

Program Announcement To DOE National Laboratories LAB 03-05

Genomes to Life

SUMMARY: The Office of Biological and Environmental Research (OBER) and the Office of Advanced Scientific Computing Research (ASCR) of the Office of Science (SC), U.S. Department of Energy (DOE), hereby announce their interest in receiving proposals for research in the following areas that support the Genomes to Life research program (<http://www.doe-genomes-to-life.org/>):

- 1) Technologies and strategies to image individual proteins and multi-protein complexes in microbes and to image complex microbial communities;
- 2) Technologies for the high-throughput synthesis of proteins and their biological characterization;
- 3) Molecular tags to identify individual proteins and to characterize multi-protein complexes in microbial cells;
- 4) High resolution, quantitative microbial biochemistry;
- 5) New genomic strategies and technologies for studying complex microbial communities;
- 6) Pathway inference in prokaryotes;
- 7) Implications for society, the law, education, and technology transfer; and
- 8) Other novel and innovative technologies and research strategies to address the core goals of the Genomes to Life research program.

DATES: Statements of intent to apply, including information on collaborators, areas of proposed research and technology development, and a short (one page) summary of the proposed research should be submitted by Tuesday, January 7, 2003.

Formal research proposals are due by 4:30 PM E.D.T. Tuesday, April 22, 2003.

ADDRESS: Statements of intent to apply should be sent to Ms. Joanne Corcoran by E-mail at: joanne.corcoran@science.doe.gov with copies to Dr. David Thomassen at: david.thomassen@science.doe.gov and Dr. Gary Johnson at: gary.johnson@science.doe.gov.

Formal proposals in response to this solicitation are to be submitted as PDF files on CDs. Three CDs should be submitted for each proposal. Color images should be submitted as a separate file in PDF format and identified as such. These images should be kept to a minimum due to the limitations of reproducing them. They should be numbered and referred to in the body of the technical scientific proposal as Color image 1, Color image 2, etc.

The CDs, referencing Program Announcement LAB 03-05, should be sent to: Life Sciences Division, SC-72/Germantown Building, Office of Biological and Environmental Research,

Office of Science, U.S. Department of Energy, 1000 Independence Avenue, SW, Washington, D.C. 20585-1290, ATTN: Program Announcement LAB 03-05.

When submitting by U.S. Postal Service Express Mail, any commercial mail delivery service, or when hand carried by the researcher, the following address must be used: Life Sciences Division, SC-72, Office of Biological and Environmental Research, Office of Science, U.S. Department of Energy, 19901 Germantown Road, Germantown, MD 20874-1290, ATTN: Program Announcement LAB 03-05.

FOR FURTHER INFORMATION CONTACT: Dr. David Thomassen, telephone: (301) 903-9817, E-mail: david.thomassen@science.doe.gov, Office of Biological and Environmental Research, SC-72/Germantown Building; U.S. Department of Energy; 1000 Independence Avenue, SW; Washington, DC 20585-1290.

SUPPLEMENTARY INFORMATION: Biology has entered a new era - the era of systems biology - in which we will understand entire living organisms and their interactions with the environment. While scientists have long tried to understand the workings of individual genes or small groups of genes this new era in biology will focus research on entire networks of genes and even entire biological systems - small, single celled organisms at first and later more complex creatures ultimately including humans.

This dramatic advance is possible, in large part, because of the scientific and technical successes of the Human Genome Project. The information and technology now available to all scientists on the human genome and on a rapidly growing list of the genomes of other organisms from microbes to plants to worms to mice not only gives us new perspectives on the inner workings of biological systems but provides new opportunities to use this knowledge to solve problems in energy.

The Genomes to Life program is a systems biology research program that offers the possibility of biotechnology solutions that can give us abundant sources of clean energy yet control greenhouse gases like carbon dioxide, a key factor in global climate change, and that can help us clean up past contamination of the environment.

The overall goals of the Genomes to Life program include understanding:

1. Natural, multi-protein molecular machines of complex living systems.
2. Complex networks that control the assembly and operation of these machines.
3. The organization and biochemical capabilities of complex microbial communities.

These three goals will only be achieved if we develop:

4. A computational infrastructure for systems biology that enables the development of computational models for complex biological systems that can predict the behavior of these complex systems and their responses to the environment.

The Genomes to Life program supports a combination of large, well integrated, multidisciplinary research teams and smaller, focused research projects. This solicitation will support smaller, focused research projects to develop new technologies, research strategies, or research resources needed by the Genomes to Life program. Future solicitations will likely request proposals for both large, well integrated, multidisciplinary research teams and smaller, focused research projects.

Information on the research projects currently funded by the Genomes to Life program and a description of project goals and overall program organization can be found at: <http://www.doe-genomestolife.org/>.

