

Appendix 8

Genomics:GTL Systems Biology Knowledgebase Workshop:

Agenda, Participant List, and Biosketches

Agenda

Preworkshop Assignments		
Participants will come with a 1- to 2-page position paper addressing particular charge questions.		
Participants will come with a 5- to 7-minute presentation.		
Tuesday, May 27, 2008		
Time	Presentation	Speakers
7:30 – 9:00 p.m.	Registration and Refreshments	
Wednesday, May 28, 2008		
8:00 – 8:05 a.m.	Welcome and Introductions	Susan K. Gregurick, Program Manager, Biological Systems Science Division, DOE Office of Biological and Environmental Research
8:05 – 8:20 a.m.	Introduction and OBER Overview	Anna Palmisano, Associate Director, DOE Office of Biological and Environmental Research
8:20 – 8:30 a.m.	Introduction and GTL Overview	David Thomassen, Acting Director, DOE Life and Medical Sciences Division, Office of Biological and Environmental Research
8:30 – 9:00 a.m.	Overview of Knowledgebase Workshop	Rick Stevens, Associate Laboratory Director for Computing and Life Sciences, Argonne National Laboratory, and University of Chicago
9:00 – 9:30 a.m.	Systems Analysis of Microbial Solar Energy Capture and Utilization	Tim Donohue, Great Lakes Bioenergy Research Center
9:30 – 10:00 a.m.	Nitrogen Regulatory Networks and Plant Systems Biology	Manpreet Katari, New York University
10:00 – 10:30 a.m.	Systems Biology Overview	Andrei Osterman, Burnham Institute
10:30 – 10:45 a.m.	Break	
11:00 – 2:30 p.m.	Science Drivers Breakouts: Basic Research Needs and Use Cases	
	<ul style="list-style-type: none"> Systems Biology for Bioenergy Solutions 	Brian Davison, Paul Adams, and Tim Donohue, DOE Bioenergy Research Centers
	<ul style="list-style-type: none"> Systems Biology for Carbon Cycle Understanding 	Mick Follows, Massachusetts Institute of Technology Grant Heffelfinger, Sandia National Laboratories
	<ul style="list-style-type: none"> Systems Biology Core 	Nitin Baliga, Institute for Systems Biology Andrei Osterman, Burnham Institute
2:30 – 3:00 p.m.	Reconvene as Larger Group to Report on Basic Research Needs and Use Cases	



3:30 – 3:45 p.m.	Break	
Time	Presentation	Speakers
3:45 – 6:00 p.m.	Science Drivers Breakouts: Path Forward	
	<ul style="list-style-type: none"> Systems Biology for Bioenergy Solutions 	Brian Davison, Paul Adams, and Tim Donohue, DOE Bioenergy Research Centers
	<ul style="list-style-type: none"> Systems Biology for Carbon Cycle Understanding 	Mick Follows, Massachusetts Institute of Technology
	<ul style="list-style-type: none"> Systems Biology Core 	Nitin Baliga, Institute for Systems Biology Andrei Osterman, Burnham Institute
6:00 p.m.	Group Dinner and Writing	
Thursday, May 29, 2008		
8:30 – 9:00 a.m.	Introduction and ASCR Overview	Michael Strayer, Associate Director, DOE Office of Advanced Scientific Computing Research
9:00 – 9:30 a.m.	CaBIG Overview	Cheryl Marks, National Institutes of Health
9:30 – 10:00 a.m.	European Efforts in Knowledgebase Development	Dawn Field, Natural Environment Research Council, Oxford, U.K.
10:00 – 10:15 a.m.	Break	
10:15 – 11:00 a.m.	Report out by Chairs on Day 1 Science Drivers	Chairs of Science Driver Breakouts
11:00 – 2:30 p.m.	Technology Driver Breakouts: Charge Questions	
	<ul style="list-style-type: none"> Knowledgebase Community and User Issues 	Owen White, University of Maryland Jennifer Reed, University of Wisconsin
	<ul style="list-style-type: none"> Data, Metadata, and Information 	Denise Schmoyer, Oak Ridge National Laboratory Gordon Anderson, Pacific Northwest National Laboratory
	<ul style="list-style-type: none"> Architecture and Technology 	Folker Meyer, Argonne National Laboratory Ed Uberbacher, Oak Ridge National Laboratory
	<ul style="list-style-type: none"> Data Integration 	Ross Overbeek, FIG, Argonne National Laboratory Nagiza Samatova, Oak Ridge National Laboratory and North Carolina State University
2:30 – 3:00 p.m.	Reconvene as larger Group to Report on Charge Questions	
3:30 – 3:45 p.m.	Break	

Participant List

- *Paul Adams**
Lawrence Berkeley National Laboratory
- **Gordon Anderson**
Pacific Northwest National Laboratory
- Rolf Apweiler**
European Bioinformatics Institute
- *Nitin Baliga**
Institute for Systems Biology
- Jacek Becla**
Stanford Linear Accelerator Center
- Richard Bonneau**
New York University, Courant
- Olga Brazhnik**
National Institutes of Health
- Thomas Brettin**
Los Alamos National Laboratory
- C. Robin Buell**
Michigan State University
- Dylan Chivian**
Lawrence Berkeley National Laboratory
- Bob Cottingham**
Oak Ridge National Laboratory
- Terence Critchlow**
Pacific Northwest National Laboratory
- *Brian Davison**
Oak Ridge National Laboratory
- Matt DeJongh**
Hope College
- Patrik D'haeseleer**
Lawrence Livermore National Laboratory
- *Tim Donohue**
University of Wisconsin, Madison
- Scott Elliott**
Los Alamos National Laboratory
- Dawn Field**
Centre of Ecology and Hydrology
- *Michael (Mick) Follows**
Massachusetts Institute of Technology
- Peg Folta**
Lawrence Livermore National Laboratory
- James K. Fredrickson**
Pacific Northwest National Laboratory
- Damian Gessler**
National Center for Genome Resources
- Stephen Goff**
University of Arizona
- Yakov Golder**
Lawrence Berkeley National Laboratory
- Ian Gorton**
Pacific Northwest National Laboratory
- Grant Heffelfinger**
Sandia National Laboratories
- Tatiana Karpinets**
Oak Ridge National Laboratory
- Manpreet Katari**
New York University
- Ken Kemner**
Argonne National Laboratory
- Les Klimczak**
Great Lakes Bioenergy Research Center
- Cheryl Kuske**
Los Alamos National Laboratory
- Mary Lipton**
Pacific Northwest National Laboratory
- Michael Lomas**
Bermuda Institute of Ocean Sciences
- John L. Markley**
University of Wisconsin, Madison
- Cheryl Marks**
National Cancer Institute
- Celeste Matarazzo**
Lawrence Livermore National Laboratory
- Raymond McCord**
Oak Ridge National Laboratory
- Lee Ann McCue**
Pacific Northwest National Laboratory
- Peter McGarvey**
Georgetown University Medical Center
- **Folker Meyer**
Argonne National Laboratory
- Bob Morris**
University of Washington
- Sean Murphy**
J. Craig Venter Institute
- Ilya Nemenman**
Los Alamos National Laboratory
- Gary J. Olsen**
University of Illinois
- *Andrei Osterman**
Burnham Institute for Medical Research
- **Ross Overbeek**
Argonne National Laboratory
- George N. Phillips, Jr.**
University of Wisconsin, Madison
- David Pletcher**
Joint BioEnergy Institute
- **Jennifer Reed**
University of Wisconsin, Madison
- **Nagiza F. Samatova**
Oak Ridge National Laboratory

* DOE Mission Session Cochair.

** Knowledgebase Session Cochair.

Appendix 8

****Denise Schmoyer**

Oak Ridge National Laboratory

Blake A. Simmons

Joint BioEnergy Institute

Tom Slezak

Lawrence Livermore National
Laboratory

Rick Stevens

Argonne National Laboratory

Michael R. Sussman

University of Wisconsin, Madison

Ronald Taylor

Pacific Northwest National Laboratory

Gerald A. Tuskan

Oak Ridge National Laboratory

****Edward C. Uberbacher**

Oak Ridge National Laboratory

****Owen White**

University of Maryland, College Park

John Wooley

University of California, San Diego

Alex Worden

Monterey Bay Aquarium Research
Institute

Cathy Wu

Georgetown University

Liming Yang

National Center for Research Resources

William S. York

University of Georgia

Biosketches

Paul D. Adams

Lawrence Berkeley National Laboratory

Paul Adams studied biochemistry at Edinburgh University where he received a doctorate in structural biology in 1992. Adams is a senior scientist and deputy director of the Physical Biosciences Division at Lawrence Berkeley National Laboratory, head of the Berkeley Center for Structural Biology, vice president for technology at the Joint BioEnergy Institute, and an adjunct professor in the bioengineering department at the University of California, Berkeley. His current research interests span computation, structural biology, and biofuels. Much of Adams's research is focused on developing new algorithms and computational methods for addressing problems in structural biology. He also leads development of the technology portal for the Protein Structure Initiative Knowledge Base.

