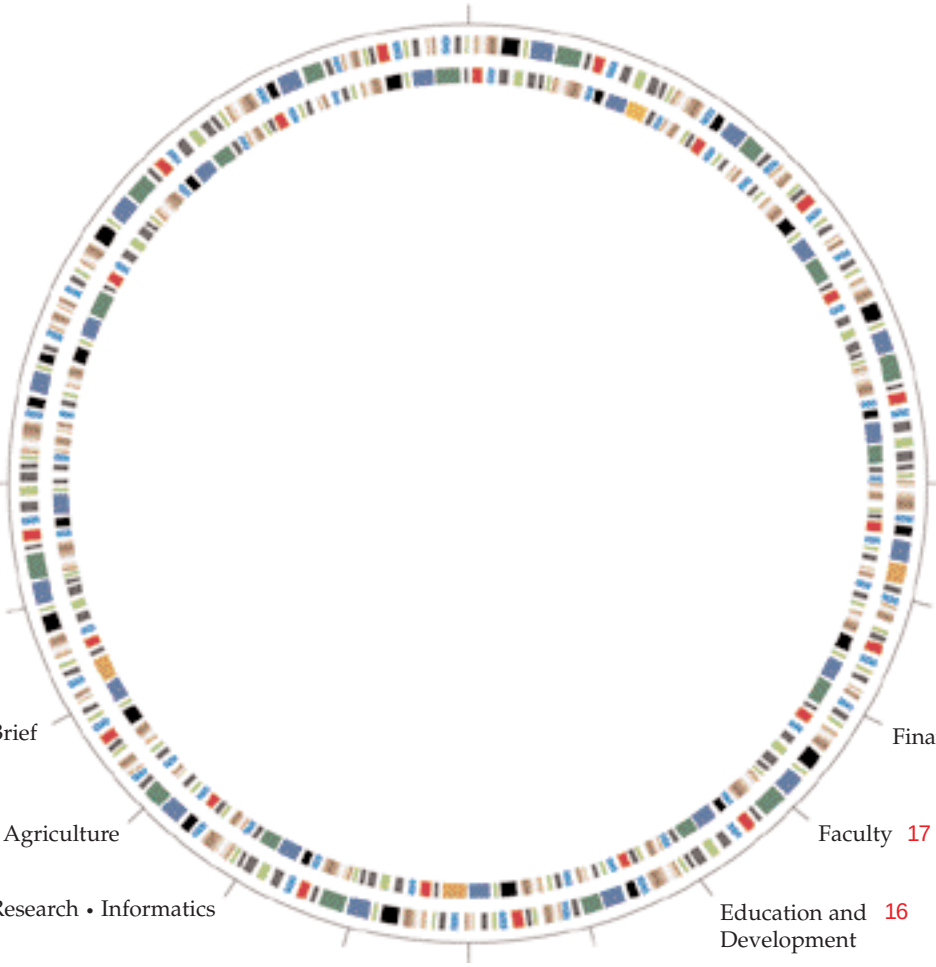
Contents



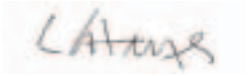
Message from the President

Throughout the ages, light has symbolized knowledge, education, and enlightenment. I'm pleased to report that the "lamp of knowledge" shines ever more brightly at Virginia Tech as exemplified by Virginia Bioinformatics Institute's substantial progress in its first two years of operation. This document highlights the progress being made in bioinformatics research and education—progress directed toward achieving Virginia Tech's goal of becoming a top-30 research university by this decade's end. In the years to come, Virginia Bioinformatics Institute (VBI) will continue to enhance Virginia Tech's stature as an internationally recognized research university, as indicated by competitively awarded extramural support, the quality of VBI's faculty, and the expansion of bioinformatics graduate education programs.

At Virginia Bioinformatics Institute, the promise for a better world drives the focus on interdisciplinary research and collaboration. Through this connectivity, we see not by the light of our own lamps, but by the greater illumination that results when knowledge is shared, refined, and expanded within the larger society. Virginia Tech is pleased to make this unique bioinformatics research platform available to all stakeholders in industry, government, and academe. The dissemination of new knowledge is truly vital in forging a global community.

All who labor and learn at Virginia Bioinformatics Institute will, in effect, bask in the glow of the lamp of knowledge. We have embarked on a path that will enable us to meet and surpass the expectations of those whose lives are touched by the ever-brightening radiance of Virginia Tech. We are pleased to share our results and vision espoused in this annual report. We are certain that VBI's role will continue to expand and, if VBI's history is any predictor, we look forward to what the coming years have to offer.

Regards,



Dr. Charles Steger
President, Virginia Polytechnic Institute and State University