Other useful web sites include:

Microbial Genome Program Home Page - <http://www.sc.doe.gov/ober/microbial.html>

DOE Joint Genome Institute Microbial Web Page - http://www.jgi.doe.gov/JGI_microbial/html/

Microbes of Interest to DOE. The initial focus of Genomes to Life is on microbes (including fungi) directly relevant to DOE mission needs in energy (cleaner energy, biomass conversion, carbon sequestration) or the environment (cleanup of metals and radionuclides at DOE sites). Research in Goals 1 and 2 takes advantage of and focuses on microbes whose complete DNA sequence is already known. Research in Goal 3 focuses on microbes or microbial communities of interest to, directly relevant to, or that will contribute substantially to an ability to address DOE mission needs. Selected, well-justified research using yeast is appropriate as a means of quickly generating data that addresses the needs of the Genomes to Life Program. However, the use of yeast as a long-term research focus will not be encouraged.

Data and Other Results. Any data and results generated through the investigations into Goals 1 through 4 that are appropriate to share with the broader community should be provided in timely, open, and machine-readable format where possible or appropriate. Microbial DNA sequence data will be publicly released according to the "Data Release Requirements: Microbial Genome Sequencing Projects" (<http://www.sc.doe.gov/production/ober/EPR/data.html>).

Software Development and Distribution. Software developed by research teams that is appropriate for distribution beyond the research team shall be made available to the biological and computational community. It is our intent that this software be accessible, useful, affordable, and interoperable with other software and with data. Proposals should include plans for assuring availability, stating whether: the software will be available as binary or source code, a fee will be charged for the use of the software, some users (e.g., commercial) will be charged while others not, in what way derivative products will be treated, etc. Statements such as that by the International Society for Computational Biology on Bioinformatics Software Availability, <http://www.iscb.org/pr.shtml>, may be used for reference.

Research Focus

(1) Technologies and strategies to image individual proteins and multi protein complexes in microbes and to image complex microbial communities.

This solicitation will promote the development of imaging technology (probes, instrumentation and computational methodology) needed to accomplish the Genomes to Life program goals. Applications or development of imaging technology should be directed to or easily adapted to the study of microbes. Development of probes and instrumentation should be complementary to and facilitate completion of Genomes to Life program goals, including currently funded projects (see currently funded projects at: <http://www.doegenomestolife.org/>).

Additional information on the projected imaging needs of the Genomes to Life program can be found at:

<http://www.doegenomestolife.org/technology/imaging/workshop2002/GTLimaging2002.pdf>.

Specific research needs include:

- Development of novel probes (fluorescent, electron dense, vibrational tags, etc.) with optimum physico-chemical properties that enable:
 - Visualization, tracking, assembly and disassembly of multi-protein molecular machines and their individual components. Multifunctional probes that measure structure, including post-translational modification and function in real time, are needed.
 - Rapid visualization and quantitation of intracellular processes with high spatial resolution.
 - Visualization and quantitation of microbial populations and communities with respect to their structure, functions, stability and response to environmental stress. Probes should be developed to determine the spatial and temporal concentration of nutrients, metabolites, signaling molecules, elements, extra cellular matrices and other biomolecules critical to maintaining microbial community structure and function. This should also include dynamic measuring of oxidative states and energy transfer kinetics.

Probes should be selective, non-perturbative, and resistant to degradation and should have unique spectroscopic signatures. Unambiguous experimental systems to validate probe performance should be presented.

- Development of new high-throughput tagging methods for chromophores, electron dense and other probes. Methods should be capable of being transported to the broader scientific community.
- Development of innovative optical and non-optical instrumentation that will visualize and quantitate dynamic aspects of molecular machines over a wide range of dimensions and time scales; enable simultaneous co-localization of different intra-cellular processes with high spatial resolution; and/or permit visualization of bacterial community composition and functions in the field as well as in the laboratory.
- Development of computational methods for rapid processing, storing, reconstructing, and three dimensional modeling of large image data sets, e.g., from cryoelectron microscopy.

Computational methods are needed that can predict capabilities and limitations of various probes and instruments over a wide range of size and time scales. Novel computational tools are needed to integrate cellular image data sets derived from different instruments and technologies. Models of bacterial community structure, growth, functions and adaptive responses should be constructed based on experimental data and should facilitate development of alternative experimental approaches.

(2) Technologies for the high-throughput synthesis of proteins and their biophysical characterization.

This solicitation seeks to promote the development of techniques and protocols for high-throughput, low-cost synthesis of full-length proteins directly from coding sequence and for their subsequent biophysical characterization. Availability of proteins will enable the production and confirmation of selective probes and molecular tags needed to address the broad goals of the Genomes to Life program.

An essential early requirement for turning genome information into biological understanding is having access to purified samples of at least the majority of proteins encoded in the genomes of interest. Even within the microbial-focus of Genomes to Life, this requirement is daunting. It must encompass, within the next decade, hundreds of different microbes and therefore many tens of thousands of proteins. Both the production and characterization goals are significantly broader than those of structural genomics programs. In those programs the goals are limited to the structural characterization of a relatively small fraction of proteins, and often protein fragments, that represent structurally novel motifs.

It is recognized that no satisfactory general approach currently exists and that not all proteins will likely yield to the same techniques. It is expected that a variety of both cell-free and cell-based systems will be required, as well as multiple characterization methods. Production and characterization technologies should be scalable, economic, and sufficiently robust to meet the production goal of milligram quantities of approximately 10,000 proteins per year.