Gordon Anderson

Pacific Northwest National Laboratory

Gordon Anderson has more than 30 years' experience in developing systems for instrument control, high-performance data acquisition, and data management. His experience and skills have been applied to high-throughput proteomic research at Pacific Northwest National Laboratory (PNNL). Proteomics produces large volumes of multidimensional data that must be organized and processed using a combination of commercial software and custom-designed tools. Anderson has assembled a multidisciplinary team at PNNL where he has led development of proteomic data management and analysis. The development of hardware and software has enabled advanced instrument control schemes for state-of-the-art, high-performance mass spectrometers at the Environmental Molecular Sciences Laboratory located at PNNL. Anderson's efforts in software development have enabled proteomics capabilities in the areas of complex spectral analysis and feature detection.

The informatics group at PNNL consists of 12 staff members responsible for data management and knowledge extraction from the raw data resulting from analysis of biological samples. Anderson holds 2 R&D 100 awards and 7 patents and has authored or coauthored more than 100 journal articles. He received his bachelor's degree in electrical engineering from Washington State University in 1985.

Rolf Apweiler

European Bioinformatics Institute

Rolf Apweiler studied biology in Heidelberg, Germany, and Bath, United Kingdom. He worked 3 years in drug

discovery in the pharmaceutical industry and has been involved in bioinformatics since 1987. Apweiler started his bioinformatics career working on Swiss-Prot at the European Molecular Biology Laboratory (EMBL) in Heidelberg. He joined EMBL's European Bioinformatics Institute (EBI) in Hinxton, United Kingdom, in 1994 and is now joint head of the Protein and Nucleotide Data Group (PANDA) at EBI (<http://www.ebi.ac.uk/panda/>). This group coordinates UniProt activities, InterPro, GOA, Reactome, PRIDE, IntAct, Ensembl, the EMBL nucleotide sequence database, and other projects at EBI.

Nitin Baliga

Institute for Systems Biology

Nitin Baliga received a master's degree in marine biotechnology from Goa University, India, and has a doctorate in microbiology from the University of Massachusetts, Amherst. He is an associate professor at the Institute for Systems Biology where he leads a multidisciplinary team in deciphering quantitative systems-scale models for complete gene regulatory circuits of diverse prokaryotic organisms. With a special focus on organisms such as *Halobacterium salinarum* NRC-1, Baliga's long-term goal is to tap into the extraordinary biological potential of extremophiles.

Jacek Becla

Stanford Linear Accelerator Center

Jacek Becla earned a master's degree in electronics engineering from the University of Science and Technology in Poland in 1995. He joined the Stanford Linear Accelerator Center in 1997 as an information systems specialist. Becla's primary expertise is developing systems for managing very large datasets, and he leads efforts related to architecting the petabyte database for the Large Synoptic Survey Telescope astronomical survey. Becla also organizes XLDB (Extremely Large Databases) workshops and helps coordinate the open-source science database, SciDB. In the past, he was one of the main designers of the BaBar database and was the manager of the BaBar database group.

Richard Bonneau

New York University

Richard Bonneau is a joint assistant professor in both the New York University biology department and the computer science department at the Courant Institute for Mathematical Sciences. He also serves as an affiliate faculty member at the Institute for Systems Biology in Seattle, Washington. Bonneau is the technical lead on two grid-computing collaborations with IBM—the first and second phases of

the Human Proteome Folding Project. Rich also oversees TACITUS's (<http://www.tacitus.com>) approach to data gaming for all applications that focus on genomics, computational biology, and cell biology and is part of multiple groups developing open-source tools for data visualization. His research focuses on three main topics: (1) learning dynamical regulatory and signaling networks from functional genomic data, (2) using state-of-the-art structure prediction and design methodologies (e.g., Rosetta) to predict protein function and to design new functions, and (3) conducting multiple data and multiple species biclustering (data integration). All these activities are united by a common motivation: developing novel computational tools that extract genome-wide mechanistic models from large functional genomic datasets.

Olga Brazhnik

National Institutes of Health

Olga Brazhnik is a program manager at the National Center for Research Resources within the National Institutes of Health. She started her career as a physicist applying theoretical and computational methods in biology and medicine. In 1993, she was awarded a research grant by the U.S. National Research Council and Academy of Sciences and joined the James Franck Institute at the University of Chicago. Brazhnik then took a position at Virginia Tech and in 1998 transitioned into information technology, searching for capabilities to enable effective transformation of abundant scientific data into knowledge. In 2000, she joined the Virginia Bioinformatics Institute where her work resulted in the creation of several bioinformatics databases (e.g., ESTAP, DOME, and SeedGenes). In 2002, Brazhnik became the chief database architect for the Epidemic Outbreak Surveillance Project and later for the COHORT project on real-time integration of clinical systems with the U.S. Air Force Surgeon General Office. Her work involved integrating clinical and biological data; designing and developing database systems in Oracle, SQL-2000, PostgreSQL, and Access; and participating in development of protocols for study design and data collection, analysis of microarray data, and implementation of MIAME and HL7 standards. Brazhnik joined the National Institutes of Health in 2004, and she is an affiliate associate professor at George Mason University.

Thomas Brettin

Los Alamos National Laboratory

Thomas Brettin is the bioinformatics team leader in the genome sciences group at Los Alamos National Laboratory (LANL). He currently is serving on a change of station to the Department of Energy (DOE) Joint Genome Institute's

Production Genomics Facility where he works as a software systems architect. Brettin has a master's degree in genetics; his more than 15 years' experience in genomics includes hands-on work and leadership roles in high-throughput sequencing laboratory automation, sequence annotation and analysis, software engineering, and information technology. Brettin is the principal investigator for a 5-year pathogen sequencing project funded by the Office of the Chief Scientist (formerly the Intelligence Technology Innovation Center). He also is principal investigator of the oral pathogens database project, now in its eighth year of funding as an interagency agreement among the National Institute for Dental and Craniofacial Research, the National Institutes of Health (NIH), and DOE. Brettin is a member of the Information Science and Technology Center's science council at LANL; serves on the scientific advisory board for the viral bioinformatics resource funded by NIH's National Institute of Allergy and Infectious Diseases; and is a longtime member of the scientific leadership team within LANL's Bioscience Division. He came to LANL and the DOE Joint Genome Institute from the Whitehead Institute/MIT Center for Genome Research (now the Broad Institute) where he retrained in computer science by taking night classes at Boston University. He became a software architecture professional from the Software Engineering Institutes at Carnegie Mellon University in 2005 and has more than 10 years of experience in software engineering. Brettin has taught computer science at the University of New Mexico, Los Alamos, since fall 2000 and has received several distinguished performance awards at LANL.

C. Robin Buell

Michigan State University

C. Robin Buell is an associate professor of plant biology at Michigan State University in East Lansing, Michigan. Buell joined Michigan State from the Institute for Genomic Research in Rockville, Maryland, where she was on the faculty for nearly 9 years. She has been involved in the genome sequencing of *Arabidopsis*, rice, and potato and led the sequencing effort for *Pseudomonas syringae*. Her current research focuses on plant and plant-pathogen genomics. Research projects in her group include annotation of the rice genome, potato sequencing and annotation, comparative sequencing of *Pythium ultimum*, and development of a comprehensive database for plant-pathogen genome sequences. Components of these projects—which are funded through several federal grants—include the generation of public resources such as large-scale sequence and annotation data, as well as bioinformatics resources like databases and data-mining websites for the greater scientific community.

