analysis. Appl. Environ. Microbiol. 61:926-936.

Bohannan, B.J.M. and R.E. Lenski. 2000. Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. Ecology Letters 3:362-377.

Boivin-Jahns, V., A. Bianchi, R. Ruimy, J. Carcin, S. Daumas, and R. Christen. 1995. Comparison of phenotypical and molecular methods for the identification of bacterial strains isolated from a deep subsurface environment. Appl. Environ. Microbiol. 6:3400-3406.

Bouma, J.E. and R.E. Lenski. 1988. Evolution of a bacteria/plasmid association. Nature 335:351-352.

Bowmann, J.P., S.A. McCammon, M.V. Brown, D.S. Nichols, and T.A. Mcmeekin. 1997. Diversity and association of psychrophilic bacteria in Antarctic sea ice. Appl. Environ. Microbiol. 63:3068-3078.

Bull, J.J. and I.J. Molineux. 1992. Molecular genetics of adaptation in an experimental model of cooperation. Evolution 46:882-895.

Bull, J.J., M.R. Badgett, H.A. Wichman, J.P. Huelsenbeck, D.M. Hillis, A. Gulati, C. Ho, and I.J. Molineux 1997. Exceptional convergent evolution in a virus. Genetics 147:1497-1507.

Bull, J.J., M.R. Badgett, and H.A. Wichman. 2000. Big-benefit mutations in a bacteriophage inhibited with heat. Mol. Biol. Evol. 17:942-950.

Burch, C.L. and L. Chao. 1999. Evolution by small steps and rugged landscapes in the RNA virus ϕ6. Genetics 151:921-927.

Burch, C.L. and L. Chao. 2000. Evolvability of an RNA virus is determined by its mutational neighbourhood. Nature 406:625-628.

Cameron, R.E. and F.A. Morelli. 1974. Viable microorganisms from ancient Ross Island and Tayler Valley drill core. Antarctic J. of the U.S. 9:113-116.

Carpenter, J.F. and J.H. Crowe. 1988. The mechanism of cryoprotection of proteins by solutes. Cryobiology 25:244-255

Cavaignac, S.M., S.J. White, G.W. de Lisle, and D.M. Collins. 2000. Construction and screening of *Mycobacterium paratuberculosis* insertional mutant libraries. Arch. Microbiol. 173:229-231.

Chakrabartty, A. and C.L. Hew. 1991. The effect of enhanced α-helicity on the activity of winter flounder antifreeze peptide. Eur. J. Biochem. 202:1057-1063.
Chamot, D., W.C. Magee, E. Yu, and G.W. Owttrim. 1999. A cold shock-induced cyanobacterial RNA helicase. J. Bacteriol. 181:1728-1732.
Chao, L., B.R. Levin, and F.M. Stewart. 1977. A complex community in a simple habitat: an experimental study with bacteria and phage. Ecology 58:369-378.
Chao, L., C. Vargas, B.B. Spear, and E.C. Cox. 1983. Transposable elements as mutator genes in evolution. Nature 303:633-635.
Chao, L. 1990. Fitness of RNA virus decreased by Muller's ratchet. Nature 348:454-455.
Chao, H., L.L. Davies, and J.F. Carpenter. 1996. Effects of antifreeze proteins on red blood cell survival during cryopreservation. J. Exp. Biol. 199:2071-2076.
Chen, Y.J., X.Y. Jin, D. Misek, R. Hinderer, S.H. Hanash, and D.M. Lubman. 1999. Identification of proteins from two-dimensional gel electrophoresis of human erythroleukemia cells using capillary high performance liquid chromatography/electrospray-ion trap-reflectron time-of-flight mass spectrometry with two-dimensional topographic map analysis of in-gel tryptic digest products. Rapid Commun. Mass Spectrom. 13:1907.
Clewell, D.B., S.E. Flannagan, and D.D. Jaworski. 1995. Unconstrained bacterial promiscuity: the Tn*916*-Tn*1545* family of conjugative ransposons. Trends Microbiol. 3:229-236
Close, T.J. 1997. Dehydrins: a commonality in the response of plants to dehydration and low temperature. Physiol Plant 100:291-296.
Cohan, F.M., E.C. King, and P. Zawadzki. 1994. Amelioration of the deleterious pleiotropic effects of an adaptive mutation in *Bacillus subtilis*. Evolution 48:81-95.
Cooper, V.S. 2000. Consequences of ecological specialization in long-term evolving populations of *Escherichia coli*. Ph.D. dissertation, Michigan State University, East Lansing.
Cooper, V.S. and R.E. Lenski. 2000. The population genetics of ecological specialization in evolving *E. coli* populations. Nature, in press.
Csonka, L.N. 1981. Proline over-production results in enhanced osmotolerance in *Salmonella typhimurium*. Mol. Gen. Genet. 182:82-86.
de Saizieu, A., C. Gardes, N. Flint, C. Wagner, M. Kamber, T.J. Mitchell, W. Keck, K.E. Amrein, and R. Lange. 2000. Microarray-based identification of a novel Streptococcus pneumoniae regulon controlled by an autoinduced peptide. J. Bacteriol. 182:4696-4703.
De Visser, J.A.G.M., C.W. Zeyl, P.J. Gerrish, J.L. Blanchard, and R.E. Lenski. 1999. Diminishing returns from mutation supply rate in asexual populations. Science 283:404-406.
Delille, D. 1992. Marine bacterioplancton at the Weddell Sea ice edge: distribution of psychrophilic and psychtrophic populations. Polar Biol. 12:205-210.
Delille, D. 1993. Seasonal changes in the abundance and composition of marine heterotrophic bacterial communities in an Antarctic coastal area. Polar Biol. 13:463-470.
Dersch, P., S. Kneip, and E. Bremer. 1994. The nucleoid-associated DNA-binding protein H-NS is required for the efficient adaptation of *Escherichia coli* K-12 to a cold environment. Mol. Gen. Genet. 245:255-259.
Devries, A.L. 1986. Antifreeze glycopeptides and peptides: interactions with ice and water. Meth. Enzymol. 127:293-303.