An essential early need is the development of improved techniques for predicting from sequence what production and purification approaches are most likely to succeed with each protein. Thus, informatics is an integral component. Algorithms based on data from successful and failed protein expressions are expected to substantially inform and improve future protein production efficiency.

Informatics coupled with biophysical characterizations are expected to provide functional insights that may also explain why such a large number of biologically important, full-length proteins either can not be expressed in soluble form, or have structures that cannot be determined once expressed. These proteins may include substantial disordered regions that adopt structures only after interaction with appropriate protein binding partners. Reliable predictive algorithms based on expression and characterization databases are therefore needed to predict disorder and binding partners.

Areas in which improvements are sought include:

- Optimization of cloning and clone validation techniques to support the protein production process.
- Optimization of cell-free and cellular expression methods.
- Optimization of protein purification protocols.
- Improved strategies for increasing the fraction of proteins that can be synthesized by automated methods. This may include sequence-based predictions of methods most likely to succeed and insights for optimization of expression protocols.

High-throughput, economical approaches for characterizing synthesized protein to assess product quality and to predict protein function are also solicited. A goal is to provide multiple benchmark biophysical characterizations for each protein under several conditions. These approaches are expected to include:

- Biophysical techniques, e.g., mass spectrometry, circular dichroism, calorimetry, partial proteolysis, deuterium exchange, surface plasmon resonance, neutron scattering, nuclear magnetic resonance.
- Improved techniques for predicting, from protein sequence, ordered and disordered domains and for predicting solubility properties of proteins and protein domains.
- Integrated data acquisition and management tools for tracking all steps of the production and characterization process and for supporting detailed QC/QA procedures.
- Improved high-throughput methods to predict, then rapidly test, and finally to confirm binding partners for proteins so that the nearly infinite number of potential interactions is reduced to experimentally testable subset.

(3) Molecular tags to identify individual proteins and to characterize multi-protein complexes in microbial cells.

This solicitation seeks advances in technology needed to mass-produce molecular tags for proteins and protein complexes as tools to be used for determining function. As a top priority, technologies are sought for mass-producing specific protein recognition tags capable of functioning as:

- capture reagents in affinity extraction and purification protocols, and as.
- labeling reagents for intracellular and '*in situ*' localization and mapping studies.

These technologies must be scalable to permit tens of thousands of successful tags to be produced and characterized per year at affordable costs. It is recognized that none of the many approaches under development to address this problem have yet demonstrated compelling promise - even as generally effective laboratory-scale methods. Yet for the purposes of Genomes to Life and for modern biology altogether, very high-throughput, industrially robust methods to address this problem are required.

For the purposes of this solicitation, it is assumed that purified protein 'targets' will be provided to the researchers in micro-gram to milli-gram quantities so that tags can be optimized and characterized. Tags that interfere with function as well as those that do not interfere with protein

function are both needed to help better define the biological roles of proteins. Areas in which technological improvements are sought include:

- Scalable methods for producing 'epitope-directed' affinity reagents of high specificity and affinity for proteins capable of functioning either as affinity extraction and capture reagents or as intra-cellular labeling reagents. High success ratios (fraction of protein epitopes yielding useful reagents) are essential.
- Improvements in protein-directed affinity tag design to improve tag utility, e.g., to facilitate subsequent purification and or/imaging, to facilitate release of the tagged protein, to image with and without disrupting activity, etc.
- Improved methods for developing tags directed specifically to protein complexes as distinct from their component proteins. Labeling complexes with and without disrupting interactions amongst protein components will provide important functional insights.
- Improved strategies for predicting, from sequence data, what potential protein epitopes are likely to be successful targets for tagging with and without interfering with function, and for predicting what tag development methods are likely to work for a particular protein/epitope.
- Imaging and labeling methods for multiplex mapping of proteins within cells. Simultaneously monitoring multiple labeled proteins will provide more comprehensive views of multi-protein complexes and their activities.
- Informatics tools both for managing tag production processes and for managing the data resulting from their use.

(4) High resolution, quantitative microbial biochemistry.

As noted above, the initial focus of Genomes to Life is on microbes (including fungi) directly relevant to DOE mission needs in energy (cleaner energy, biomass conversion, carbon sequestration) or the environment (cleanup of metals and radionuclides at DOE sites). To this end, development of novel technologies are encouraged to support the characterization of the internal environment and organization of prokaryotic microbes relevant to DOE missions and the Genomes to Life program and to explore how the characteristics of a microbe's internal environment affect its metabolism and physiology.

Very little is known of the internal "milieu" of any cell. A microbial cell is not likely to be a "bag of dilute salt water" within which metabolites and gene products freely diffuse. There is internal organization due to structural cytoskeletal components, partitioning of gene products in different parts of the cell so that they can efficiently mediate their appropriate pathways, concentration gradients of proteins and small molecules across the volume of the cell, and physical effects caused by the cell membrane and intracellular constituents including the viscosity of a cell's cytoplasm.

A protein's localization within a cell, its relationships with other proteins, concentrations, and subcellular dynamics are critically important parameters in determining its function, for identifying functional networks of proteins in a morphological context, and for expanding our understanding of whole-cell function. Thus, studies on the topological, physical, and chemical

properties of cellular cytoplasm, their effects on protein dynamics, on flux rates of metabolites, on protein-protein and protein-ligand interactions, and ultimately, on protein function are needed.