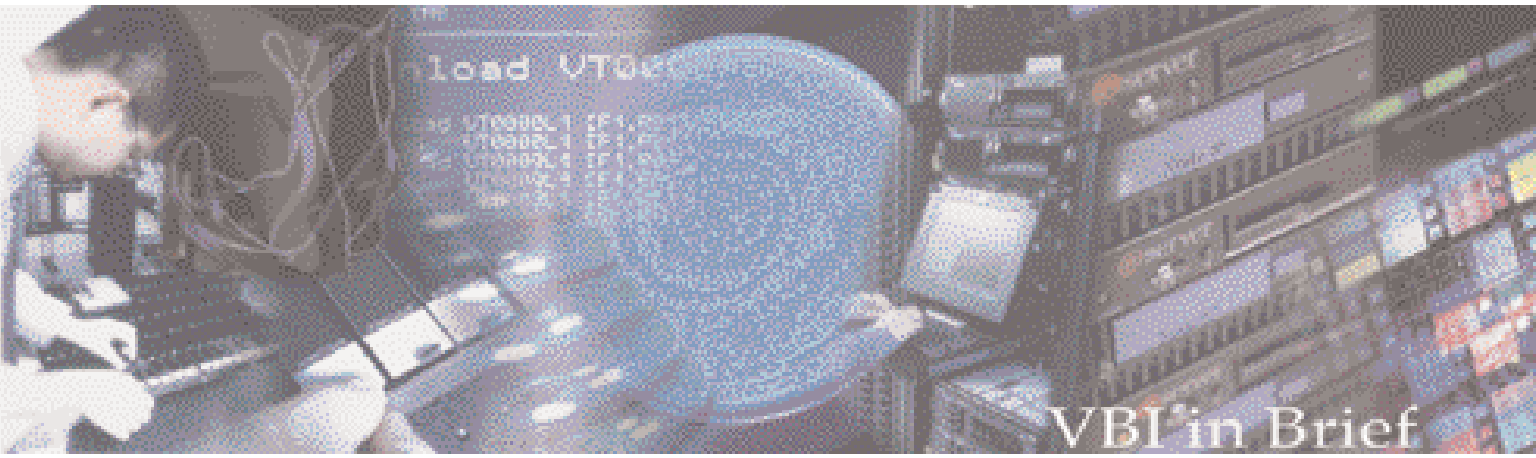
Letter from the Director

Virginia Bioinformatics Institute strives to embody Virginia Tech's focus on "putting knowledge to work". Bioinformatics—an interdisciplinary science that merges information technology and biology—allows scientists to interpret and apply vast amounts of biological data generated from basic research. VBI faculty and staff have been extremely productive as shown by our ever- increasing grant portfolio and many research collaborations. Our research staff has increased by an order of magnitude, joining to find cures for diseases like malaria, brucellosis, and tuberculosis; create high-yield, insect- and disease-resistant crops; and provide information and scientific tools supporting further discoveries.

Our second year marks an important milestone for VBI. We have settled into our new home in Building XV of the Corporate Research Center, and our Core Laboratory and Core Computational Facilities are fully online to provide a wide range of scientific services to partners in Virginia and beyond. Our faculty researchers have secured over $24 million in funding, and more is sure to come, as we continue to evolve new research projects and expand existing knowledge. VBI is excited to undertake the voyage of discovery that bioinformatics offers us. We hope you will join us in that journey as we continue to venture beyond information in the years ahead.

Regards,

Dr. Bruno Sobral
Director, Virginia Bioinformatics Institute



VBI is a Commonwealth of Virginia shared resource, founded at Virginia Tech in July 2000. In two short years, we have combined our experience and knowledge of the biological sciences with our understanding and expertise in information technology to secure over $24 million in grants and contracts. Our bioinformatics systems merge information technology and life sciences and have laid the groundwork for lasting collaborations with government, industry, and academia. With this unique portfolio, VBI is planting the seeds of promise across Virginia and beyond.

Core Facilities

VBI is comprised of two main core facilities—the **Core Laboratory Facility (CLF)** and the **Core Computational Facility (CCF)**—which provide services to collaborators at Virginia Tech, in Virginia, and worldwide. The synergy of information technology and biotechnology at VBI's core facilities provides the best in bioinformatics research.

Core Laboratory Facility • CLF

The Core Laboratory Facility at VBI is a multi-user resource dedicated to the development and delivery of state-of-the-art technologies for the large-scale discovery of biological macromolecules. Industrial-scale biotechnologies are expensive for small laboratories to acquire and maintain, yet they are of paramount importance to biologists, engineers, mathematicians, statisticians, and computer scientists who wish to remain on the cutting edge of their fields. In the past two years, we have served over 200 clients, many of those from 80 individual labs at Virginia Tech.

The CLF currently focuses on three primary areas of laboratory research: genomics (DNA), transcriptomics (RNA), and proteomics (proteins). Metabolomics services will be added in the coming year. For genomics research, sequencing, fragment analysis, and quantitative real-time PCR are now available. The CLF provides GeneChip® and custom microarray platforms for transcriptomics research. Proteomic analysis capabilities include 2-D gel analysis and spot-picking, robotic handling of the digestion and clean-up steps in protein processing, and MALDI-TOF and LC-MS/MS analysis components. In addition, the CLF will set up instrumentation for the investigation of specific protein interactions, including the elucidation of kinetic parameters.

Dissemination, education, and training related to industrial-scale biotechnologies are important functions of the CLF. Staff members are available to give seminars and workshops on broad views of established and emerging technologies, or particular technologies in-depth. The CLF also provides a test bed for emerging technologies, from sample handling and processing, to data production and data handling steps. Innovation is the major secondary role of the CLF, and VBI welcomes both academic and commercial development partners in this endeavor.

Core Laboratory Facility staff from left to right: Leonard Comaratta, Jennifer Fick, Dr. John Lennon, Amanda McWatters, Nickole Kaufman and Annie Foster.

Proteomics at VBI

Proteomics is the most recent addition to the variety of services available at VBI's CLF. This discipline focuses on the characterization of proteins that are identified as significant from the entire protein profile of a given sample, either from a cell type or organism, under a specific set of conditions. Therefore, gel electrophoresis and mass spectrometry are fundamental, and the CLF is equipped to generate 1-D and 2-D gels using a variety of stains. State-of-the-art gel scanning and image analysis provide the necessary data to drive automated spot excision from a gel. Samples can then be prepared for mass spectrometric analysis by performing automated in-gel protein digestion with the CLF's liquid handling robot. Digested samples are finally introduced into one of two mass spectrometers which provide data that can be used to search relevant databases through VBI's CCF.