Dimmic, M.W. and R.A. Goldstein. 2000. Modeling evolution at the amino acid level using a general fitness model, *in* Pacific Symposium on Biocomputing 2000, Altman, Dunker, Hunger, Lauderdale, and Klein, eds., World Scientific, Singapore, pps. 18-29

Ducret, A., I. Van Oostveen, J.K. Eng, J.R. Yates, and R. Aebersold. 1998. High throughput protein characterization by automated reverse-phase chromatography electrospray tandem mass spectrometry. Protein Sci. 7:706.

Dure III, L. 1993. A repeating 11-mer amino acid motif and plant desiccation. Plant J. 3:363-369.

Dykhuizen, D. 1978. Selection for tryptophan auxotrophs of *Escherichia coli* in glucose-limited chemostats as a test of the energy conservation hypothesis of evolution. Evolution 32:125-150.

Dykhuizen, D.E. 1990. Experimental studies of natural selection in bacteria. Annual Reviews of Ecology and Systematics 21:378-398.

Elena, S.F., V.S. Cooper, and R.E. Lenski. 1996. Punctuated evolution caused by selection of rare beneficial mutations. Science 272:1802-1804.

Emerson, D. and C. L. Moyer. 1997. Isolation and characterization of novel lithotrophic iron-oxidizing bacteria that grow near neutral pH. Appl. Environ. Microbiol. 63:4784-4792.

Etchegaray, J.P. and M. Inouye. 1999. CspA, CspB, and CspG, major cold shock proteins of *Escherichia coli*, are induced at low temperature under conditions that completely block protein synthesis. J Bacteriol. 181:1827-1830.

Ewart, K.V., Q. Lin, and C.L. Hew. 1999. Structure, function and evolution of antifreeze protiens. Cell. Mol. Life Sci. 55:271-283.

Franzmann, P.D., P.P. Deprez, A.J. McGuire, T.A. McMeekin, and H.R. Burton. 1990. The heterotrophic bacterial microbiota of Burton Lake, Antarctica. Polar Biol. 10:261-264.

Franzmann, P.D. 1996. Examination of Antarctic prokaryotic diversity through molecular comparisons. Biodivers. Conserv. 5:1295-1306.

Friedman, E.I. 1982. Endolithic microorganisms in the Antarctic cold desert. Science 215:1045-1053.

Friedman, I. 1994. Permafrost as a microbial habitat. *In*: D. Gilichinsky (ed.) Viable Microorganisms in Permafrost, Russian Acad. Sci. pp. 21-26.

Gilichinsky, D.A., G.M. Khlebnikova, D.G. Zvyagintsev, D.C. Fedorov-Davydov, and N.N. Kudryatseva. 1988. The use of microbiological charcteristics of rocks in geocryology. Proc. Vth Int. Conf. Permafrost. Trondheim, Norway 1:749-753.

Gilichinsky, D.A., E.A. Vorobyova, L.G. Erokhina, D.G. Fedorov-Davydov, and N.R. Chaiovskaya. 1992. Long term preservation of microbial ecosystems in permafrost. Adv. Space Res. v. 12, N 4:225-263.

Gilichinsky, D. and S. Wagener. 1994. Microbial life in permafrost. *In*: D. Gilichinsky (ed.) Viable Microorganisms in Permafrost, Russian Acad. Sci. pp. 7-20.

Gilichinsky, D. and S. Wagener. 1995. Microbial Life in Permafrost, A Historical Review. Permafrost and Periglacial Processes. 5:243-250.

Gilmour, S.J., C. Lin, and M.F. Thomashow. 1996. Purification and properties of *Arabidopsis thaliana COR* (cold- regulated) gene polypeptides COR15am and COR6.6 expressed in *Escherichia coli*. Plant Physiol. 111:293-299.

Gilmour, S.J., D.G. Zarka, E.J. Stockinger, M.P. Salazar, J.M. Houghton, and M.F. Thomashow. 1998. Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced *COR* gene expression. Plant J. 16:433-442.

Goldstein J., N.S. Pollitt, and M. Inouye. 1990. Major cold shock protein of *Escherichia coli*. Proc. Natl. Acad. Sci. USA 87:283-287.

Gould, S.J. 1989. Wonderful Life: The Burgess Shale and the Nature of History. Norton, New York.