Research is needed that furnishes information on the dynamic behavior of these various molecules as the "molecular machines" perform their functions and on the distribution, localization, movement, and temporal variations of the molecules and complexes inside individual microbes as they carry out reactions of relevance to DOE missions and the Genomes to Life Program. Research is also needed to characterize topological, physical, and chemical characteristics underlying cellular responses to external stimuli, e.g., nutrients, toxins, or changes in environmental conditions. Similarly, computational algorithms designed to recognize regulatory networks or patterns of gene expression under different circumstances are needed that can provide insights into co-regulated genes.

New methods that accomplish any of several aims are solicited:

- Techniques to map the spatial distribution and concentrations of proteins and metabolites within prokaryotes.
- Techniques to assess fluxes and changes in concentrations of metabolites as a function of intracellular parameters and spatial location.
- Techniques to effectively map the immediate environment surrounding specific proteins, protein complexes, or other structural components within prokaryotes.
- Techniques to measure changes in enzyme-catalyzed reaction rates (catabolic and anabolic) and fluxes, as a function of the internal cell milieu, e.g., distance from the inner membrane surface, proton concentration, temperature, etc.
- Techniques to quantitate intracellular protein-protein association/dissociation rates as a function of ion concentrations, dielectric constants, protein concentrations, small molecule (metabolite, cofactor, ligand, etc.) concentrations, or temperature.
- Techniques to link data from experiments addressing the above aims to the broader goals of the Genomes to Life Program.
- Techniques to exploit computational methods to interrogate resulting datasets in order to suggest experimental priorities and derive insights into the underlying biology.

(5) New genomic strategies and technologies for studying complex microbial communities.

Microorganisms are the largest reservoir of genetic and biochemical diversity on earth. New methods for examining microbial communities have revealed that uncultured microbes make up more than 99% of many natural microbial communities. DNA isolated directly from environmental samples is a tremendous resource for examining the structure and function of microbial communities. The science of microbial ecology will be advanced by understanding the distribution, diversity, relative abundance, and interactions of the microorganisms in these communities.

A goal of the Genomes to Life Program (Goal 3) is to dramatically extend current scientific and technical understanding of the genetic diversity and metabolic capabilities of microbial communities in the environment, especially those related to remediation, biogeochemical cycles, climate changes, energy production, and biotechnology. A challenge to achieving this objective,

however, is the difficulty in characterizing the complexity of microbial communities in nature. For example, it has been estimated that there may be thousands of different species in surface soils. Thus, new strategies and technologies are needed to help define and assess the repertoire of metabolic capabilities as embodied in the collective community's genomic sequence.

We need new technologies that enable us to:

- Determine whole-genome sequences of dominant uncultured microorganisms to estimate their genetic diversity and interrelationships. Novel technologies and strategies are needed to use the genome sequence to identify the genes, metabolic pathways, regulatory network and proteins needed for survival, growth and adaptation to the environment.
- Identify the extent, patterns and spatial distribution of genetic diversity in microbial communities of interest to the DOE mission areas. In particular, we need to understand how microbial diversity supports community structure and function, and the relationship of genetic diversity to key environmental parameters. For example, one strategy for understanding the extent and pattern of genetic diversity in microbial communities is to sequence bacterial artificial chromosome (BAC) clones from individual microbial communities by the shotgun approach. Comparing BAC clone sequences should lead to insights into community genetic diversity and metabolic capacity.
- Understand the ecological functions of the uncultured microorganisms. We need to identify the metabolic functions that these genomes encode and to understand how those functions contribute to the community's ecological role in the environment. Of particular interest is the unique role of novel uncultured microorganisms in ecosystems relevant to DOE's missions in bioremediation, carbon sequestration, global climate change, energy production, and biotechnology.
- Determine cellular and biochemical functions of genes discovered in uncultured community members. This includes determining the protein complexes unique to uncultured microorganisms in ecosystems of DOE relevance, and whether their unique characteristics can be used for protein engineering.
- Understand the genetic basis of microbial community functional stability and adaptation in environments important to DOE missions. We need to understand the relationship between genetic diversity and microbial community stability. For example, the genetic basis and factors controlling microbial community stability and adaptation is of great importance in managing microbial communities to bioremediate contaminated sites, sequester carbon from the atmosphere, and contribute to sustainable energy production.

Key technologies needed to achieve these goals include, but are not limited to:

- New approaches for recovering RNA and high-molecular-weight DNA from environmental samples.
- New approaches for isolating single cells of uncultured microorganisms.
- New parallel comparative approaches that allow unique microbial community DNA fragments to be identified and the community to be characterized in automated high-throughput ways.

- Novel technologies and approaches for defining the patterns of expression and functions of genes from microbial communities with large numbers of uncultured microorganisms, under different environmental conditions.
- Advanced methods for community genome sequence assembly, genome comparison, microarray data analysis, and data management.