Core Computational Facility • CCF

Powerful computers with massive data storage capacity, vizualization capabilities, and high-speed data connections are necessary to achieve the goals of bioinformatics. VBI's Core Computational Facility incorporates computers with multi-terabytes of expandable data storage that use symmetric multiprocessor and parallel processor cluster technologies. Powerful supercomputers, including a Sun Microsystems E15000 and an IBM 204-processor Linux cluster, are the computational heart of the CCF. The facility also includes an IBM Storage Area Network that initially provides 1.7 terabytes each of disk and tape storage, which ensures the integrity, backup capacity, and availability of CCF services.

Visualization systems allow users to view data through various interfaces, either remotely or at VBI. The CCF utilizes gigabit Ethernet as its backbone and has a dedicated, scalable, and high-speed connection to the Internet. With such capacity, researchers can download, manipulate, and visualize data, as well as easily collaborate with fellow scientists.

The CCF offers the following bio-computational services:

• Computational processing – Processing time on both parallel and symmetric multi-processing supercomputers provides computational resources appropriate to a wide variety of application types.

• Data storage and backup – Data storage provides users with reliable disk storage. Robotic tape backup systems ensure data availability.

• Compound services – Multi-service packages are available which combine computational processing, storage, and database administration.

Highly - skilled systems administrators ensure that VBI's CCF runs in a reliable, professional manner. A seamless connection with VBI's CLF ensures that data entering databases accurately reflect laboratory results.

VBI's Core Computational Facility staff provides the vital computing base for bioinformatics research. Staff members include (from left to right): Nick Galloway, Ryan Chase, Dustin Machi, Mark DiFilippo, Ivan Morozov, and Dominik Borkowski.

The Core Computational Facility Database Group

The Core Computing Facility Database Group (CCFDG) works with principal scientific investigators, their staffs, and collaborators to provide the highest quality, lowest cost database design, analysis, development, and maintenance services in support of basic and applied research. Requirements analysis, design, and development are performed using several software tool sets that employ industry recognized database standards. Mission critical systems are created on UNIX and Linux servers hosting both Oracle and IBM DB2 relational database engines. Scientific databases are designed to support each project's complete life cycle. Data are fully protected by an automated suite of Oracle, IBM, and third-party backup and recovery systems. A charter member of the Virginia Tech Database User's Group, the CCFDG manages a Web page, sponsors technical workshops, and routinely coordinates special-interest database issues among participating university staff.

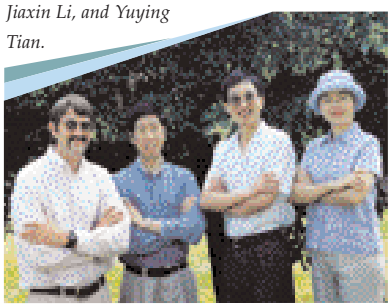
VBI's research efforts center on understanding the "disease triangle" of host-pathogen-environment interactions. Interpreting complex biological data now requires an integrated approach which includes a full array of analysis from the genomic to the metabolomic level. Computer and mathematical modeling provide the necessary tools for researchers to decipher these data. Collaborations both within and outside of Virginia Tech foster further bioinformatic and economic development. The many projects in which our research teams collaborate epitomize the multidisciplinary approach to which the Institute is dedicated.

Agriculture

Humans have manipulated genetic traits of plants and animals for millennia by painstakingly breeding for desired traits. Such methods, however, are difficult to evaluate and reproduce consistently. Biotechnology offers the opportunity to produce agricultural products through methods that are more precise and quicker than traditional techniques. As human populations continue to rise and arable land becomes increasingly scarce, genetically engineered crops may also be the only viable solution to feed the growing population and still maintain ecosystem biodiversity.

Unlocking the keys to plant metabolic processes will give us the tools to harvest the vast potential all plants carry in their genomes. We will be able, for example, to create nitrogen-fixing cereals as well as disease- and stress-resistant plants that will reduce our dependency on pesticides, herbicides, and chemical fertilizers. VBI researchers have undertaken a number of projects which will provide baseline information to catalyze this agricultural revolution.

Dr. Allan Dickerman with project team members Longlong Yang, Jiaxin Li, and Yuying Tian.



Plant Duplication Events

Genetically engineered varieties are already a major component of US-grown corn, cotton, and soybean crops, and, in the future, we can expect even more and greater innovations. To make this potential a reality, geneticists and molecular biologists must identify genes that can be manipulated to achieve desired traits. This can be a daunting task, due to the complex genetics of many plant varieties.

Plant genomes are often complicated by their polyploid ancestry, which further hampers attempts at genetic research. In polyploidy, each gene occurs in multiple copies, though many of these duplicates are lost or become specialized to perform other functions. Thus, it becomes difficult to accurately assess which genes are responsible for the desired trait.

In the Plant Duplication Events project, Drs. Sobral and Dickerman are developing an informatics system that will sort out these multiplicities. To perform these analyses, the researchers are building a database of published phylogenetic trees of agricultural plants. These phylogenetic trees detail the lineage of plants in relationship to one another, and demarcate the approximate time when new species branched from their predecessors. The phylogenetic tree database will help Drs. Sobral and Dickerman address functional outcomes of major genome duplication events. The database will also further knowledge of how and when these events occurred and allow us to pinpoint targeted sequence data for use in further agricultural development. Cold Spring Harbor Laboratory will be one of the first to make use of the results generated at VBI.