Gounot, A.-M. 1991. Bacterial life at low temperature: physiological aspects and biotechnological implications. J. Appl. Bacteriol. 71:386-397.

Goverde, R.L.J., J.H.J. Huis in't Veld, J.G. Kusters and F.R. Mooi. 1998. The psychrotrophic bacterium *Yersinia enterocolitica* requires expression of *pnp*, the gene for polynucleotide phosphorylase, for growth at low temperature (5°C). Mol. Microbiol. 28:555-569.

Graumann, P., K. Shcröder, R. Schmid, and M.A. Marahiel. 1996. Cold shock stress-induced proteins in *Bacillus subtilis*. J. Bacteriol. 178:4611-4619.

Graumann, P., T.M. Wendrich, M.H. Weber, K. Schroder, and M.A. Marahiel. 1997. A family of cold shock proteins in *Bacillus subtilis* is essential for cellular growth and for efficient protein synthesis at optimal and low temperatures. Mol. Microbiol. 25:741-756.

Graumann, P.L. and M.A. Marahiel. 1999. Cold shock proteins CspB and CspC are major stationary-phase-induced proteins in *Bacillus subtilis*. Arch. Microbiol. 171:135-138.

Graumann, P. and M.A. Marahiel. 1999. Some like it cold: Response of microorganisms to cold shock. Archives Microbiol. 171:135.

Griffith, M., M. Antikainen, W-C Hon, K. Pihakaski-Maunsbach, X-M Yu, Y.U. Chun, and S.C. Yang. 1997. Antifreeze proteins in winter rye. Physiol. Plant 100:327-332.

Grossmann, S. and G.S. Dieckmann. 1994. Bacterial standing stock, activity, and carbon production during formation and growth of sea ice in the Weddell Sea, Antarctica. Appl. Environ. Microbiol. 60:2746-2753.

Guy, C., D. Haskell, and Q.B. Li. 1998. Association of proteins with the stress 70 molecular chaperones at low temperature: evidence for the existence of cold labile proteins in spinach. Cryobiology 36:301-314.

Haines, R.B. 1931. Growth of microorganisms on chilled and frozen meat. J. Soc. Chem. Ind. 50:223-227.

Hajela, R.K., D.P. Horvath, S.J. Gilmour, and M.F. Thomashow. 1990. Molecular cloning and expression of *cor* (cold-regulated) genes in *Arabidopsis thaliana*. Plant Physiol. 93:1246-1252.

Hall, B.G. 1983. Evolution of new metabolic functions in laboratory organisms. Pp. 234-257 *in* M. Nei, R.K. Koehn, eds. Evolution of Genes and Proteins. Sinauer, Sunderland, Mass.

Harris, S.A. 1986. The Permafrost Environment. Barnes and Nobles, Totowa, N. Y. pp 276.

Hartl, D.L. and D.E. Dykhuizen. 1979. A selectively driven molecular clock. Nature 281:230-231.

Harvey, P.H. and M.D. Pagel. 1991. The Comparative Method in Evolutionary Biology. Oxford University Press, Oxford.

Hays, L.M., R.E. Feeney, L.M. Crowe, J.H. Crowe and A.E. Oliver. 1996. Antifreeze glycoproteins inhibit leakage from liposomes during thermotropic phase transitions. Proc. Natl. Acad. Sci. 93:6835-6840.

Helmke, E. and H. Weyland. 1995. Bacteria in sea ice and underlying water of the Eastern Weddell Sea in midwinter. Mar. Ecol. Progr. Ser. 117:269-287.

Hillis, D.M., J.J. Bull, M.E. White, M.R. Badgett, and I.J. Molineux. 1992. Experimental phylogenetics: generation of a known phylogeny. Science 255:489-592.