In addition, there are many computational challenges to characterizing the composition and functional capabilities of microbial communities. New algorithms for DNA sequence assembly and annotation will be required to analyze the multiorganism sequence data, and new modeling methods will be required to predict the behavior of microbial communities. Computational methods needed include the ability to deconvolute mixtures of partial genomes sampled in the environment and to identify individual organisms; to facilitate multiple-organism shotgun-sequence assembly; to improve comparative approaches to microbial sequence annotation and gene finding; to reconstruct pathways from sequenced or partially sequenced genomes; and to evaluate the combined metabolic capabilities of heterogeneous microbial populations. Importantly, computational methods are needed to correlate genomic, physiological, and biogeochemical site parameters, as well as their spatial and temporal distribution. Finally, methods to integrate regulatory-network, pathway, and expression data into integrated models of microbial community function are needed.

(6) Pathway Inference in Prokaryotes.

Many of the future solutions to the problems of supplying energy without net greenhouse gas emissions, managing the atmosphere's carbon budget, and remediating environmental contamination from metals, radionuclides, and toxic chemicals, will be based on biotechnology. Most of the new biotechnologies will almost certainly arise from fundamental advances in our understanding the 'microbial world'. This is primarily due to two facts. First, the metabolism of naturally occurring microorganisms plays a major role, often a dominant one, in many of the key chemical and energy fluxes of the planet. Second, virtually all of the biochemical transformations needed for safe energy production, carbon management, and environmental cleanup are part of the natural repertoire of one or more microorganisms. The challenge therefore is to explore and understand the immense chemical processing power that the microbial world possesses and uses. Achieving the needed understanding will require a nearly complete predictive mastery of the microbial cell from a 'systems' point of view - including their metabolic and signaling pathways, their regulatory networks, their material and energy flow constraints, etc. Data sets of considerable size and complexity must be obtained, managed, and mined. In addition, entirely new realms of modeling and simulation must be mastered.

The research requested in this section builds on advances in both computation and data base management as well as the extraordinary increase in the speed and capacity - and a corresponding reduction in the cost - of genome sequencing. Most fundamentally, it builds on the new and massive investment in the systems-level genomic-style study of microbial cells and microbial communities being undertaken as part of the Genomes to Life initiative.

The research requested in this section will facilitate the use of data obtained from the genomic and 'systems-level' experimental study of microbes (primarily prokaryotes) and microbial

communities. It will in particular assist in using these data to predict the role played by each of the proteins encoded in the microbe's genome, the microbe's signaling and metabolic pathways, its regulatory mechanisms, and its biochemical capacities. This research will help enable the re-annotation of incorrectly annotated genomes, the prediction of functions for unknown genes, and discovery of known functions for which no genes have been identified. Biochemical capacities with direct relevance to DOE missions, such as energy production, carbon fixation, bioremediation, etc. are of particular interest.

Pathway Inference: Information on regulatory, metabolic, and signaling pathways in prokaryotes is growing rapidly. Just as the use of similarity searches, such as Basic Local Alignment Search Tool, across genomes of multiple organisms has provided extraordinarily useful information regarding the imputed function of the target gene sequences, the research requested in this section is intended to facilitate similar inferences through probes of pathways in other organisms, primarily microbes. Although, some new knowledge may be required experimentally, the emphasis is on providing a computational infrastructure for this homology searching. Investigators may propose the construction of specific databases, research on knowledge representation, and/or tools to measure similarity or provide inference. Any proposed databases should contain references to the source of the data, including measures of presumed accuracy, based partly on whether annotations were derived from experimental results or computational analogy. Research may be proposed on data structures and data access tools for the integrative storage of pathway, signaling, and regulation information needed to support 'knowledge' extraction and in particular the computation of inferences about pathway structure and function. This goal presents questions concerning the types of data that should be stored and how they are to be interrelated, queried, presented, etc. Research also may be proposed to develop tools and resources that will support computational methods for inferring the existence and function of signaling, regulatory, and metabolic pathways. The research in this element initially may be conducted on organisms chosen for their utility to the research rather than for their importance to DOE, but the proposed research should show that it will be transferable to prokaryotes and pathways of DOE interest.

(7) Implications for Society, the Law, Education, and Technology Transfer.

Scientific research takes place in a context of ongoing societal concerns and expectations. Headlines about DNA, genes, and the new powers of science to analyze and manipulate fundamental elements of life vie for our attention daily. The dazzling diversity of applications of DNA science to fields ranging from medicine and agriculture to forensics and environmental restoration are having and will continue to have profound impacts on society and the lives of our citizenry. Many recent discoveries stem from data and tools generated by the Human Genome Project, whose goal is to describe in intricate detail the DNA from humans and other selected organisms by 2003. DNA is the information molecule that carries instructions for creating and maintaining all life. Resources and analytical technologies generated by the Human Genome Project and other genetic research can be applied to the DNA of all other organisms including those that are currently centerpieces of Genome to Life research. Thus, it is important for the Genomes to Life program to address some of the ethical, legal, and social issues that may arise from the project.