Rhizobia-Legume Symbioses

Legumes form nitrogen-fixing nodules in association with symbiotic bacteria called rhizobia. Natural nitrogen fixation is important to agricultural systems because of the high cost of fixing nitrogen industrially. In addition, symbiotic nitrogen fixation is of great scientific interest from evolutionary and developmental standpoints. During development of nitrogen-fixing nodules, rhizobia redirect plant developmental machinery to access intracellular spaces for nodule establishment. Early events in rhizobia-legume symbioses seem to use some of the same plant genes as those required for establishment of associations between plants and mycorrhiza, the beneficial fungi that aid in plant root development. This is of interest because arbuscular mycorrhiza-plant symbioses are much older and more widespread than legume-rhizobia symbioses. Recently discovered plant receptor kinase-like proteins seem to be involved in the early steps of plant-microbe signaling. Similar proteins have been implicated in plant disease resistance and animal immune system responses, as well as in plant and animal developmental programs. Of additional interest is the recent discovery, through comparative analysis of sequenced genomes of *Sinorhizobium meliloti* and *Brucella* spp. (intracellular pathogens of humans, cows, and pigs), of high levels of similarity between *Brucella* and rhizobial genomes. This suggests that certain features required for achieving intracellular life may be conserved across bacteria species.

VBI researchers are working to understand why only legumes form nitrogen-fixing symbioses with plants and what might be required to allow other plants to engage in such symbioses. In addition, comparing the genes, proteins, and metabolites of productive and mutant legume-rhizobial associations should allow researchers to understand the similarities and differences between the networks, genes, proteins, and metabolites that are affected throughout development. Further comparisons to other systems, such as *Brucella* and its hosts, will increase our understanding of intracellular bacterial pathogens, with potential applications in agriculture and human health.

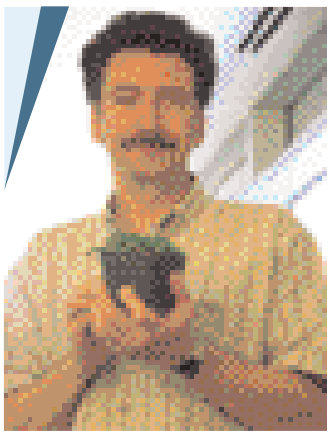
The Sobral Laboratory focuses on rhizobia-legume symbioses. Dr. Bruno Sobral, Nicole McMaster, and Dr. Sonal Malhotra examine bio-engineered plants.

SeedGenes

Arabidopsis thaliana, also known as mouse-eared cress, is the perfect model plant for genetic research because of its small, tidy genome. Identifying the functions of all genes in this species is the goal of NSF's *Arabidopsis* 2010 Project, of which the SeedGenes Project is one component. The SeedGenes Project is identifying and analyzing every *Arabidopsis* gene essential for seed development. The long-term goal is to use knowledge developed in the experimentally tractable *Arabidopsis* genome to understand how to improve and maintain those plants which are important to the sustenance of human civilization.

SeedGenes will focus on genes that give a viable seed phenotype when disrupted by mutation. The network of genes chosen for analysis is therefore not defined by shared biochemical or molecular features but rather by the fact that they are essential. *A. thaliana* contains an estimated 500 such *EMB* (embryo mutant) genes required for normal seed development and another 200 genes required for normal seed pigmentation. A coordinated bioinformatics effort will collect, analyze, and present information on the function of these genes based on molecular, genetic, and phenotypic studies from various public and private sources and synthesize the results for efficient use by the community. In addition, VBI will use whole-genome mRNA expression assays with Affymetrix technology to study the timing of when these genes function in early developmental stages. The resulting insights into biological functions of these diverse genes will complement biochemical research in other labs on specific gene products. The result will be an integrated view of essential gene functions within cellular, organismal, and evolutionary contexts.

The SeedGenes project builds upon an existing collaboration between Oklahoma State University and Syngenta that has identified over 350 tagged *emb* mutants and preliminary sequence data for these mutants in the past four years. This represents a significant corporate investment that includes substantial release of materials with minimal restrictions and provides a mechanism for collecting and analyzing community-wide information on essential genes. Such efforts are needed at the start of the NSF *Arabidopsis* 2010 project to focus attention on genes with important biological functions.



Dr. Vladimir Shulaev is one of many VBI researchers studying the model plant Arabidopsis thaliana.

Phytophthora

Phylogenetically distinct from plants, animals, and fungi, stramenopiles form a major branch of the evolutionary tree. They are perhaps best represented by the golden-brown algae called kelp, diatoms with their siliceous skeletons, and the pathogenic, fungi-like oomycetes. Stramenopiles are vital to proper carbon cycling and form the foundation for food webs in many ecosystems. They also can have deadly consequences for forests, crops, and humans. *Phytophthora* species are extremely destructive to forest ecosystems. *P. cinnamomi* completely annihilated American chestnut trees in the southern US 50 years before the chestnut blight destroyed the northern US population, and a newly emerged *Phytophthora* species is attacking the oak forests of California. *P. sojae* (a soybean pathogen) and *P. infestans* (the pathogen responsible for the infamous Irish potato famine) both pose major difficulties for the agricultural industry. In humans, the intestinal oomycete parasite, *Blastocystis hominis*, also affects many immuno-suppressed persons who suffer from AIDS, cancer, or similar diseases. As global warming continues, such organisms will gain an even greater foothold. And since September 11, the use of *Phytophthora* as a bioweapon is a potential threat.