Hong, B., R. Barg, and T.H. Ho. 1992. Developmental and organ-specific expression of an ABA- and stress- induced protein in barley. Plant Mol. Biol. 18:663-674.
Horowitz, N.H., R.E. Cameron, and J.S. Hubbard. 1972. Microbiology of the Dry Valleys of Antarctica. Science 176:242-245.
Houde, M., R.S. Dhindsa, and F. Sarhan. 1992. A molecular marker to select for freezing tolerance in Gramineae. Mol. Gen. Genet. 234:43-48.
Huey, R.B., G.W. Gilchrist, M.L. Carlson, D. Berrigan, and L. Serra. 2000. Rapid evolution of a geographic cline in size in an introduced fly. Science 287:308-309.
Hughes, M.A. and M.A. Dunn. 1996. The molecular biology of plant acclimation to low temperature. J. Exp. Bot. 47:291-305.
Ingraham, J.L. and A.G. Marr. 1996. Effect of temperature, pressure, pH, and osmotic stress on growth. Pp. 1570-1578 *in* F. C. Neidhardt, editor-in-chief. Escherichia coli and Salmonella: Cellular and Molecular Biology, 2nd edition. ASM Press, Washington.
Jaglo-Ottosen, K.R., S.J. Gilmour, D.G. Zarka, O. Schabenberger, and M.F. Thomashow. 1998. *Arabidopsis CBF1* overexpression induces *COR* genes and enhances freezing tolerance. Science 280:104-106.
Jenke, D.R. 1996. Practical examination of a nonporous silica stationary phase for reversed-phase fast LC applications. J. Chromatogr. Sci. 34:362.
Jolly, W.T., E.G. Lidiak, J.H. Schellekens, and H. Santos. 1998. Volcanism, tectonics, and stratigraphic correlation in Puerto Rico, *in* Lidiak, E. G., and Larue, D. K., eds., Tectonics and Geochemistry of the Northeastern Caribbean: Boulder, Colorado, Geological Society of America Special Paper 322.
Jones, P.G., R.A. VanBogelen, and F.C. Neidhardt. 1987. Induction of proteins in response to low temperatures in *Escherichia coli*. J. Bacteriol. 169:2092-2095.
Jones, P.G. and M. Inouye. The Cold-Shock Response – A Hot Topic. 1994. Mol. Microbiol. 11:811.
Junge, K., J.J. Gosink, H.G. Hoppe, and J.T. Staley. 1998. *Arthrobacter, Brachybacterium* and *Planococcus* isolates identified from Antarctic sea ice brine. Description of *Planococcus mcmeekinii* sp. nov. Syst. Appl. Microbiol. 21:306-314.
Kaan, T., B. Jurgen, and T. Schweder. 1999. Regulation of the expression of the cold shock proteins CspB and CspC in *Bacillus subtilis*. Mol. Gen. Genet. 262:351.
Kalghatgi, K. and C. Horvath. 1990. Micropellicular Sorbents for Rapid Reversed-Phase Chromatography of Proteins and Peptides. ACS Sym. Ser. 434:162.
Kandror, O. and A.L. Goldberg. 1997. Trigger factor is induced upon cold shock and enhances viability of *Escherichia coli* at low temperatures. Proc. Natl. Acad. Sci. USA. 94:4978-4981.
Kann, M. and R.A. Goldstein. 2000. Optimizing for success: A new score function for distantly related protein sequence comparison, *in* RECOMB 2000, Shamir, Miyano, Istrail, Pevzner, and Waterman, eds., ACM Press, New York, pps. 177-182
Kann, M. and R.A. Goldstein. 2001. Optimization of a new score function for the detection of remote homologs. Proteins. In press.
Kappen, L. 1989. Field measurements of carbon dioxide exchange of the Antarctic lichen *Usnea sphacelata* in the frozen state. Antarctic Science 1:31-34.
Kasuga, M., Q. Liu, S. Miura, K. Yamaguchi-Shinozaki, and K. Shinozaki. 1999. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nature Biotech. 17:287-291.

Khlebnikova, G.M., D.A. Gilichinsky, D.G. Fedorov-Davydov, and E.A. Vorobyova. 1990. Quantitative evaluation of microorganisms in permafrost deposits and buried soils. Microbiology 59:106-112.

Kleckner, N., J. Bender, and S. Gottesman. 1991. Uses of transposons with emphasis on Tn*10*. Meth. Enzymol. 204:139-180.

Koshi, J.M. and R.A. Goldstein. 1995. Context-dependent optimal substitution matrices. Protein Engineering 8:641-645.

Koshi, J.M. and R.A. Goldstein. 1996. Probabilistic reconstruction of ancestral protein sequences. J. of Molecular Evolution 42:313-320.

Koshi, J.M. and R.A. Goldstein. 1997. Mutation matrices: Correlations and implications. Proteins 27:336-344.

Koshi, J.M., D.P. Mindell, and R.A. Goldstein. 1997. Beyond mutation matrices: Physical-chemistry based evolutionary models, *in* Genome Informatics 1997, Miyano and Tagagi, eds., Universal Academy Press, pps. 80-89.

Koshi, J.M. and R.A. Goldstein. 1998. Models of natural mutations including site heterogeneity. Proteins 32:289-295

Koshi, J.M., D.P. Mindell, and R.A. Goldstein. 1999. Using physical-chemistry based substitution models in phylogenetic analyses of HIV-1 subtypes. Molecular Biology and Evolution 16:173-179.

Koshi, J.M. and R.A. Goldstein. 2001. Analyzing rate heterogeneity during protein evolution, *in* Pacific Symposium on Biocomputing 2001. In press.

Koster, K.K. and D.V. Lynch. 1992. Solute accumulation and compartmentation during the cold acclimation of puma rye. Plant Physiol. 98:108-113.

Larkin, J.M. and J.L. Stokes. 1968. Growth of psychrophilic microorganisms at subzero temperatures. Can. J. Microbiol. 14:97-101.

Larson, A. and J.B. Losos. 1996. Phylogenetic systematics of adaptation. Pp. 187-220 *in* M. R. Rose and G. V. Lauder, eds. Adaptation. Academic Press, San Diego.

Lee, S.J., A. Xie, W. Jiang, J. Etchegaray, P.G. Jones, and M. Inouye. 1994. Family of the major cold-shock protein, CspA (CS7.4), of *Escherichia coli*, whose members show a high sequence similarity with the eukaryotic Y-box binding proteins. Mol. Microbiol. 11:833-839.

Lenski, R.E. and B.R. Levin. 1985. Constraints on the coevolution of bacteria and virulent phage: a model, some experiments, and predictions for natural communities. Am. Nat. 125:585-602.