The Genomes to Life program initially focuses on nonpathogenic microbes of environmental importance and those that have potential to address DOE missions such as bioremediation, energy production, global climate change processes and biotechnology. To this end, research is solicited into the Implications for Society, the Law, Education, and Technology Transfer from the research being conducted under the Genomes to Life program. Investigations are encouraged that focus on:

- Defining the range, nature and scope of issues raised by Genomes to Life research or the applications of that research;
- Exploring legal issues such as intellectual property protection and commercialization practices that may be relevant to advances in the Genomes to Life program;
- Exploring potential economic sequelae to the introduction of Genomes to Life scientific developments into the marketplace, e.g., impacts on the biotechnology sector and other industries;
- Educational challenges from the Genomes to Life mediated "paradigm shift" from reductionist science to a more "reconstructionist" science, e.g., the need to present science as more of a synthetic activity requiring insights from different scientific disciplines.

The scope of research on the Implications for Society, the Law, Economics and Education is a work in progress and emphases will evolve as opportunities are identified to explore the consequences of Genomes to Life science for society.

(8) Other novel and innovative technologies and research strategies to address the core goals of the Genomes to Life research program.

Many different technologies, research strategies, and data resources will be required to successfully address the core goals of the Genomes to Life program. Proposals will be accepted that propose to develop additional tools, research strategies, or resources that will help speed success in reaching the core goals of the Genomes to Life program. In most cases, these new technologies and research strategies should be scalable and automatable for genome-scale analyses. A strategy for or demonstration of scalability and automatability should be described. The relevance to Genomes to Life goals should be clearly described.

Program Funding

Up to \$10 million is available in Fiscal Year 2003, contingent upon availability of appropriated funds. It is anticipated that individual research awards will be funded at a level of \$250,000 to \$1,000,000 per year.

Proposals

Electronic access to required forms is made available via the World Wide Web at: <http://www.science.doe.gov/production/grants/Forms-E.html>. DOE is under no obligation to pay for any costs associated with the preparation or submission of proposals if an award is not made.

The proposal must contain an abstract or project summary, letters of intent from collaborators, and short curriculum vitas consistent with NIH guidelines for all Principal and co-Principal Investigators.

Adherence to type size and line spacing requirements is necessary for several reasons. No researchers should have the advantage, or by using small type, of providing more text in their proposals. Small type may also make it difficult for reviewers to read the proposal. Proposals must have 1-inch margins at the top, bottom, and on each side. Type sizes must be 10 point or larger. Line spacing is at the discretion of the researcher but there must be no more than 6 lines per vertical inch of text. Pages should be standard 8 1/2" x 11" (or metric A4, i.e., 210 mm x 297 mm).

As noted above, color images should be submitted as a separate file in PDF format on the CD and identified as such. These images should be kept to a minimum due to the limitations of reproducing them. They should be numbered and referred to in the body of the technical scientific proposal as Color image 1, Color image 2, etc.

Researchers are expected to use the following ordered format to prepare proposals in addition to following instructions given below in the Office of Science Guide for Preparation of Scientific/Technical Proposals to be Submitted by National Laboratories. Proposals must be written in English, with all budgets in U.S. dollars.

- Field Work Proposal (FWP) Format (Reference DOE Order 5700.7C) (DOE ONLY)
- Project abstract (no more than one page) including the name of the researcher, mailing address, phone, Fax, and E-mail
- Budgets for each year and a summary budget page for the entire project period (using DOE F 4620.1)
- Budget explanation
- Budgets and budget explanation for each collaborative subproject, if any
- Project description (includes goals, background, research plan, preliminary studies and progress, and research design and methodologies) not to exceed 20 pages.
 - o Goals
 - o Background
 - o Research plan
 - o Preliminary studies and progress (if applicable)
 - o Research design and methodologies
- Literature cited
- Collaborative arrangements (if applicable)
- Biographical sketches (limit 2 pages per senior investigator)
- Description of facilities and resources
- Current and pending support for each senior investigator

Any recipient of an award from the Office of Science, performing research involving recombinant DNA molecules and/or organisms and viruses containing recombinant DNA

molecules shall comply with the National Institutes of Health "Guidelines for Research Involving Recombinant DNA Molecules," which is available via the World Wide Web at: <http://www.niehs.nih.gov/odhsb/biosafe/nih/rdna-apr98.pdf>, (59 FR 34496, July 5, 1994), or such later revision of those guidelines as may be published in the Federal Register. DOE policy requires that potential researchers adhere to 10 CFR Part 745 "Protection of Human Subjects" (if applicable), or such later revision of those guidelines as may be published in the Federal Register. DOE requirements for reporting, protection of human and animal subjects and related special matters can be found on the World Wide Web at: <http://www.science.doe.gov/production/grants/Welfare.html>.

The instructions and format described below should be followed. Reference Program Announcement LAB 03-05 on all submissions and inquiries about this program.

OFFICE OF SCIENCE GUIDE FOR PREPARATION OF SCIENTIFIC/TECHNICAL PROPOSALS TO BE SUBMITTED BY NATIONAL LABORATORIES

Proposals from National Laboratories submitted to the Office of Science (SC) as a result of this program announcement will follow the Department of Energy Field Work Proposal process with additional information requested to allow for scientific/technical merit review. The following guidelines for content and format are intended to facilitate an understanding of the requirements necessary for SC to conduct a merit review of a proposal. Please follow the guidelines carefully, as deviations could be cause for declination of a proposal without merit review.