Dylan Chivian*Lawrence Berkeley National Laboratory*

Dylan Chivian received his doctorate from the University of Washington where he worked with David Baker on methods for protein structure prediction, including the creation of the Robetta server for protein structure prediction. He conducted his postdoctoral work with Adam Arkin at Lawrence Berkeley National Laboratory (LBNL) where he studied environmental and comparative genomics of bacteria and archaea, discovering the first single-species ecosystem deep within the Earth. Chivian is currently a scientist at LBNL where he leads the bioinformatics team for the Department of Energy Joint BioEnergy Institute, studying ways of engineering microbes to adopt capabilities ordinarily accomplished by communities in nature.

Bob Cottingham*Oak Ridge National Laboratory*

Bob Cottingham is one of the pioneers of bioinformatics. In the 1970s, he began his career as a software developer on some of the first programs for genetic linkage analysis applied to mapping human disease traits. In 1989, Cottingham became directeur informatique at the Centre d'Etude du Polymorphisme Humain (CEPH) in Paris. There he oversaw the database of CEPH family genotypes, a resource ultimately used by more than 1000 laboratories in an international consortium to construct the first genetic maps of the human genome. Cottingham then joined the U.S. Human Genome Project, first as codirector of the Informatics Core in the Baylor College of Medicine Human Genome Center, then as operations director of the Genome Database at Johns Hopkins University School of Medicine. Subsequently, he became vice president of computing at Celltech Chiroscience, a biopharmaceutical company in the United Kingdom that develops drugs based on gene targets. In 2000, he cofounded Vizx Labs, a bioinformatics company that developed GeneSifter, the first Web-based gene expression microarray analysis service now used worldwide by hundreds of laboratories. In 2008, Cottingham joined Oak Ridge National Laboratory where he leads the computational biology and bioinformatics group currently working on projects for the Department of Energy's BioEnergy Science Center and Genomics:GTL program.

Terence Critchlow*Pacific Northwest National Laboratory*

Terence Critchlow is the chief scientist and associate division director for scientific data management in the Computational Sciences and Mathematics Division at Pacific Northwest National Laboratory (PNNL). Critchlow earned his bachelor of science degree from the University

of Alberta in 1990; he received his master's and doctorate in computer science from the University of Utah in 1992 and 1997, respectively. Critchlow worked at Lawrence Livermore National Laboratory (LLNL) from 1997 to 2007, spending time as a postdoc, individual contributor, and principal investigator. He led several projects while at LLNL, such as data management efforts supporting the Advanced Simulation and Computing program and several Department of Homeland Security (DHS) programs, including the Biodefense Knowledge Center. Critchlow joined PNNL in April 2007. He currently is the technical group manager for the Scientific Data Management group, is thrust area lead for the Scientific Process Automation area within the Department of Energy's SciDAC Scientific Data Management Center, and is a principal investigator for a DHS S&T data management and analysis project. Critchlow's current research interests are data analysis, data integration, metadata, and large-scale data management.

Brian Davison*Oak Ridge National Laboratory*

Brian Davison is chief scientist for systems biology and biotechnology at Oak Ridge National Laboratory (ORNL), and in fall 2009, he will begin serving as chief scientist for the Department of Energy's (DOE) Genomics:GTL program. Davison is a deputy lead in the recently awarded DOE BioEnergy Science Center (<http://www.bioenergycenter.org>). He also is an adjunct professor of chemical engineering at the University of Tennessee. Davison recently served 2 years as director of ORNL's Life Sciences Division, and he previously was a Distinguished Researcher and BioChemical Engineering Research group leader. In his 24 years at ORNL, Davison has performed biotechnology research in a variety of areas, including bioconversion of renewable resources (e.g., ethanol, organic acids, and solvents); non-aqueous biocatalysis; systems analysis of microbes (cultivation and proteomics); biofiltration of volatile organic compounds; mixed cultures; immobilization of microbes and enzymes; metal biosorption; and extractive fermentations. The theme connecting his work is life at the interface of solid, liquid, and gas phases between biocatalysts and their environments, and this research has resulted in 100 publications and 6 patents. Davison received his doctorate in chemical engineering from the California Institute of Technology and his bachelor's degree in chemical engineering from the University of Rochester.

He led a multilaboratory team that in 1997 received an R&D 100 Award for "Production of Chemicals from Biologically Derived Succinic Acid." Davison also cochaired

the 15th to 26th Symposia on Biotechnology for Fuels and Chemicals, served as editor of *Proceedings in Applied Biochemistry and Biotechnology* from 1994 to 2005, and received the 2006 C.D. Scott Award by the Society of Industrial Microbiology. Davison has served as chairman of the ORNL Institutional Biosafety Committee from 2001 to present, and he was named a fellow in the American Institute for Medical and Biological Engineering in 2006.

Matt DeJongh

Hope College

Matt DeJongh received his doctorate in artificial intelligence from Ohio State University. He worked as a senior software engineer in the bioinformatics software industry before joining the faculty of Hope College in Holland, Michigan, where he is an associate professor of computer science. DeJongh is active in bioinformatic research with undergraduate students at Hope College in reconstructing and modeling cellular metabolic systems.

Patrik D'haeseleer

Lawrence Livermore National Laboratory

Patrik D'haeseleer received a master's degree in electrical engineering from Ghent University in Belgium, a master's in computer science from Stanford University, and a doctorate in computer science from the University of New Mexico. His research includes metabolic and regulatory networks, large-scale comparative genomics, and metagenomics. D'haeseleer currently is a research scientist in the Microbial Systems Biology Group at Lawrence Livermore National Laboratory. He also is part of the microbial communities team at the Department of Energy Joint BioEnergy Institute where he studies metabolic processes in natural biomass-degrading microbial organisms and communities.

Tim Donohue

University of Wisconsin, Madison

Tim Donohue has a bachelor of science degree from Polytechnic Institute of Brooklyn and earned a master's degree and doctorate from Pennsylvania State University in 1977 and 1980, respectively. Donohue has been a member of the bacteriology department at the University of Wisconsin, Madison, for more than 20 years. In 2007, he was named director of the Department of Energy's Great Lakes Bioenergy Research Center.

Scott Elliott

Los Alamos National Laboratory

Scott Elliott began his career as a laboratory marine chemist then shifted to atmospheric photochemistry and aerosol microphysics modeling in the 1990s. In this role, he participated in regional simulations of megacity and Asian

air pollution and was involved in elucidation of heterogeneous reactions within the Antarctic ozone hole. After joining Los Alamos National Laboratory, Elliott worked on various defense- and security-oriented environmental chemistry projects, including studies of plume composition of boost phase missiles, degradation of nerve agents in urban atmospheres, and hyperspectral infrared imaging for remote identification.

In the late 1990s, the opportunity arose for Elliott to contribute his modeling skills to a team developing an ultrafast, fine-resolution marine general circulation model—the Parallel Ocean Program (POP). Elliott introduced global biogeochemistry modules into the code and now specializes in simulation of geocycling for dissolved, climate-relevant trace gases. Demonstrations have included the computation of total marine distributions for methane, nitrous oxide, nonmethane hydrocarbons, and organohalogenes. Development has progressed farthest with mechanisms for dimethyl sulfide and carbon monoxide, which influence tropospheric cloud nucleus and ozone fields, respectively. Over the past decade, POP has evolved into the core ocean model in the primary U.S. Earth System simulator—the Department of Energy and National Science Foundation's Community Climate System Model (CCSM). Elliott's trace gas mechanisms are now running within CCSM in a coupled surface ocean-to-atmosphere mode, both for preindustrial and contemporary situations. Simulations of the upcoming period of global warming are now under way, and projects planned for the medium term involve incorporating polar ice algal biogeochemistry and global bacterial population dynamics into the CCSM framework.

Dawn Field

Natural Environment Research Council's Centre for Ecology and Hydrology

Dawn Field is head of the Molecular Evolution and Bioinformatics section of the Natural Environment Research Council's (NERC) Centre for Ecology and Hydrology. She is principal investigator on a NERC project to develop a new genomic data standard to capture a richer set of information about genome sequences. She also is principal investigator on a NERC-funded effort to understand the evolution and function of microsatellites in microbial species. One outcome of this project thus far is Msatfinder, a Perl script designed to allow the identification and characterization of microsatellites in a comparative genomic context. In addition, Field is coinvestigator on the Marine Metagenomics project, which is being undertaken by an integrated consortium of United Kingdom microbiologists who will use postgenomics to investigate aquatic microbial

assemblages that control biogeochemical cycles. She also is participating in the Floral Genome Project, which aims to investigate the origin, conservation, and diversification of the genetic architecture of the flower and to develop conceptual and real tools for evolutionary functional genomics in plants. Field is director and founding member of the NERC Environmental Bioinformatics Centre, which provides bioinformatic and data management solutions for environmental genomic research.