Lenski, R.E. 1988. Experimental studies of pleiotropy and epistasis in *Escherichia coli*. II. Compensation for maladaptive pleiotropic effects associated with resistance to virus T4. Evolution 42:433-440.

Lenski, R.E., M.R. Rose, S.C. Simpson, and S.C. Tadler. 1991. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. Am. Nat. 138:1315-1341.

Lenski, R.E. and A.F. Bennett. 1993. Evolutionary response of *Escherichia coli* to thermal stress. Am. Nat. 142:S47-S64.

Lenski, R.E. and M. Travisano. 1994. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. Proc. Natl. Acad. Sci. USA 91:6808-6814.

Lenski, R.E. 1995a. Molecules are more than markers: new directions in molecular microbial ecology. Mol. Ecol. 4:643-651.

Lenski, R.E. 1995b. Evolution in experimental populations of bacteria. Pp. 193-215 *in* S. Baumberg, J.P.W. Young, E.M.H. Wellington, and J.R. Saunders, editors. Population Genetics of Bacteria. Cambridge University Press, Cambridge.

Lenski, R.E. 2000. Evolution, theory and experiments. Vol. 2, pp. 283-298 *in* J. Lederberg, editor. Encyclopedia of Microbiology, 2nd edition. Academic Press, San Diego.

Leroi, A.M., R.E. Lenski, and A.F. Bennett. 1994a. Evolutionary adaptation to temperature. III. Adaptation of *Escherichia coli* to a temporally varying environment. Evolution 48:1222-1229.

Leroi, A.M., A.F. Bennett, and R.E. Lenski. 1994b. Temperature acclimation and competitive fitness: an experimental test of the beneficial acclimation assumption. Proc. Natl. Acad. Sci. USA 91:1917-1921.

Lin, C. and M.F. Thomashow. 1992. DNA sequence analysis of a complementary DNA for cold-regulation *Arabidopsis* gene *cor15* and characterization of the COR15 polypeptide. Plant Physiol. 99:519-525.

Liquid Chromatography: Tandem System Rivals 2-D Gels. 1999. High Tech Separation News, 12:3.

Liu, Q., M. Kasuga, Y. Sakuma, H. Abe, S. Miura, K. Yamaguchi-Shinozaki, K. Shinozaki. 1998. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. Plant Cell 10:1391-1406.

Los, D.A., M.K. Ray, and N. Murata. 1997. Differences in the control of the temperature-dependent expression of four genes for desaturases in *Synechocystis* sp. PCC 6803. Mol. Microbiol. 25:1167-1175.

Maguin, E., H. Prevost, S.D. Ehrlich, and A. Gruss. 1996. Efficient insertional mutagenesis in lactococci and other gram-positive bacteria. J. Bacteriol. 178:931-935.

Mayr, B., T. Kaplan, S. Lechner, and S. Scherer. 1996. Identification and purification of a family of dimeric major cold shock protein homologs from the psychrotrophic *Bacillus cereus* WSBC 10201. J. Bacteriol. 178:2916-2925.

Mazur, P. 1970. Cryobiology: the freezing of biological systems. Science 168:939-949.

McCammon, S.A., B.H. Innes, J.P. Bowman, P.D. Franzmann, S.J. Dobson, S.P.E. Holloway, J.H. Skerratt, P.D. Nichols, and L.M. Rankin. 1998. *Flavobacterium hibernum* sp. nov., a lactose-utilizing bacterium from a freshwater Antarctic lake. Int. J. Syst. Bacteriol. 48:1405-1412.

McKersie, B.D. and S.R. Bowley. 1997. Active oxygen and freezing tolerance in transgenic plants., p. 203-214. *In* P. H. Li and T. H. H. Chen (ed.), Plant Cold Hardiness. Molecular Biology, Biochemistry, and Physiology. Plenum Press, New York.

Michel, V., I. Lehoux, G. Depret, P. Anglade, J. Labadie, and M. Hebraud. 1997. The cold shock response of the psychrotrophic bacterium *Pseudomonas fragi* involves four low-molecular-mass nucleic acid-binding proteins. J. Bacteriol. 179:7331-7342.

Mongold, J.A., A.F. Bennett, and R.E. Lenski. 1996. Evolutionary adaptation to temperature. IV. Adaptation of *Escherichia coli* at a niche boundary. Evolution 50:35-43.

Mongold, J.A., A.F. Bennett, and R.E. Lenski. 1999. Evolutionary adaptation to temperature. VII. Extension of the upper thermal limit of *Escherichia coli*. Evolution 53:386-394.

Monroy, A.F., Y. Castonguay, S. Laberge, F. Sarhan, L.P. Vezina, and R.S. Dhindsa. 1993. A new cold-induced alfalfa gene is associated with enhanced hardening at subzero temperature. Plant Physiol. 102:873-879.

Mortz, E., O. Vorm, M. Mann, and P. Roepstorff. 1994. Identification of Proteins in Polyacrylamide Gels by Mass Spectrometric Peptide-Mapping Combined with Database Search. Biol. Mass Spectrom. 23:249.