Researchers at VBI hope to shortcut the spread of *Phytophthora* through in-depth genetic analysis. Despite their importance, stramenopile genomes have not yet been sequenced. Genetic and genomic studies are most advanced for the oomycetes *P. sojae* and *P. infestans*. Genetic maps, DNA transformation systems (including gene silencing), BAC libraries, and extensive EST collections exist for both species. The *Phytophthora* Genome Initiative (PGI), a collaborative effort of researchers across the globe, seeks to obtain a complete genome sequence of a *Phytophthora* species. VBI will provide a bioinformatics platform for access to and analysis of the sequence by the global research community, which will annotate and finish the sequencing. VBI will also conduct functional genomics research to discover how *Phytophthora* genes operate during infection.

Informatics

One of VBI's primary research goals is to develop informatics tools that will allow biologists and other life scientists to more thoroughly understand the volumes of data produced by their experiments. In previous genetic research, for instance, mutagenesis was the only feasible way of producing desired results. Scientists are discovering more accurate methods for finding the genes that control certain traits. One of the ways to do this is to model gene systems, using data generated by laboratory experiments and analyzing them with high-performance supercomputers.

Most natural systems, when reduced to their constituent components, are too complex for the human mind to comprehend. A cell, for instance, is comprised of many factors—RNA, DNA, metabolites, and proteins—just to name a few. All of these molecules interact with one another in many biochemical networks, which are usually very robust. VBI researchers are studying the sub-networks that form these larger networks. The goal is to create artificial networks with sets of properties that can be compared with real networks. Reference points can then be found for the behaviors scientists observe, which will help identify and quantify the networks. Such studies provide data that will give scientists a very precise understanding of the systems in question, thus allowing them to engineer organisms for specific traits.

Modeling also enables greater experimental efficiency by simulating experiments and allowing scientists to choose the experiment best suited for their needs. Models further allow the scientist to generate hypotheses which aid in experimental design, giving scientists better predictive power. Any cell —human, animal, or plant—can be modeled, which may, in the future, alleviate our current need to perform research on living organisms.

Currently, VBI researchers use two forms of modeling to accomplish these goals. The Biochemical Networks Modeling Group, led by Dr. Pedro Mendes, uses differential equations and probabilistic models via the software simulators Gepasi and COPASI (Complex Pathway Simulator). Dr. Reinhard Laubenbacher's Mathematics Group uses combinatorial algebra and topological spaces as another form of modeling.

Gepasi/COPASI

Dr. Pedro Mendes developed the software Gepasi to model biochemical networks. Its successor, COPASI, is an evolved version of Gepasi which allows researchers to take advantage of high performance computers when analyzing data. VBI and European Media Laboratory (EML) of Heidelberg, Germany, are collaborating to develop COPASI. VBI is providing the legacy code from Gepasi to build COPASI's supercomputing capabilities, while EML supplies the probabilistic models. COPASI incorporates a model generator, new simulation techniques, optimization routines, and methods from nonlinear dynamics. Both Gepasi and COPASI have the added advantage of user-friendly visualization platforms, enabling biologists to focus on the biological side of their research rather than worry with developing differential equations and probabilistic models on their own.

Discrete Methods in Mathematical Biology: Modeling of Biological Networks

The Gepasi/COPASI modeling method provides detailed models of biochemical networks, representing the concentrations of a variety of relevant chemical compounds. A complementary view of a network would describe a coarse high-level logical structure, which can also be used for both prediction and simulation. Dr. Laubenbacher's group is working on a novel approach to such models, using methods from discrete dynamical systems theory and symbolic computation. In particular, these methods allow the exploration of the whole parameter space of such models that fit a given time series of experimental data. Of special interest are metabolic and gene regulatory networks.

The Biochemical Networks Modeling Group (from left to right): Alberto de la Fuente, Aejaaz Kamal, Xingjing Li, Dr. Stefan Hoops, Dr. Pedro Mendes, Dr. Paul Brazhnik, Dianjing Guo, Mrinmayee Kulkarni, Christine Lee, and Dr. Dingjun Chen.

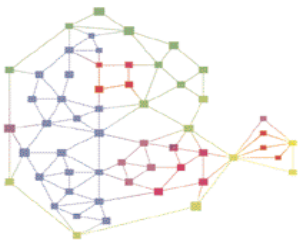
Theory and Practice of Computer Simulation

Large-scale computer simulations play an important role in mathematical biology and bioinformatics. Many practical issues related to simulations are difficult to answer without a theoretical underpinning. In collaboration with Los Alamos National Laboratory, Dr. Laubenbacher's group is working on a mathematical foundation for computer simulation. Of particular interest are applications of the theory to the simulation of biological networks.

Analysis of Interaction Pattern in Networks

This project uses tools from combinatorial topology and computational algebra to detect and measure dynamic and static interaction patterns in different types of networks.

As part of this project, Dr. Laubenbacher's group is developing graph-theoretic visualization methods, together with the Laboratoire Bordelais de Recherche Informatique (LaBRI) in Bordeaux, France, as well as new algebraic methods to represent and analyze geometric objects with Dr. B. Sturmfels from the Department of Mathematics at the University of California, Berkeley.



A planar graph created by Tulip, a program developed by LaBRI.

The Mathematics Group at VBI is comprised of (from left to right) Jignesh Shah, Brandilyn Stigler, Dr. Reinhard Laubenbacher, and Omar Colon-Reyes.