1. Evaluation Criteria

Proposals will be subjected to formal merit review (peer review) and will be evaluated against the following criteria which are listed in descending order of importance:

Scientific and/or technical merit of the project

Appropriateness of the proposed method or approach

Competency of the personnel and adequacy of the proposed resources

Reasonableness and appropriateness of the proposed budget

The evaluation will include program policy factors such as the relevance of the proposed research to the terms of the announcement, the uniqueness of the proposer's capabilities, and demonstrated usefulness of the research for proposals in other DOE Program Offices as evidenced by a history of programmatic support directly related to the proposed work.

2. Summary of Proposal Contents

Field Work Proposal (FWP) Format (Reference DOE Order 5700.7C) (DOE ONLY)
Proposal Cover Page

Table of Contents
Abstract
Budget and Budget Explanation
Narrative
Literature Cited
Other support of investigators
Biographical Sketches
Description of facilities and resources
Appendix

2.1 Number of Copies to Submit

An original and seven copies of the formal proposal/FWP must be submitted. (Unless otherwise instructed in this Program Announcement.)

3. Detailed Contents of the Proposal

Proposals must be readily legible, when photocopied, and must conform to the following three requirements: the height of the letters must be no smaller than 10 point with at least 2 points of spacing between lines (leading); the type density must average no more than 17 characters per inch; the margins must be at least one-half inch on all sides. Figures, charts, tables, figure legends, etc., may include type smaller than these requirements so long as they are still fully legible.

3.1 Field Work Proposal Format (Reference DOE Order 5700.7C) (DOE ONLY)

The Field Work Proposal (FWP) is to be prepared and submitted consistent with policies of the investigator's laboratory and the local DOE Operations Office. Additional information is also requested to allow for scientific/technical merit review.

Laboratories may submit proposals directly to the SC Program office listed above. A copy should also be provided to the appropriate DOE operations office.

3.2 Proposal Cover Page

The following proposal cover page information may be placed on plain paper. No form is required.

Title of proposed project
SC Program announcement title
Name of laboratory
Name of principal investigator (PI)
Position title of PI
Mailing address of PI
Telephone of PI
Fax number of PI

Electronic mail address of PI
Name of official signing for laboratory*
Title of official
Fax number of official
Telephone of official
Electronic mail address of official
Requested funding for each year; total request
Use of human subjects in proposed project:
 If activities involving human subjects are not planned at any time during the proposed project period, state "No"; otherwise state "Yes", provide the IRB Approval date and Assurance of Compliance Number and include all necessary information with the proposal should human subjects be involved.
Use of vertebrate animals in proposed project:
 If activities involving vertebrate animals are not planned at any time during this project, state "No"; otherwise state "Yes" and provide the IACUC Approval date and Animal Welfare Assurance number from NIH and include all necessary information with the proposal.
Signature of PI, date of signature
Signature of official, date of signature*

*The signature certifies that personnel and facilities are available as stated in the proposal, if the project is funded.

3.3 Table of Contents

Provide the initial page number for each of the sections of the proposal. Number pages consecutively at the bottom of each page throughout the proposal. Start each major section at the top of a new page. Do not use unnumbered pages and do not use suffices, such as 5a, 5b.

3.4 Abstract

Provide an abstract of no more than 250 words. Give the broad, long-term objectives and what the specific research proposed is intended to accomplish. State the hypotheses to be tested. Indicate how the proposed research addresses the SC scientific/technical area specifically described in this announcement.

3.5 Narrative

The narrative comprises the research plan for the project and is limited to 25 pages. It should contain the following subsections:

Background and Significance: Briefly sketch the background leading to the present proposal, critically evaluate existing knowledge, and specifically identify the gaps which the project is intended to fill. State concisely the importance of the research described in the proposal. Explain the relevance of the project to the research needs identified by the Office of Science. Include

references to relevant published literature, both to work of the investigators and to work done by other researchers.

Preliminary Studies: Use this section to provide an account of any preliminary studies that may be pertinent to the proposal. Include any other information that will help to establish the experience and competence of the investigators to pursue the proposed project. References to appropriate publications and manuscripts submitted or accepted for publication may be included.

Research Design and Methods: Describe the research design and the procedures to be used to accomplish the specific aims of the project. Describe new techniques and methodologies and explain the advantages over existing techniques and methodologies. As part of this section, provide a tentative sequence or timetable for the project.

Subcontract or Consortium Arrangements: If any portion of the project described under "Research Design and Methods" is to be done in collaboration with another institution, provide information on the institution and why it is to do the specific component of the project. Further information on any such arrangements is to be given in the sections "Budget and Budget Explanation", "Biographical Sketches", and "Description of Facilities and Resources".

3.6 Literature Cited

List all references cited in the narrative. Limit citations to current literature relevant to the proposed research. Information about each reference should be sufficient for it to be located by a reviewer of the proposal.