Muller, S. 1943. Permafrost or permanently frozen ground and related engineering problems. Strategic Studies, 62. United States Army.

Nakashima, K., K. Kanamaru, T. Mizuno, and K. Horikoshi. 1996. A novel member of the *cspA* family of genes that is induced by cold shock in *Escherichia coli*. J. Bacteriol. 178:2994-2997.

Nakatsu, C.H., R. Korona, R.E. Lenski, F.J. de Bruijn, T.L. Marsh, and L.J. Forney. 1998. Parallel and divergent genotypic evolution in experimental populations of *Ralstonia* sp. J. Bacteriol. 180:4325-4331.

Nanjo, T., M. Kobayashi, Y. Yoshiba, Y. Kakubari, K. Yamaguchi-Shinozaki, and K. Shinozaki. 1999. Antisense suppression of proline degradation improves tolerance to freezing and salinity in *Arabidopsis thaliana*. FEBS Lett 461:205-210.

Newman, T., F.J. de Bruijn, P. Green, K. Keegstra, H. Kende, L. McIntosh, J. Ohlrogge, N. Raikhel, S. Somerville, M. Thomashow, E. Retzel, and C. Somerville. 1994. Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones. Plant Physiol. 106:1241-1255.

Ng, H., J.L. Ingraham, and A.G. Marr. 1962. Damage and depression in *Escherichia coli* resulting from growth at low temperatures. J. Bacteriol. 84:331-339.

Olien, C.R. and M.N. Smith. 1977. Ice adhesions in relation to freeze stress. Plant Physiol. 60:499-503.

Palmisano, A.C. and D.L. Garrison. 1993. Microorganisms in Antarctic sea ice, pp. 167-218. *In*. E.I. Friedmann (eds.), Antarctic Microbiology. Wiley-Liss, New York, NY.

Panoff, J.M., D. Corroler, B. Thammavongs, and P. Boutibonnes. 1997. Differentiation between cold shock proteins and cold acclimation proteins in a mesophilic gram-positive bacterium, *Enterococcus faecalis* JH2-2. J. Bacteriol. 179:4451.

Panoff, J.M., B. Thammavongs, M. Gueguen, and P. Boutibonnes. 1998. Cold stress responses in mesophilic bacteria. Cryobiology 36:75-83.

Panoff, J.M., B. Thammavongs, and M. Gueguen. 2000. Cryoprotectants lead to phenotypic adaptation to freeze-thaw stress in *Lactobacillus delbrueckii* ssp. *bulgaricus* CIP 101027T. Cryobiology. 40:264-269.

Papadopoulos, D., D. Schneider, J. Meier-Eiss, W. Arber, R.E. Lenski, and M. Blot. 1999. Genomic evolution during a 10,000-generation experiment with bacteria. Proc. Natl. Acad. Sci. USA 96:3807-3812.

Park, J.I., C.M. Grant, M.J. Davies, and I.W. Dawes. 1998. The cytoplasmic Cu,Zn superoxide dismutase of *Saccharomyces cerevisiae* is required for resistance to freeze-thaw stress: generation of free radicals during freezing and thawing. J. Biol. Chem. 273:22921-22928.

Priscu, J.C., C.H. Fritsen, E.E. Adams, S.J. Giovannoni, H.W. Pearl, C.P. McKay, P.T. Doran, D.A. Gordon, B.D. Lanoil, and J.L. Pinckney. 1998. Perennial Antarctic Lake ice: an oasis for life in a polar desert. Science 280:2095-2098.

Rainey, P.B. and M. Travisano. 1998. Adaptive radiation in a heterogeneous environment. Nature 394:69-72.

Richmand, C.S., J.D. Glasner, R. Mau, H.F. Jin, F.R. Blattner. 1999. Genome-wide expression profiling in Escherichia coli K-12. Nuc. Acids Res. 27:3821-3835.

Rodriquez, G.G., D. Phipps, K. Ishiguro, and H.F. Ridgway. 1992. Use of a fluorescent redox probe for direct visualization of actively respiring bacteria. Appl. Environ. Micro. 58:1801-1808.

Rosenzweig, R.F., R.R. Sharp, D.S. Treves, and J. Adams. 1994. Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. Genetics 137:903-917.

Rudolph, A.S. and J.H. Crowe. 1985. Membrane stabilization during freezing: the role of two natural cryoprotectants, trehalose and proline. Cryobiology 22:367-377.

Russell, N.J. 1990. Cold adaptation of micro-organisms. Phil. Trans. R. Soc. Lond. B 326:595-611.

Sakamoto, T. and D.A. Bryant. 1997. Temperature-regulated mRNA accumulation and stabilization for fatty acid desaturase genes in the cyanobacterium *Synechococcus* sp. strain PCC 7002. Mol. Microbiol. 23:1281-1292.

Sakamoto, T., G. Shen, S. Higashi, N. Murata, D.A. Bryant. 1998. Alteration of low-temperature susceptibility of the cyanobacterium *Synechococcus* sp. PCC 7002 by genetic manipulation of membrane lipid unsaturation. Arch. Microbiol. 169:20-28.