Michael (Mick) Follows

Massachusetts Institute of Technology

Michael (Mick) Follows received his doctorate from the University of East Anglia, United Kingdom, in 1991 and is a senior research scientist in the Department of Earth, Atmospheric, and Planetary Sciences at the Massachusetts Institute of Technology. He uses idealized and numerical models to explore and better understand the interactions of ocean circulation, chemistry, and biology that regulate the productivity of the oceans and marine biogeochemical cycles of key elements, including carbon. His recent work focuses on the relationship of marine microbial communities and their environment.

Peg Folta

Lawrence Livermore National Laboratory

Peg Folta is the associate department head for Computing in Biology at Lawrence Livermore National Laboratory. Bioinformatics and computational biology research within the department are focused primarily on bioenergy and biodefense. Large-scale genomic and proteomic analyses are designed to predict function, identify and characterize unique regions, and determine metabolic pathways. Large-scale data integration and automated high-throughput sequencing also are emphasized. In recent years, Folta was the interim department head at the Department of Energy's Joint Genome Institute and was leader of the computational biology thrust area within the Chemical and Biological Countermeasures Program. She received her master's degree in applied mathematics from the University of Missouri, Rolla, and her bachelor of science in mathematics degree at Truman State University.

James K. Fredrickson

Pacific Northwest National Laboratory

James K. Fredrickson specializes in microbial ecology and environmental microbiology. He received a master's degree in soil chemistry in 1982 and a doctorate in soil microbiology in 1984 from Washington State University. Since joining Pacific Northwest National Laboratory (PNNL) in 1985, he has focused his research efforts in subsurface

microbiology and biogeochemistry. Fredrickson has been responsible for laboratory and field research programs investigating the microbial ecology and biogeochemistry of geologically diverse subsurface environments and is recognized nationally and internationally for these efforts. He also has served as subprogram coordinator for the Department of Energy's (DOE) Subsurface Science Program from 1991 to the present. In this role, Fredrickson coordinated the technical aspects of DOE's deep subsurface microbiology subprogram at the national level and assisted DOE program managers in setting programmatic research directions. This subprogram involved more than 15 projects at universities and national laboratories nationwide and focused on multidisciplinary, field-scale research. At the request of DOE, he currently serves as the national coordinator for the *Shewanella* Microbial Cell Project, part of DOE's Genomics:GTL program. Additionally, Fredrickson was appointed chief scientist for GTL in 2005 and serves as a spokesman for the program to the scientific community. He is a senior chief scientist (laboratory fellow, Level VI) within the Biological Sciences Division, Fundamental and Computational Sciences Directorate, at PNNL.

Damian Gessler

National Center for Genome Resources

Damian Gessler earned degrees in biology and mathematics at Beloit College, Wisconsin, and received his doctorate in population genetics from the University of California, Santa Cruz. Gessler's biological expertise is in evolution and population genetics, as studied via computational techniques. He has used these skills to delineate conditions favorable for the evolution of recombination and meiosis and to quantify the rate of Muller's ratchet in populations unable to achieve mutation-selection balance. Gessler continues research in the evolution of recombination. His informatics expertise is in simulation, modeling, and data integration, and he has more than 20 year of experience in computer programming and systems operations. Gessler's recent work focuses on the challenges of integrating data and services from across the Web in a semantic Web architecture. This complements ongoing work to build a new class of data-driven simulation designs aimed at constructing better predictive models.

Stephen Goff

University of Arizona

Stephen Goff received his bachelor's degree in biology from the University of California, Santa Cruz, in 1978; he earned a doctorate from Harvard University in 1985. His graduate research focused on cell and molecular physiology, and his research training involved molecular genetics of bacteria and bacteriophage, molecular biology, and mammalian cellular

physiology. Goff continued research at Biogen Inc. in Cambridge, Massachusetts, and Geneva, Switzerland, and then joined Tuft's Medical School as a research associate where he focused on transcriptional control mechanisms in mammalian cells. In 1997, he shifted his research focus to gene expression in plants at the Plant Gene Expression Center, a collaboration between the U.S. Department of Agriculture and the University of California, Berkeley. Goff continued this research at the Institute for Molecular Biology at the University of Oregon; he then joined Ciba Biotechnology in Research Triangle Park, North Carolina, in 1992 as a senior scientist. After building up a group involved in gene discovery in plant and animal systems, Goff continued his research in gene discovery and function as director of genome technology at the Torrey Mesa Research Institute, a subsidiary of Novartis/Syngenta in San Diego, California, originally funded by the Novartis Foundation. He initiated and led a large effort to improve genomics technologies to better understand both model and crop plants, especially *Arabidopsis* and rice. This effort resulted in Goff being awarded Research Leader of the Year by *Scientific American* magazine in 2002. From 2003 to 2007, he worked with corporate business development at Syngenta as a senior Syngenta Fellow and senior technical analyst. Goff then became science advisor for Syngenta's corn and soybean business and focused on molecular approaches to enhancing yield and understanding hybrid vigor. Goff also advised Syngenta's vegetable business on appropriate scientific targets for vegetable improvement. At the end of April 2008, Goff left Syngenta and joined the iPlant Collaborative (where he currently is project director) at the University of Arizona's BIO5 Institute.

Yakov Golder

Lawrence Berkeley National Laboratory

Yakov Golder joined the Department of Energy Joint Genome Institute (JGI) in April 2007 and oversaw the Informatics Department until fall 2008. Golder received his bachelor's degree in computer science from City College of New York and a master's degree in computer science from the New York Institute of Technology. He has more than 20 years of technical leadership experience at both established and startup companies in the delivery of complex, high-performance software applications for social networking, workflow management, investment management, customer relationship management; and health care. Prior to joining JGI, Golder served as vice president of technology at CNET Networks where he oversaw the engineering organization in the Online Community Division. There he was responsible for the high-performance photosharing website (<http://www.webshots.com>), which boasted more than

1 billion monthly page views. Prior to his work at CNET, Golder was responsible for two critical areas of eBay's complex Web infrastructure: application data persistence and messaging. Golder's experience in designing enterprise-class software systems for the private and public sector builds upon previous efforts in creating both software-as-a-service websites and more traditional software product development in the engineering, industrial automation, and online communities markets.

Ian Gorton

Pacific Northwest National Laboratory

Ian Gorton is the associate division director in the Computational Sciences and Mathematics Division at Pacific Northwest National Laboratory (PNNL). Gorton also serves as the chief architect for PNNL's Data Intensive Computing Initiative. Gorton received his doctorate in computer science from Sheffield Hallam University, United Kingdom, in 1988. Before coming to PNNL, from March 2004 to July 2006 he led software architecture research and development at National Information and Communications Technology Australia in Sydney. Gorton was PNNL's chief architect in Information Sciences and Engineering, a group of more than 200 software developers who created applications that ranged from full-production deployments to advanced research prototypes and demonstrators. Gorton was responsible for infusing component-based development approaches into Information Sciences and Engineering projects, promoting best-practice architecture designs and review processes, acting as technical lead on several key client projects, and pursuing an R&D agenda to develop new infrastructure technology for data integration and content-based messaging. In addition, he holds the position of honorary associate at the School of Information Technologies at the University of Sydney in Australia.

Grant S. Heffelfinger

Sandia National Laboratories

Grant S. Heffelfinger is deputy director for Materials Science and Technology in the Materials and Process Sciences Center at Sandia National Laboratories in Albuquerque, New Mexico. His graduate research in molecular physics led to a doctorate in chemical engineering from Cornell University in 1988. Since that time, Heffelfinger has held various staff and management positions at Sandia. His research achievements include co-inventing the dual control volume molecular dynamics simulation method for modeling diffusion in molecular systems with chemical potential gradients, such as diffusion through membranes. Heffelfinger was the principal author and technical leader for "Accelerating Biology with Advanced Algorithms and Massively Parallel Computing,"

a cooperative research and development agreement between Sandia National Laboratories and Celera Genomics that was signed in January 2001. He also is the principal investigator in the Department of Energy Genomics:GTL project, Carbon Fixation in *Synechococcus* sp.: From Molecular Machines to Hierarchical Modeling, which is developing advanced computational biology tools and prototyping these tools to understand how marine cyanobacteria fix carbon.