3.7 Budget and Budget Explanation

A detailed budget is required for the entire project period, which normally will be three years, and for each fiscal year. It is preferred that DOE's budget page, Form 4620.1 be used for providing budget information*. Modifications of categories are permissible to comply with institutional practices, for example with regard to overhead costs.

A written justification of each budget item is to follow the budget pages. For personnel this should take the form of a one-sentence statement of the role of the person in the project. Provide a detailed justification of the need for each item of permanent equipment. Explain each of the other direct costs in sufficient detail for reviewers to be able to judge the appropriateness of the amount requested.

Further instructions regarding the budget are given in section 4 of this guide.

* Form 4620.1 is available at web site: <http://www.sc.doe.gov/production/grants/Forms.html>

3.8 Other Support of Investigators

Other support is defined as all financial resources, whether Federal, non-Federal, commercial or institutional, available in direct support of an individual's research endeavors. Information on

active and pending other support is required for all senior personnel, including investigators at collaborating institutions to be funded by a subcontract. For each item of other support, give the organization or agency, inclusive dates of the project or proposed project, annual funding, and level of effort devoted to the project.

3.9 Biographical Sketches

This information is required for senior personnel at the laboratory submitting the proposal and at all subcontracting institutions. The biographical sketch is limited to a maximum of two pages for each investigator.

3.10 Description of Facilities and Resources

Describe briefly the facilities to be used for the conduct of the proposed research. Indicate the performance sites and describe pertinent capabilities, including support facilities (such as machine shops) that will be used during the project. List the most important equipment items already available for the project and their pertinent capabilities. Include this information for each subcontracting institution, if any.

3.11 Appendix

Include collated sets of all appendix materials with each copy of the proposal. Do not use the appendix to circumvent the page limitations of the proposal. Information should be included that may not be easily accessible to a reviewer.

Reviewers are not required to consider information in the Appendix, only that in the body of the proposal. Reviewers may not have time to read extensive appendix materials with the same care as they will read the proposal proper.

The appendix may contain the following items: up to five publications, manuscripts (accepted for publication), abstracts, patents, or other printed materials directly relevant to this project, but not generally available to the scientific community; and letters from investigators at other institutions stating their agreement to participate in the project (do not include letters of endorsement of the project).

4. Detailed Instructions for the Budget

(DOE Form 4620.1 "Budget Page" may be used)

4.1 Salaries and Wages

List the names of the principal investigator and other key personnel and the estimated number of person-months for which DOE funding is requested. Proposers should list the number of postdoctoral associates and other professional positions included in the proposal and indicate the number of full-time-equivalent (FTE) person-months and rate of pay (hourly, monthly or annually). For graduate and undergraduate students and all other personnel categories such as secretarial, clerical, technical, etc., show the total number of people needed in each job title and

total salaries needed. Salaries requested must be consistent with the institution's regular practices. The budget explanation should define concisely the role of each position in the overall project.

4.2 Equipment

DOE defines equipment as "an item of tangible personal property that has a useful life of more than two years and an acquisition cost of \$25,000 or more." Special purpose equipment means equipment which is used only for research, scientific or other technical activities. Items of needed equipment should be individually listed by description and estimated cost, including tax, and adequately justified. Allowable items ordinarily will be limited to scientific equipment that is not already available for the conduct of the work. General purpose office equipment normally will not be considered eligible for support.

4.3 Domestic Travel

The type and extent of travel and its relation to the research should be specified. Funds may be requested for attendance at meetings and conferences, other travel associated with the work and subsistence. In order to qualify for support, attendance at meetings or conferences must enhance the investigator's capability to perform the research, plan extensions of it, or disseminate its results. Consultant's travel costs also may be requested.

4.4 Foreign Travel

Foreign travel is any travel outside Canada and the United States and its territories and possessions. Foreign travel may be approved only if it is directly related to project objectives.

4.5 Other Direct Costs

The budget should itemize other anticipated direct costs not included under the headings above, including materials and supplies, publication costs, computer services, and consultant services (which are discussed below). Other examples are: aircraft rental, space rental at research establishments away from the institution, minor building alterations, service charges, and fabrication of equipment or systems not available off-the-shelf. Reference books and periodicals may be charged to the project only if they are specifically related to the research.

a. Materials and Supplies

The budget should indicate in general terms the type of required expendable materials and supplies with their estimated costs. The breakdown should be more detailed when the cost is substantial.

b. Publication Costs/Page Charges

The budget may request funds for the costs of preparing and publishing the results of research, including costs of reports, reprints page charges, or other journal costs (except costs for prior or early publication), and necessary illustrations.

c. Consultant Services

Anticipated consultant services should be justified and information furnished on each individual's expertise, primary organizational affiliation, daily compensation rate and number of days expected service. Consultant's travel costs should be listed separately under travel in the budget.

d. Computer Services

The cost of computer services, including computer-based retrieval of scientific and technical information, may be requested. A justification based on the established computer service rates should be included.

e. Subcontracts

Subcontracts should be listed so that they can be properly evaluated. There should be an anticipated cost and an explanation of that cost for each subcontract. The total amount of each subcontract should also appear as a budget item.

4.6 Indirect Costs

Explain the basis for each overhead and indirect cost. Include the current rates.