Schneider, D., E. Duperchy, E. Coursange, R.E. Lenski, and M. Blot. 2000. Long-term experimental evolution in *Escherichia coli*. IX. Characterization of insertion sequence-mediated mutations and rearrangements. Genetics 156, in press.

Schrag, S.J., V. Perrot, and B.R. Levin. 1997. Adaptation to the fitness costs of antibiotic resistance in *Escherichia coli*. Proc. Roy. Soc. Lond. B 264:1287-1291.

Shaw, M.K., A.G. Marr, and J.L. Ingraham. 1971. Determination of the minimum temperature for growth of *Escherichia coli*. J. Bacteriol. 105:683-684.

Shyamala, V., and G. Ferro-Luzzi Ames. 1989. Genome walking by single-specific primer polymerase chain reaction: SSP-PCR. Gene 84:1-8.

Sidebottom, C., S. Buckley, P. Pudney, S. Twigg, C. Jarman, C. Holt, J. Telford, A. McArthur, D. Worrall, R. Hubbard, and P. Lillford. 2000. Heat-stable antifreeze protein from grass. Nature 406:256.

Silberg, J.J., K.G. Hoff, and L.E. Vickery. 1998. The Hsc66-Hsc20 chaperone system in *Escherichia coli*: chaperone activity and interactions with the DnaK-DnaJ-GrpE system. J. Bacteriol. 180:6617-6624.

Smith, K. and P.Youngman. 1992. Use of a new integrational vector to investigate compartment-specific expression of the *Bacillus subtilis spoIIM* gene. Biochimie 74:705-711.

Sniegowski, P.D., P.J. Gerrish, and R.E. Lenski. 1997. Evolution of high mutation rates in experimental populations of *Escherichia coli*. Nature 387:703-705.

Steponkus, P.L., M. Uemura, and M.S. Webb. 1993. Membrane destabilization during freeze-induced dehydration. Curr. Topics Plant Physiol. 10:37-47.

Steponkus, P.L., M. Uemura, R.A. Joseph, S.J. Gilmour, and M.F. Thomashow. 1998. Mode of action of the *COR15a* gene on the freezing tolerance of *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 95:14570-14575.

Stockinger, E.J., S.J. Gilmour, and M.F. Thomashow. 1997. *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-

repeat/DRE, a *cis*-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. Proc. Natl. Acad. Sci. USA 94:1035-1040.

Storad, B.C. 1990. Forever Frozen. ASU Res. 5:22-25.

Straka, R.P. and J.L. Stokes. 1960. Psychrophilic bacteria from Antarctica. J. Bacteriol. 80:622-625.

Strauss, G. and H. Hauser. 1986. Stabilization of lipid bilayer vesicles by sucrose during freezing. Proc. Natl. Acad. Sci. USA 83:2422-2426.

Strom, A.R., P. Falkenberg, and B. Landfald. 1986. Genetics of osmoregulation in *Escherichia coli*: uptake and synthesis of organic osmolytes. FEMS Microbiol. Rev. 39:79-86.

Talaat, A.M., P. Hunter, and S.A. Johnston. 2000. Genome-directed primers for selective labeling of bacterial transcripts for DNA microarray analysis. Nature Biotech. 18:679-682.

Tao, H., C. Bausch, C. Richmond, F.R. Blattner, and T. Conway. 1999. Functional genomics: expression analysis of *Escherichia coli* growing on minimal and rich media. J. Bacteriol. 181:6425-6440.

Tauch, A., Z. Zheng, A. Puhler, and J. Kalinowski. 1998. *Corynebacterium striatum* chloramphenicol resistance transposon Tn*5564*: genetic organization and transposition in *Corynebacterium glutamicum*. Plasmid. 40:126-139.

Thammavongs, B., D. Corroler, J.M. Panoff, Y. Auffray, and P. Boutibonnes. 1996. Physiological response of *Enterococcus faecalis* JH2-2 to cold shock: growth at low temperatures and freezing/thawing challenge. Lett. Appl. Microbiol. 23:398-402.

Thieringer, H.A., P.G. Jones, and M. Inouye. 1998. Cold shock and adaptation. BioEssays 20:49-57.

Thomashow, M.F., S. Hugly, W. Buchholz and L.S. Thomashow. 1986. Molecular basis for the auxin-independent phenotype of crown gall tumor tissues. Science 231:616-618.

Thomashow, M.F. 1998. Role of cold-responsive genes in plant freezing tolerance. Plant Physiol. 118:1-7.

Thomashow, M.F. 1999. Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Annu. Rev. Plant Physiol. Plant Mol. Biol. 50: 571-599.

Thompson, M.J. and R.A. Goldstein. 1996. Constructing amino-acid residue substitution classes maximally indicative of local protein structure. Proteins 25:28-37.

Thompson, M.J. and R.A. Goldstein. 1996. Predicting solvent accessibility: Higher accuracy using Bayesian statistics and optimized residue substitution classes. Proteins 25:38-47.

Thompson, M.J. and R.A. Goldstein. 1997. Predicting protein secondary structure with probabilistic schemata of evolutionarily-derived information. Protein Science 6:1963-1975.

Tomich, P.K., F.Y. An, and D.B. Clewell. 1980. Properties of erythromycin-inducible transposon Tn*917* in *Streptococcus faecalis*. J. Bacteriol. 141:1366-1374.