Tatiana Karpinets

Oak Ridge National Laboratory

Tatiana Karpinets is a research scientist in the Computer Science and Mathematics Division at Oak Ridge National Laboratory. She also is an adjunct professor in the plant sciences department the University of Tennessee. Karpinets received a master's degree in biophysics from Kharkov State University in Ukraine. From 1991 to 2001, she worked as a research scientist and then as chief scientist on computational and mathematical modeling of biological systems in the All-Russian Scientific Research Institute of Agriculture. In 2002, Karpinets joined the physics department at Wright State University in Dayton, Ohio, as a postdoctoral research scientist. There she worked on two projects: bioinformatics support for toxicogenomics and simulating the interactions of genes, proteins, and metabolites in cell-like entities. Karpinets specializes in bioinformatics, computational biology, biostatistics, and mathematical modeling of biological systems. She has dozens of publications in these areas of research.

Manpreet Katari

New York University

Manpreet Katari is a postdoctoral fellow and manager of bioinformatics at New York University's plant systems biology laboratory. He received his bachelor's degree in biochemistry from State University of New York, Buffalo, in 1996 and his doctorate in genetics from State University of New York, Stony Brook, in 2004. Katari's research interests include systems biology, comparative genomics, and software and database development. Specifically, his research focuses on identifying networks of genes involved in regulating different metabolic pathways and development stages in *Arabidopsis*. Katari uses both computational and experimental methods to solve biological questions. He is participating in several software projects, including VirtualPlant (<http://www.virtualplant.org>), a system containing a set of data integration, analysis, and visualization tools to support plant systems biology investigations, and Vicogenta (VlEwer for COmparing GENomes to *Arabidopsis*, <http://www.vicogenta.org>), a data-mining tool that allows users to simultaneously search sequence databases

for multiple taxa to find closest matches to the *Arabidopsis* genome based on sequence similarity.

Ken Kemner

Argonne National Laboratory

Ken Kemner is leader of the Molecular Environmental Science (MES) Group at Argonne National Laboratory. He received his doctorate in physics from the University of Notre Dame in 1993. A main emphasis during creation and growth of the MES Group has been development of an internationally recognized and integrated multidisciplinary scientific team focused on investigating fundamental biogeochemical questions. Members of the group have expertise in areas such as high-energy X-ray physics, environmental chemistry, environmental microbiology, and radiolimnology. Additional expertise in geomicrobiology, electron microscopy, and X-ray microscopy often is provided by collaborations with scientists outside the MES group.

Kemner's group uses numerous analytical techniques (e.g., inductively coupled plasma atomic emission spectroscopy, high-performance liquid chromatography, ion chromatography, kinetic phosphorescence analysis, X-ray diffraction, and electron microscopy) to better understand the role of minerals, microbes, and microbial exudates in determining carbon and contaminant mobility and fate in the environment. The group also uses and develops several synchrotron-based X-ray techniques to advance scientists' understanding of processes occurring at physical, geological, chemical, and biological interfaces that determine fate and transport. Kemner's group has begun integrating metagenomic sequencing and bioinformatic approaches to understand microbial community evolution during biostimulation of terrestrial environments.

Les Klimczak

Great Lakes Bioenergy Research Center

Les Klimczak is the chief informatics officer at the Department of Energy's Great Lakes Bioenergy Research Center (GLBRC). Prior to his work at GLBRC, Klimczak was a research informatics consultant at several biotechnology and research organizations. He served as senior director of bioinformatics and information technologies at Psychiatric Genomics Inc., a genomics-based drug discovery company that develops and creates small-molecule drugs for the treatment of psychiatric diseases. Klimczak also was program coleader of bioinformatics at Akkadix Corporation, an agricultural biotechnology company that uses functional genomics, bioinformatics, and other approaches to discover novel plant genes and agrochemicals. In addition to bioinformatics and information technology, Klimczak's

expertise includes cheminformatics, genomics, data mining, databases, statistics, biotechnology startups, knowledge management, laboratory information management systems, biomedical research, and biofuels. He was educated at the University of Würzburg in Germany.

Cheryl Kuske

Los Alamos National Laboratory

Cheryl Kuske has a doctorate in plant pathology and molecular biology and 27 years of research experience in microbial ecology, plant-microbe interactions, and pathogen epidemiology. Her professional experience has included positions in academic, industrial, and national laboratory settings. Over the past 15 years, Kuske has developed and applied molecular methods to study microbial communities and their functions in the environment. Her research portfolio while at Los Alamos National Laboratory (LANL) has focused on two goals: (1) understanding the diversity, structure, and functions of soil microbial communities with applications to Department of Energy missions in climate change, carbon cycling, and environmental remediation and (2) developing technology for rapid, accurate detection of pathogens in the environment and understanding their ecology when not associated with a host. Kuske has published about 50 peer-reviewed manuscripts and 14 LANL unclassified reports and holds 4 patents. She has mentored 9 postdoctoral scientists and more than 30 undergraduate and graduate students.

Mary Lipton

Pacific Northwest National Laboratory

Mary Lipton is a senior scientist in systems biology at Pacific Northwest National Laboratory where she specializes in mass spectrometry and ultrasensitive approaches for globally and quantitatively monitoring gene product expression at the protein level. She received her bachelor's degree in chemistry from Juniata College in 1988 and her doctorate in biochemistry from the University of South Carolina in 1993. She has additional research expertise in Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) for biological research; proteomics of *Yersina pestis*; *Rhodospseudomonas palustris* microbial cell; comprehensive analysis of the proteome of *Deinococcus radiodurans*; determination of metal reduction by *Shewanella onedensis*; and direct characterization of DNA damage from ionizing radiation.

Michael Lomas

Bermuda Institute of Ocean Sciences

Michael Lomas received his doctorate in biological oceanography in 1999 from the University of Maryland where he

studied the nitrogen metabolism of marine phytoplankton in response to variable light, and therefore cellular energy, environments. He was a postdoctoral scholar at Horn Point Laboratory in the Harmful Algal Research Group before joining in 2001 the Bermuda Institute of Ocean Sciences' Bermuda Atlantic Time-Series Study (BATS). His primary interest is studying the ecological linkages between phytoplankton functional diversity and nutrient biogeochemical cycling. Lomas currently is involved in several projects, including examining long-term patterns in phytoplankton diversity at BATS and relationships to ocean carbon cycling; investigating linkages among interannual variability in sea ice, phytoplankton diversity, and primary production in the eastern Bering Sea; studying dissolved organic phosphorus utilization by phytoplankton taxonomic groups; and linking phytoplankton diversity to variability in carbon export in the Sargasso Sea and the subarctic North Pacific.

John L. Markley

University of Wisconsin, Madison

John L. Markley is the Steenbock professor of biomolecular structure in the biochemistry department at the University of Wisconsin, Madison. Markley, who received his doctorate from Harvard University, uses biophysical and biochemical approaches, principally nuclear magnetic resonance (NMR) spectroscopy, to investigate the structure and function of proteins. He also is active in the field of metabolomics. Markley is director of both the BioMagResBank, which is the NMR component of the Worldwide Protein Data Bank, and the National Magnetic Resonance Facility at Madison. He is the principal investigator for the Center for Eukaryotic Structural Genomics and is a fellow of both the American Association for the Advancement of Science and the Biophysical Society. Markley is an honorary member (and silver medal recipient) of the Nuclear Magnetic Resonance Society of Japan and has authored more than 400 research publications, mainly in the field of structural biology.

Cheryl Marks

National Cancer Institute

Cheryl Marks is associate director of the Division of Cancer Biology at the National Cancer Institute where she also serves as director of the Mouse Models of Human Cancers Consortium Program.

Celeste Matarazzo

Lawrence Livermore National Laboratory

Celeste Matarazzo is a computer scientist at Lawrence Livermore National Laboratory where she is participating in the the Advanced Simulation and Computing (ASCI) program. Matarazzo has more than 15 years of experience

in software development and is a research program manager and leader of the data science research group in the Center for Applied Scientific Computing. Matarazzo also leads the ASCI Scientific Data Management project, which aims to provide intelligent assistance in managing terabytes of complex scientific data through development of data models and tools and integration of databases, storage, networks, and other computing resources. Her previous work experience includes developing software for climate modeling simulations, output devices, and defense applications. Matarazzo has a bachelor's degree in mathematics and computer science from Adelphi University.