The *Medicago* Project

The *Medicago* Project, funded by the National Science Foundation, is a prime example of using simulation software to compare models with results from actual experiments. *Medicago truncula* is a close relative of *M. sativa*, the world's most important forage legume, also known as alfalfa. It is a rich source of natural products such as flavonoids and isoflavonoids, both of which have high nutritional value. *M. truncula*'s genome is smaller and thus easier to sequence and utilize in experiments. VBI is collaborating with the Samuel R. Noble Foundation on this project to understand how *M. truncula* resists disease. A functional genomics data set will be produced which will encompass expressed sequence information and the associated mRNA, protein, and metabolite identities and concentrations. Such information will also have implications for alfalfa and future plant research.

It is imperative that this project establish integrative models and software to facilitate analysis of the data relative to itself and to previous knowledge on sequences and pathways. A relational database is under construction that will store all data. An expandable analysis server will process the data with several statistical and numerical algorithms. A Web interface will integrate the previous components, making them accessible and user-friendly to researchers. By the end of the project, COPASI will be used to construct a predictive model which will interpret the regulation of the complex biological processes underlying disease resistance in *M. truncula*. Scientists will also be able to pinpoint key genes of possible agronomic importance, allowing them to reduce dependency on chemical fertilizers and synthesize natural products artificially, in addition to enhancing plant disease resistance.

Huafang Lai, Dr. Biswarup Mukhopadhyay, Kristen Boswell, and Jessica Kraszewski lead microbial research at VBI.

Biomedicine

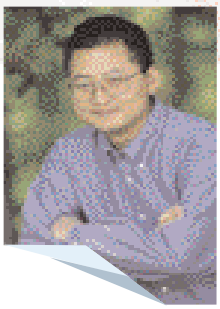
With a draft sequence of the human genome complete, scientists are now in the perfect position to reap the potential rewards from understanding our genetic blueprint. But with an estimated 30,000 genes comprised of 3 billion base pairs, how are we ever to analyze the mountain of data? Bioinformatics will answer that challenge by providing the computational tools and genomic expertise to interpret the data.

In 1989, William Wulf of the National Science Foundation (NSF) defined a collaboratory as "a center without walls, in which the nation's researchers can perform their research without regard to geographical location, interacting with colleagues, accessing instrumentation, sharing data and computational resources, [and] accessing information in digital libraries." VBI has laid the groundwork for several unique collaboratories which will help researchers understand infectious diseases, answering questions such as how *Mycobacterium tuberculosis*, responsible for tuberculosis in humans, resuscitates in the lungs after years of dormancy.

Kiss of the Mosquito

Annually, there are 300-500 million new cases of malaria and 1.5-2.7 million deaths. Malaria, in fact, has caused soldiers to lose more days from duty than bullets in every conflict in this century where malaria is endemic. The organism that causes malaria, *Plasmodium falciparum*, has proven very resistant to most forms of treatment, including chloroquine, which was once hailed as the solution to the malaria problem. Researchers at Johns Hopkins University's Bloomberg School of Public Health have found that some of this resistance may stem from the fact that *P. falciparum* contains a plasmid-like structure in its cell makeup, something that is normally only found in plants.

Bioinformatics, however, may give us the keys to solving the malaria puzzle. A deeper understanding of the human, *Anopheles* mosquito, and *P. falciparum* genomes, and how these genomes interact, may help us develop a DNA-based vaccine that will give immunity to those who would otherwise die of the disease. The problem lies in finding the proper protein targets for development of a vaccine. With the epidemiological expertise of our collaborators at JHU and VBI's microarray technology and vast computational facilities, we may be able to isolate the gene or genes that could stamp out malaria and make a mosquito bite little more than a nuisance, rather than the kiss of death.



Dr. Yongqun (Oliver) He, Project Manager for the PathPort Project and member of the Brucella *Microarray Research Group.*

Combating Infectious Diseases

Worldwide, millions of people succumb to infectious and parasitic diseases every year. The simple fact of so much death from disease suggests the need for new approaches to the problem. Bioinformatics is a critical component missing from infectious disease research and treatment. Through understanding how the genes of hosts and pathogens interact across a broad array of species and environments, VBI will provide information to aid in rapid detection, identification, and remediation.

VBI has begun a five-year collaboration with Johns Hopkins University's Bloomberg School of Public Health to study some of the most deadly infectious diseases—tuberculosis (TB), acquired immune deficiency syndrome (AIDS), malaria, toxoplasmosis, and measles. While Johns Hopkins researchers will provide vital data, VBI's bioinformatics capabilities will allow broad comparisons between various host responses to different pathogens. Microarray technologies will determine the expression patterns of genes and correlate these with susceptibility and resistance to the disease.

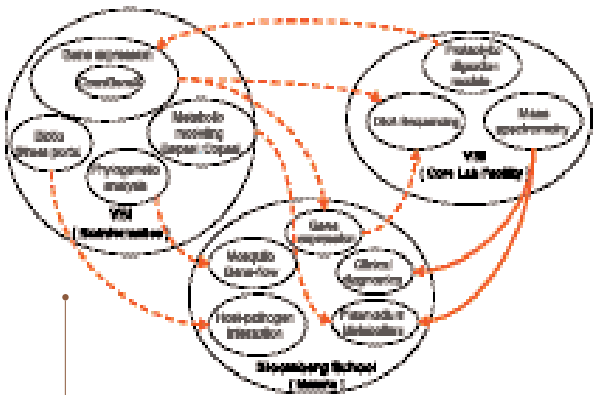
Dr. Stefan Hoops, Dr. Dana Eckart and Dave Sebring have been instrumental in bringing the PathPort project to fruition. At right, the multidisciplinary PathPort group is comprised of experts from the fields of computational biology, veterinary medicine, and computer science, to name a few.