Travisano, M., J.A. Mongold, A.F. Bennett, and R.E. Lenski. 1995. Experimental tests of the roles of adaptation, chance, and history in evolution. Science 267:87-90.

Tschistajakow, F.M. and G.L. Botescharowa. 1938 a. The effect of low temperatures on microorganisms. II (in Russian). Mikrobiologiya (USSR) 7:498-524.

Tschistajakow, F.M. and G.L. Botescharowa. 1938 b. The effect of low temperatures on microorganisms. IV (in Russian). Mikrobiologiya (USSR) 7: 838-842.

Turner, P.E., V.S. Cooper, and R.E. Lenski. 1998. Tradeoff between horizontal and vertical modes of transmission in bacterial plasmids. Evolution 52:455-469.

Turner, P.E. and L. Chao. 1999. Prisoner's dilemma in an RNA virus. Nature 398:441-443.
Velicer, G.J., L. Kroos, and R.E. Lenski. 1998. Loss of social behaviors by *Myxococcus xanthus* during evolution in an unstructured habitat. Proc. Natl. Acad. Sci. USA 95:12376-12380.
Velicer, G.J., L. Kroos, and R.E. Lenski. 2000. Developmental cheating in the social bacterium *Myxococcus xanthus*. Nature 404:598-601.
Versalovic, J., T. Koeuth, and J.R. Lupski, 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res. 19:6823-6831.
Vishnivetskaya, T., S. Kathariou, J. McGrath, D. Gilichinsky, and J.M. Tiedje. 2000. Low temperature recovery strategies for the isolation of bacteria from ancient permafrost sediments. Extremophiles 4:165-173.
Vorobyova, E., D. Gilichinsky, V. Soina, M.A. Gorlenko, N.E. Minkovskaya, E.M. Rivkina, and T. Vishnivetskaya. 1997. Deep cold biosphere: Facts and hypothesis. FEMS Microbiol. Rev. 20:277-290.
Vulic, M., R.E. Lenski, and M. Radman. 1999. Mutation, recombination and incipient speciation of bacteria in the laboratory. Proc. Natl. Acad. Sci. USA 96:7348-7351.
Wall, D.B., M.T. Kachman, S. Gong, R. Hinderer, S. Parus, D.E. Miset, S.M. Hanash, and D.M. Lubman. 2000. Isoelectric focusing nonporous RP-HPLC: A two-dimensional liquid-phase separation method for mapping of cellular proteins with identification using MALDI-TOF mass spectrometry. Anal. Chem. 72:1099.
Wang, N., K. Yamanaka, and M. Inouye. 1999. CspI, the ninth member of the CspA family of *Escherichia coli*, is induced upon cold shock. J. Bacteriol. 181:1603-1609.
Wichman, H.A., M.R. Badgett, L.A. Scott, C.M. Boulianne, and J.J. Bull. 1999. Different trajectories of parallel evolution during viral adaptation. Science 285:422-424.
Williams, P.J. and M.W. Smith. 1989. The Frozen Earth. Fundamentals of Geocryology. (Studies in polar research), Cambridge Univ. Press, 306 pp.
Willimsky, G., H. Bang, G. Fischer, and M.A. Marahiel. 1992. Characterization of *cspB*, a *Bacillus subtilis* inducible cold shock gene affecting cell viability at low temperatures. J. Bacteriol. 174:6326-6335.
Wouters, J.A., J.-W. Sanders, J. Kok, W.M. de Vos, O.P. Kuipers, and T. Abee. 1998. Clustered organization and transcriptional analysis of a family of five *csp* genes of *Lactococcus lactis* MG1363. Microbiol. 144:2885-2893.
Xu, H., M. Griffith, C.L. Patten, B.R. Glick. 1998. Isolation and characterization of an antifreeze protein with ice nucleation activity from the plant growth promoting rhizobacterium *Pseudomonas putida* GR 12-2. Can. J. Micro. 44:64-73.
Yamaguchi-Shinozaki, K. and K. Shinozaki. 1994. A novel cis-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. Plant Cell 6:251-264
Yoshiba, Y., T. Kiyosue, K. Nakashima, K. Yamaguchi-Shinozaki, and K. Shinozaki. 1997. Regulation of levels of proline as an osmolyte in plants under water stress. Plant Cell Physiol. 38:1095-1102
Zheng, W. and S. Kathariou. 1994. Transposon-induced mutants of *Listeria monocytogenes* incapable of growth at low temperature (4 degrees C). FEMS Microbiol Lett. 121:287-291.

Zvyagintsev, D.G., D.A. Gilichinsky, S.A. Blagodatsky, E.A. Vorobyova, G.M. Khlebnikova, A.A. Arkhangelov, and N.N. Kudryavtseva. 1985. Survival time of microorganisms in permanently frozen sedimentary rocks and buried soils. Micobiology 54:131-136.

Zvyagintsev, D.G., D.A. Gilichinsky, G.M. Khlebnikova, D.G. Fedorov-Davydov, and N.N. Kudryavtseva. 1990. Comparative characteristics of microbial cenoses from permafrost rocks of different age and genesis. Microbiolgy 59:332-338.