Raymond McCord

Oak Ridge National Laboratory

Raymond McCord has been an environmental information manager in the Environmental Sciences Division at Oak Ridge National Laboratory for 21 years. He has managed the development and operation of three major information systems supporting environmental assessment, research, and restoration. McCord also was responsible for establishing a geographic information system within the division. Currently, he is manager of the data archive for the Atmospheric Radiation Measurements Program that supports climate change research. This archive contains 8 million files (~140 TB of storage) about meteorology, solar radiation, and cloud formation. McCord received his doctorate in ecology from the University of Tennessee in 1980.

Lee Ann McCue

Pacific Northwest National Laboratory

Lee Ann McCue received a doctorate in microbiology from Ohio State University. Her research interests focus on comparative genomics, transcription regulation, and the inference of regulatory networks in prokaryotic systems. McCue is a senior research scientist in the Computational Biology and Bioinformatics Group at Pacific Northwest National Laboratory.

Peter McGarvey

Georgetown University Medical Center

Peter McGarvey has 20 years of academic and commercial experience in molecular biology, biotechnology, bioinformatics, and software development. He is interested in genomic and proteomic analysis, biological databases, data integration, and visualization. McGarvey currently is managing the data integration and website for the Biodefense Proteomics Resource, a project of the National Institute of Allergy and Infectious Diseases. He also is a funded participant in the caBIG VCDE (vocabulary and common data elements) workspace and has served as project manager for

several caBIG adopter projects. In addition, McGarvey is active in UniProt consortium activities and databases as a member of the Protein Information Resource. He received a doctorate in biological sciences from the University of Michigan in 1988 and a master's degree in technology management from the University of Maryland University College in 2004.

Folker Meyer

Argonne National Laboratory

Folker Meyer is a computational biologist at Argonne National Laboratory and a senior fellow at the Computation Institute at the University of Chicago. He was trained as a computer scientist, which led to his interest in building software systems. Meyer now focuses on building systems that advance scientists' understanding of biological datasets. In the past, he has been known best for his leadership role in developing the GenDB genome annotation system and designing and implementing a high-performance computing facility at Bielefeld University in Germany. Currently, Meyer is most interested in the comparative analysis of large numbers of microbial genomes. He received his doctorate in bioinformatics from Bielefeld University in 2001.

Bob Morris

University of Washington

Bob Morris is an assistant professor of biological oceanography at the University of Washington School of Oceanography. He received his doctorate in microbiology from Oregon State University in 2004. Morris' research interests are marine microbial ecology, bacterioplankton physiology, and microbial community interactions. His laboratory uses cultivation, genomic, and proteomic approaches to study relationships between biogeochemical cycles and microbial processes in the oceans. Morris is specifically interested in exploring the diversity and metabolism of dominant, uncultured bacterioplankton.

Sean Murphy

J. Craig Venter Institute

Sean Murphy is a software engineer at the J. Craig Venter Institute (JCVI) where he develops enterprise software applications to support bioinformatic research. He currently is a member of the Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis (termed CAMERA) team. With an extensive software background, including database design, middle-ware architecture, asynchronous messaging systems, grid-computing interfaces, model-view controllers, and website design, Murphy is experienced in developing software systems end to end. Before joining JCVI, he worked at Celera

Genomics and Intelligent Medical Imaging. Murphy has developed several products, including the Moore Microbial Genome website, the JCVI Blast Server, the JCVI resequencing primer designer, the Celera Gene Index pipeline, and machine-vision algorithms for automated pathology applications. He has bachelor's degrees in electrical engineering and biology from the Massachusetts Institute of Technology, and he earned a doctorate in neuroscience from Yale University.

Ilya Nemenman

Los Alamos National Laboratory

Ilya Nemenman received his doctorate in theoretical physics, specializing in biophysics, from Princeton University. He completed extra postdoctoral training at the NEC Research Institute and Kavli Institute for Theoretical Physics at the University of California, Santa Barbara. Nemenman was a member of the research faculty at Columbia University Medical School. In 2005, he joined the Computer, Computational, and Statistical Sciences Division at Los Alamos National Laboratory. His research interests focus on information processing in biological systems, from neural assemblies to molecular signaling and regulatory pathways.

Gary J. Olsen

University of Illinois

Gary J. Olsen is a microbiology professor at the University of Illinois, Urbana-Champaign. He received a bachelor's degree in physics at the University of California, Los Angeles, in 1975; a master's in physics from UCLA in 1976; and a doctorate in biophysics from the University of Colorado Health Sciences Center in 1983. Olsen conducted postdoctoral work in molecular and cellular biology at the National Jewish Hospital and Research Center (1983–84) and in biology at Indiana University (1984–1985). He also was an assistant scientist in biology at Indiana University from 1985 to 1988. Most of Olsen's current research focuses on two areas: (1) gene expression in archaea and its relation to corresponding systems in eucarya and bacteria and (2) genomics, with an emphasis on comparative genomics and genome evolution. His approach combines experimental work and computational analyses of genomes and proteins.

Andrei Osterman

Burnham Institute for Medical Research

Andrei Osterman is an associate professor in the Bioinformatics and Systems Biology Program at the Burnham Institute for Medical Research (BIMR). He received his doctorate in biochemistry at Moscow University in Russia. In 1993, Osterman joined the laboratory of Meg Phillips at the

University of Texas in Dallas to perform structure–functional studies of the ornithine decarboxylase enzyme family. Recognizing the new frontiers of metabolic biochemistry and enzymology enabled by the genomics revolution, Osterman joined Integrated Genomics, a startup biotechnology firm in Chicago in 1999. As a director and vice president of research at Integrated Genomics, he pioneered integration of comparative genomics with biochemical and genetic experiments for gene and pathway discovery. His research team published the first genome-scale study of gene essentiality in *Escherichia coli* by genetic footprinting. Osterman is one of the founders of the Fellowship for Interpretation of Genomes (FIG), a nonprofit research organization that launched the Project to Annotate 1000 Genomes in 2003. FIG provides the open-source integration of all publicly available genomes and tools for their comparative analysis, annotation, and metabolic reconstruction. Osterman's laboratory at BIMR focuses on fundamental and applied aspects of key metabolic subsystems in a variety of species, from bacteria to human. His group applies bioinformatic techniques followed by experimental validation to reconstruct metabolic pathways from genomic data, reveal gaps in current knowledge, and identify previously uncharacterized (missing) genes. The power of this integrative approach is illustrated by the discovery and characterization of more than 20 enzyme families in the metabolism of cofactors, carbohydrates, and amino acids. Most applications pursued by this group are related to pathogenic and environmental bacteria. New research directions include the analysis of regulatory networks and the application of proteomics and metabolomics technology for identification of novel diagnostic and therapeutic targets in cancer.

Ross Overbeek

Argonne National Laboratory

Ross Overbeek received a doctorate in computer science in 1972 from Penn State University. For 11 years, Overbeek taught mathematics and computer science at Northern Illinois University where his research focused on computational logic and database systems. From 1983 to 1998, Overbeek worked at Argonne National laboratory (ANL), focusing on parallel computation and logic programming. Overbeek collaborated with Carl Woese and helped in the founding of the Ribosomal Database Project. Overbeek participated in analysis of *Methanococcus jannaschii*, the first archaeal genome to be completely sequenced. He was the lead architect of the PUMA and WIT genomics database systems at ANL before becoming a founder of Integrated Genomics (IG) in 1998. While at IG, Overbeek participated in the sequencing and analysis of more than 50

genomes and led the company's bioinformatics effort. The most significant product was ERGO, a system to support comparative genomic analyses. In mid- 2003, Overbeek left IG to become a founding fellow of the Fellowship for Interpretation of Genomes (FIG). His efforts at FIG centered on building the SEED, an open-resource data curation system to facilitate comparative analyses of genomic data. Since 2004, Overbeek has been a coprincipal investigator of the National Microbial Pathogen Data Resource, a framework to support comparative analysis of pathogen genomes.

George N. Phillips, Jr.

University of Wisconsin, Madison

George N. Phillips, Jr. received his doctorate in biochemistry at Rice University in Houston, Texas, in 1976. He currently is professor of biochemistry and computer sciences at the University of Wisconsin, Madison. Phillips leads the Computation and Informatics in Biology and Medicine training program supported by the National Library of Medicine. He also is coinvestigator at the Center for Eukaryotic Structural Genomics and serves as the informatics and information technology manager at the Department of Energy's Great Lakes Bioenergy Research Center. Phillips' research interests are computational and structural biology.