PathPort: Bridging the Gap of Knowledge

Coupled with the need for greater understanding of the world's leading infectious diseases is the need for greater general knowledge of pathogens. Currently, there is no central database which stores genomic information on pathogens. Accessibility for government officials, health care workers, and researchers is paramount to fighting and managing both natural and engineered disease outbreaks.

One of the greatest weapons we have against such outbreaks is knowledge. A science portal providing access to molecular information about known pathogens and their near relatives would aid us in rapid assessment and abatement. To provide this platform, VBI is performing genome data acquisition, consolidation, and annotation; creating data models, graphical user interfaces, and bioinformatics tools for analysis and science portals; developing methods to validate candidate target sequences; and researching host response to pathogen models. Funded by the Department of Defense (DoD), this science portal will be regularly updated, representing a consensus list of high-priority pathogens, variant strains, and closely-related species. PathPort will import data from public databases and Web-accessible collections of pathogen sequence data. VBI will achieve this with its high-speed computational infrastructure, its world-class laboratory facilities, and its skilled researchers—a powerful combination in the war against disease.

VBI-JHU Cross-Institutional Interactions: An Example with Malaria



Education and Outreach Programs

Virginia Bioinformatics Institute is as much about preparing a world-class workforce as it is about research and discovery. Our unique organizational structure opens up new vistas and frontiers for learning and discovery. VBI currently has ten graduate students engaged in research programs via interdepartmental options in bioinformatics or the Interdepartmental Genetics Ph.D. Program. This program will soon formally include a Ph.D. program in Bioinformatics.

VBI's undergraduate and high school mentoring programs also allow a younger generation hands-on research experience in bioinformatics. For example, during the summer semester, Nick Galloway and Felix Kim, local high school students, facilitated networking in the CCF and learned to create custom gene chips in the CLF, respectively. Through this educational pipeline, VBI trains a workforce ready to meet the challenges and promises of the post-genomic age.

In addition to its educational endeavors, the Public Relations and Outreach Team continues to develop team building and outreach programs within the Institute, throughout Virginia Tech, and beyond. VBI will soon host several short courses and conferences centered on bioinformatics research, each coordinated through education and outreach efforts. Another major effort melds bioinformatics education and outreach efforts at VBI with local research programs to be established at the Institute for Advanced Learning and Research in Danville, Virginia.

Through such outreach programs, facility tours, mass media and other communication channels, we provide our colleagues and collaborators with timely information regarding research and education endeavors at VBI. We hope this information flow continues to nurture collaborations within the broader research and education communities, as well as portraying the significant contributions bioinformatics research will make in each of our daily lives.

Nick Galloway participates in High School/High Tech, a nationwide mentoring program. Nick will enter Virginia Tech as a freshman in Spring 2003.

The Public Relations and Outreach Team at VBI oversees media relations, popular publications, and VBI events, and also administers learning experiences for students from high school to graduate studies. Members include (from left to right): Tiffany Trent, Dr. Neysa Call, and Candace Baracat.

Development

Because of the high cost of advanced bioinformatics technology, VBI takes a proactive stance on development issues. The Institute has garnered over $24 million in grants and contracts, and continues to seek funding in several key areas. Our world-class computing and laboratory facilities utilize a range of equipment from terascale data storage to high-end mass spectrometers. We have been fortunate thus far to draw the best faculty, research scientists, and graduate students. Our development officer, Jeff Janosko, works in tandem with University Development and VBI to ensure our continued success.

VBI's research is solidly founded upon our world-class faculty. Experts in computer science, biology, biochemistry, mathematics, microbiology, plant pathology, and more have all joined within the Institute to create a unique center for bioinformatics research.

Allan Dickerman
Research Assistant Professor

Ph.D., Zoology
University of Wisconsin-Madison, 1992

Research

Dr. Dickerman's research addresses the broad area of genomic sequence analysis from an evolutionary perspective. He focuses on using phylogenetic analyses to describe interesting patterns of conservation or diversity in gene sequences, where "interesting" is defined by the context of the research questions. Current collaborative efforts are aimed at using comparative analysis of plant gene sequence to extrapolate functional information from the model dicot *Arabidopsis thaliana* to other species. The input of interdisciplinary research is essential to solve the large, complex problems posed by biological systems.

In addition to the biological questions addressed through comparative genomic analysis, Dr. Dickerman develops algorithms used in this arena. A common theme in complex genome analysis is the joint inference of multiple phylogenetic models, such as genes evolving within species. Adapting these methods to genome-scale analyses is one of his activities.

Research Group

Dr. Dickerman and his group are developing functional interpretations of DNA sequence data using comparative methods. Project team members are Jiaxin Li, Longlong Yang, and Yuying Tian.

Grants

PI. Essential Gene Functions in *Arabidopsis* Seed Development. National Science Foundation. 10/1/01-9/30/05: $2,326,667.

PI. Bioinformatics Prediction of Functions of Unculturable Microbes in Ecosystems. National Science Foundation. 1/1/02-12/31/02: $100,000.