David Pletcher

Joint BioEnergy Institute

David Pletcher is the director of informatics at the Department of Energy's (DOE) Joint BioEnergy Institute (JBEI). Prior to joining JBEI, Pletcher worked for nearly 7 years as a computer scientist at Lawrence Livermore National Laboratory. He also served as group leader of production informatics at the DOE Joint Genome Institute from June 2004 to August 2008. Pletcher began his career in the private sector, working as a programmer and software engineer and developer for several companies, including Rockwell Scientific, Lumisys, Visual Edge Technology, and idrive.com. He graduated from Harvey Mudd College in 1992.

Jennifer Reed

University of Wisconsin, Madison

Jennifer Reed is an assistant professor in the chemical and biological engineering department at the University of Wisconsin, Madison. She received her bachelor's and master's degrees as well as her doctorate from the University of California, San Diego. Most of Reed's research interests involve studying metabolism and regulation through the generation and subsequent analysis of metabolic models and reconciling the models with experimental data. Overall, her research group uses computational models and develops methods to study biological systems, engineer cells, and

expand scientific knowledge of the mechanisms underlying observed cellular behavior. The group specifically is interested in building, analyzing, and using metabolic and regulatory models of organisms involved in environmental remediation, biofuels, and pharmaceutical applications. Reed's laboratory also uses the developed models to identify novel gene functions or regulatory interactions. In addition to model building, her research involves computational methods for designing strains or cell lines with enhanced production yields of desired products.

Nagiza F. Samatova

Oak Ridge National Laboratory

Nagiza F. Samatova is a senior research scientist in the Computational Biology Institute, Computer Science and Mathematics Division, at Oak Ridge National Laboratory. She received her bachelor's degree in applied mathematics from Tashkent State University in Uzbekistan in 1991 and her doctorate in mathematics from the Russian Academy of Sciences in Moscow in 1993. Samatova also obtained a master's in computer science in 1998 from the University of Tennessee. She specializes in computational biology and high-performance data mining, knowledge discovery, and statistical data analysis. Samatova is the author of more than 50 publications, including 1 book, and she holds 2 patents.

Denise Schmoyer

Oak Ridge National Laboratory

Denise Schmoyer is a research staff member in the Computer Science and Mathematics Division at Oak Ridge National Laboratory. Schmoyer has worked on the design and development of several large-scale database systems for human, model organism, and microbial sequence annotation. She is the primary developer of a laboratory information management system and database for protein complexes in microbial organisms.

Blake A. Simmons

Joint BioEnergy Institute

Blake A. Simmons received a bachelor's degree in chemical engineering in 1997 from the University of Washington and attended graduate school at Tulane University where the focus of his thesis work was the synthesis and characterization of templated nanomaterials. He earned a doctorate in chemical engineering from Tulane in 2001. Simmons then joined Sandia National Laboratories in Livermore, California, as a senior member of the technical staff, working in the Materials Chemistry Department. He participated in and led various projects, including the development of cleavable surfactants, enzyme engineering for biofuel cells, microfluidics, and the synthesis of

silicate nanomaterials. In 2004, Simmons was promoted to principal member of the technical staff. He expanded his research portfolio to include the design, fabrication, integration, and testing of polymeric microfluidic devices for several lab-on-a-chip and homeland security applications. He also continued to pursue opportunities in renewable energy. In 2006, Simmons was promoted to manager of the Energy Systems Department, which focuses on developing novel, materials-based solutions to meet the United States' growing energy demands. In 2007, Simmons was named one of the principal coinvestigators of the Joint BioEnergy Institute (JBEI, <http://www.jbei.org>), a \$135 million project funded by the Department of Energy and tasked with developing next-generation biofuels produced from non-food crops. He currently is serving as vice president of the Deconstruction Division at JBEI where he leads a team of 35 researchers working on advanced methods of liberating fermentable sugars from lignocellulosic biomass. He also manages the Biomass Science and Conversion Technology Department at Sandia. Simmons has authored more than 70 publications, book chapters, and patents.

Tom Slezak

Lawrence Livermore National Laboratory

Tom Slezak has been involved with bioinformatics at Lawrence Livermore National Laboratory (LLNL) for more than 28 years. He received his bachelor's degree in computer science from the University of San Francisco and his master's in computer science from the University of California, Davis. Slezak participated in the Human Genome Project from its inception and led the informatics efforts at LLNL and then the Department of Energy's Joint Genome Institute from 1987 to 2000. In 2000, Slezak began to assemble a pathogen bioinformatics team at LLNL, pioneering a whole genome analysis approach to DNA signature design. His team developed signature targets to identify multiple human pathogens, and these targets were used as part of the biodefense measures at the 2002 Winter Olympic Games under the BASIS program. They later were adapted for use nationwide as part of the Centers for Disease Control and Prevention's (CDC) BioWatch program. Slezak's bioinformatics team has developed DNA-based signatures of virtually every biothreat pathogen (the organisms identified by CDC as high-priority threat agents) for which adequate genomic sequences are available, as well as of several other human and livestock pathogens. LLNL signatures are part of the nation's public health system and have been in use for homeland defense since fall 2001.

Rick Stevens

Argonne National Laboratory

Rick Stevens is associate laboratory director for Computing, Environment, and Life Sciences at Argonne National Laboratory and professor of computer science at the University of Chicago. His research interests are high-performance computer architectures and computational science, especially challenges in the life sciences. Stevens leads Argonne's efforts in advanced computing that target the development of exascale computing technology and applications in systems and computational biology and environmental modeling and simulation. He is a fellow of the American Association for the Advancement of Science and is also a cofounder and senior fellow of the Argonne and University of Chicago Computation Institute, a multidisciplinary institute aimed at connecting computing to all areas of inquiry at the university and laboratory.

Michael R. Sussman

University of Wisconsin, Madison

Michael R. Sussman has been a faculty member at the University of Wisconsin, Madison, for the past two decades. During that time, he has become recognized as a leading expert on signal transduction and genomics in plants. Sussman's research interests have focused on using the model higher-plant *Arabidopsis thaliana* for understanding the role of plasma membrane proteins in signal transduction and solute transport. His laboratory was the first to report on unique protein kinases found only in plants and protists and on the plasma membrane proton pump, which provides the driving force for the uptake of all nutrients. To help understand the in situ role played by these important proteins, Sussman's laboratory pioneered the development of genome-wide reverse genetics techniques. Specifically, the lab used an insertional mutagenesis scheme to isolate "knockout" plants, starting with the sequence for any one of the roughly 30,000 genes in *Arabidopsis*. For example, Sussman's laboratory was the first to demonstrate that the plant homologue for a brain potassium channel performs a nutritional role in plants (i.e., is responsible for the uptake of potassium from soil). Similar studies have been performed to identify the in planta roles of several plasma membrane hormone receptors.

In 1999, Sussman, together with colleague Franco Cerrina, a professor in the College of Engineering, developed a new instrument known as a MAS (Maskless Array Synthesizer), which makes "gene chips" that can analyze hundreds of thousands of genes at once. MAS is unique because it eliminates the requirement for expensive masks used in traditional DNA

chip technology, thus making MAS accessible to all scientists. Based on the MAS technology, Sussman and Cerrina founded in 2000 a biotechnology startup company, Nimble-Gen Systems Inc., which after 8 years, was sold to Roche Inc.

Sussman's awards have included a Fulbright research fellowship for a sabbatical in Belgium; a McKnight Foundation award; a University of Wisconsin, Madison, WARF Kellett Mid-Career Award; and selection as a fellow to the American Association for the Advancement of Science. In 1996, Sussman was appointed director of the UW Biotechnology Center (UWBC), a campus-wide facility devoted to research, outreach, and service in the area of biotechnology and genomic science and instrumentation.

Ronald Taylor

Pacific Northwest National Laboratory

Ronald Taylor earned a doctorate in bioinformatics from George Mason University in Fairfax, Virginia. He received his bachelor's degree in physics, master's degree in computer science, and master's in biology from Case Western Reserve University in Cleveland, Ohio. Taylor is a research scientist at Pacific Northwest National Laboratory (PNNL) where he develops algorithms and software for inference of biological networks. He also is involved in the development of biological databases, leading one such project at PNNL.

Gerald A. Tuskan

Oak Ridge National Laboratory

Gerald A. Tuskan is a distinguished scientist in the Plant Genomics Group within the Environmental Sciences Division at Oak Ridge National Laboratory (ORNL) where he coordinates the Department of Energy's (DOE) effort to sequence the *Populus* genome. He received a master's degree in forest genetics from Mississippi State University in 1980 and a doctorate in genetics from Texas A&M University in 1984. In addition to his work at ORNL, Tuskan is involved in the laboratory science program for the DOE Joint Genome Institute (JGI). In this role, he coordinates the solicitation and review of principal investigator-led sequencing proposals submitted through the DOE laboratory system; helps establish multiple large-genome sequencing projects that address DOE missions in biofuel development, carbon biosequestration, and global climate change; and facilitates DOE, laboratory, and JGI interactions. Tuskan also is a research professor in the entomology, plant pathology, and plant sciences departments at the University of Tennessee. His research interests include understanding the genetic basis of tree growth and development with emphasis on biomass accumulations, carbon allocation, and cell-wall chemistry;

Populus genomics, including assembly of the draft sequence, comparative genomics, and functional gene identification; and short-rotation woody crop silvicultural systems.

Edward C. Uberbacher

Oak Ridge National Laboratory

Edward C. Uberbacher received his bachelor's degree in chemistry from Johns Hopkins University in 1974 and a doctorate in physical chemistry from the University of Pennsylvania in 1979. Beginning in 1980, he conducted postdoctoral studies in the University of Pennsylvania's biophysics department and in the Biology Division of Oak Ridge National Laboratory (ORNL) and the University of Tennessee's Graduate School of Biomedical Sciences, investigating the structure and function of genetic materials using crystallography, electron microscopy, and computational modeling. In 1985, Uberbacher became an investigator at the Center for Small-Angle Scattering Research at ORNL, pursuing structural and dynamic studies of macromolecules in solution using techniques involving neutron and X-ray scattering and molecular modeling. In 1987, he also became a research assistant professor at UT's Graduate School of Biomedical Sciences and an investigator in the ORNL Biology Division, focusing on X-ray and neutron crystallography, scattering, and other biophysical methods. Uberbacher became a consultant in the ORNL Engineering Physics and Mathematics Division in 1988 to develop artificial intelligence and high-performance computing methods for genomic DNA sequence analysis; in 1991, he joined the staff of the Computer Science and Mathematics Division as the informatics group leader. In this role, he received an R&D 100 Award for developing the GRAIL DNA sequence analysis system. In 1997, Uberbacher became the head of ORNL's Computational Biology Section in Life Sciences and a codeveloper of the PROSPECT computational protein fold prediction system, which received an R&D 100 Award in 1998. Uberbacher performed part-time duties as an IPA in 2003–04 for the Department of Energy's Office of Biological and Environmental Research, contributing extensively to the Genomics:GTL computing roadmap. He is currently the program leader for Computational Biology at ORNL and is an adjunct professor in the Genome Science and Technology Program at the University of Tennessee. His scientific interests include the application of pattern recognition; artificial intelligence; concurrent processing techniques and algorithm development for computational biology; computational genome sequence analysis; mass spectrometry analysis; and macromolecular structure, dynamics, and docking.

Owen White*University of Maryland, College Park*

Owen White, professor of epidemiology and preventive medicine, is the director of bioinformatics at the University of Maryland School of Medicine. He received his doctorate in molecular biology from New Mexico State University in 1992 and is an internationally recognized expert in bioinformatics. He is the principal investigator of the Data Analysis and Coordination Center (funded by the National Human Genome Research Institute) of the Human Microbiome Project, a National Institutes of Health Roadmap Initiative. In this capacity, White is responsible for coordinating the collection, integration, standardization, analysis, and distribution of all genomic and metagenomic data related to the Human Microbiome Project. At the Institute for Genome Sciences (IGS), he leads a group of more than 20 scientists and engineers who collectively are responsible for developing nearly all IGS production-level annotation pipelines, database systems, and tools for automated and manual annotation of genomes and metagenomic datasets. White has experience in DNA sequence generation and genomic analysis of human expressed sequence tags, other eukaryotes, and prokaryotes as well as in comparative analyses.

John Wooley*University of California, San Diego*

John Wooley is associate vice chancellor for research; professor of pharmacology, chemistry, and biochemistry; and director of digitally enabled genomic medicine at the University of California, San Diego. He also is chief scientific officer of the metagenomics cyber-resource project termed CAMERA at the university. This infrastructure project focuses on linking environmental metadata to molecular data and on the development and provision of software tools in a rich computing environment to probe metagenomic data and advance microbial ecology. Wooley's current research involves bioinformatics and structural biology focused on protein structure-function relationships. He is co-principal investigator of the Joint Center for Structural Genomics, a high-throughput structural pipeline. For nearly three decades, Wooley has been focused on nurturing the interface between computing and biology. He received his doctorate in 1975 at the University of Chicago.

Alex Worden*Monterey Bay Aquarium Research Institute*

Alex Worden is a microbiologist at the Monterey Bay Aquarium Research Institute (MBARI). She earned a bachelor's degree in history from Wellesley College, with a concentration in Earth, atmospheric, and planetary sciences at the Massachusetts Institute of Technology (MIT).

Worden remained at MIT for 2 years as a research technician and then joined the University of Georgia where she received a NASA Earth systems science fellowship and completed her doctorate in ecology in 2000. Worden spent 3.5 years conducting postdoctoral research on microbial interactions at the Scripps Institution of Oceanography. She then accepted an assistant professorship at the Rosenstiel School of Marine and Atmospheric Science at the University of Miami. In 2007, Worden joined MBARI where she leads a microbial ecology research group. Her research interests include population regulation of photoautotrophic microbes, with an emphasis on carbon cycling in marine systems. Worden's laboratory uses a range of methods and technologies, from seagoing oceanography to genomics and metagenomics.

Cathy Wu*Georgetown University*

Cathy Wu is a professor in the biochemistry and molecular and cellular biology department at Georgetown University's School of Medicine. She also is a professor in the oncology department and is director of the Protein Information Resource (PIR) at Georgetown University Medical Center. Wu has master's degrees in plant pathology and computer science and received her doctorate in molecular plant pathology from Purdue University in 1984. She has conducted bioinformatic research since 1990 and has developed several protein classification systems and databases. Wu has managed large software and database projects and has led the bioinformatics effort of PIR since 1999, becoming PIR director in 2001. Her research interests include protein family classification and functional annotation, biological data integration, and literature mining.

Liming Yang*National Center for Research Resources*

Liming Yang is a health scientist administrator in the Biomedical Technology Division of the National Center for Research Resources (NCRR) within the National Institutes of Health (NIH). He manages a portfolio of grants on computational biology, software development, and genetic studies. Before joining NCRR, Yang was associate director of biomedical informatics from 2005 to 2008 at the Center for Bioinformatics within the National Cancer Institute. He led several projects to build bioinformatics infrastructures to support large genomics and proteomics initiatives. Prior to that position, Liming was an intramural scientist at NIH where he played an important role in data analysis and management for the multi-institute Lymphoma and Leukemia Molecular Profiling Project. Yang received his doctorate in pathology from the University of Utah School

of Medicine. After that, he spent 2 years as a postdoctorate fellow at NIH. Yang is from Beijing, China, where he attended Peking University for undergraduate studies and Peking Union Medical College for medical school.

William S. York

University of Georgia

William S. York received his bachelor's degree in molecular, cellular, and developmental biology in 1978 from the University of Colorado and his doctorate in biochemistry and molecular biology in 1996 from the University of Georgia. York was senior research chemist at the Complex Carbohydrate Research Center from 1985 to 1996 before beginning his faculty career at the University of Georgia. His diverse research interests include the development and application

of spectroscopic and computational methods for structural and conformational analysis of complex carbohydrates, development of bioinformatic tools to study the roles of carbohydrates in living systems, and the use of these tools to create realistic models describing the assembly and morphogenesis of the walls surrounding the cells of higher plants. York's current research includes the application of these techniques to understand the recalcitrance of biomass to saccharification. Results of this research may lead to improvement of feedstocks for the biofuel industry. His research is supported by the Department of Energy, the National Science Foundation, the National Institutes of Health, and the University of Georgia Research Foundation.