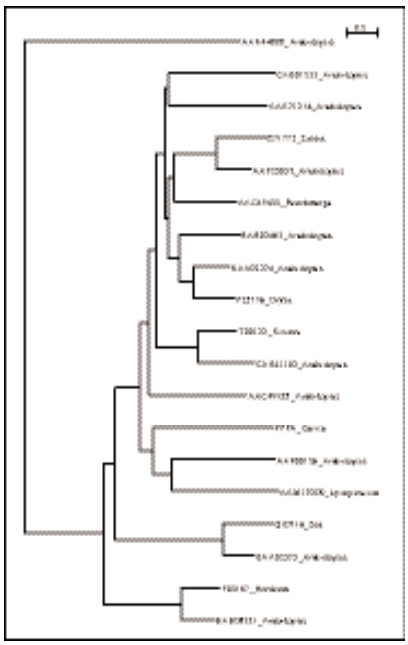
co-PI. Analysis of Plant Genome Duplication Events and Their Functional Relevance. U.S. Department of Agriculture - Cooperative State, Research, Education, and Extension Service. 4/1/01-3/31/02: $444,230.

Selected Publications

Dickerman, A. W. 1998. Generalizing phylogenetic parsimony from the tree to the forest. *Systematic Biology* 47:414-426.

Meyers, B.C., A.W. Dickerman, R.W. Michelmore, R.M. Pecherer, S. Sivaramakrishnan, B.W.S. Sobral, N.D. Young. 1999. Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *The Plant Journal* 20:317-332.

Huala, Eva, Allan W. Dickerman, Margarita Garcia-Hernandez, Danforth Weems, Leonore Reiser, Frank LaFond, David Hanley, Donald Kiphart, Mingzhe Zhuang,Wen Huang, Lukas A. Mueller, Debika Bhattacharyya, Devaki Bhaya, Bruno W. Sobral, William Beavis, David W. Meinke, Christopher D. Town, Chris Somerville, and Seung Yon Rhee. 2001. The *Arabidopsis* Information Resource (TAIR): A comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. *Nucl. Acids Res.* 29: 102-105.



Gene phylogeny of cystein proteases in Arabidopsis *and other plants.*

Reinhard Laubenbacher
Research Professor
Professor of Mathematics

Ph.D., Mathematics
Northwestern University,
1985

Research

Dr. Laubenbacher directs the Mathematics Group at VBI. This group is interested in the development and application of bioinformatics tools using discrete mathematics, dynamic systems theory, and symbolic computation. Methods from combinatorics and combinatorial topology, as well as computational polynomial algebra, are of particular interest. Active projects include: development of mathematical models for gene regulatory networks, development of theory and applications of computer simulation to the study of networks, topological analysis of interaction patterns in networks, and development of algorithms for invariants of cell complexes and other objects from combinatorics and combinatorial geometry.

Mathematical models are often used to analyze large systems, but current computational methods sometimes cannot handle the amount of data generated from genetic research. To get a clearer picture of how systems such as gene networks function, Dr. Laubenbacher uses cell complexes to represent certain aspects of biological systems. These complexes create algorithms that will allow for more computational flexibility and efficiency. Applying theories previously used to understand patterns in large-scale traffic models and the spread of pathogens, Dr. Laubenbacher and his group are modeling gene regulatory networks of yeasts as a testbed for their approach.

Research Group

Dr. Laubenbacher's research group consists of Brandilyn Stigler, Jignesh Shah, and Omar Colon-Reyes. Dr. Laubenbacher mentors six graduate students in conjunction with the Mathematics Department at Virginia Tech.

Grants

PI. Algebraic Algorithms for Cell Complexes. National Science Foundation. 5/1/02-4/30/03: $122,475.

PI. Recruitment and Educational Development of Minority Mathematics Students. Virginia Tech Diversity Initiative. 7/1/02-11/2/02: $1,000.

Selected Publications

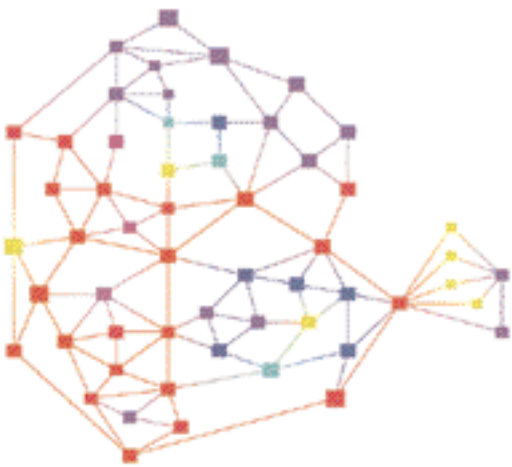
Laubenbacher, R., H. Barcelo, X. Kramer, and C. Weaver. 2001. Foundations of a connectivity theory for simplicial complexes. *Adv. Appl. Math.* 26: 97-128.

Laubenbacher, R., and B. Parageis. 2001. Equivalence relations on finite dynamical systems. *Adv. Appl. Math.* 26: 237-251.

Laubenbacher, R., E. Aguirre, A. Jarrah, J.A. Ortiz-Navarro, and R. Torrez. 2001. Generic ideals and the Moreno-Socias Conjecture. *Proceedings of the 2001 International Symposium on Symbolic and Algebraic Computation.* ACM.

Laubenbacher, R., G. McGrath, and D. Pengelley. 2001. Lagrange and the solution of numerical equations. *Historia Mathematica* 28: 220-231.

Laubenbacher, R., and C. Woodburn. 2000. A new algorithm for the Quillen-Suslin Theorem for polynomial rings. *Contributions to Algebra and Geometry* 41: 23-32.



A planar graph created by Tulip, a program developed by LaBRI.

Pedro Mendes
Assistant Research Professor
Adjunct Assistant Professor of Biochemistry

Ph.D., Biochemistry
University of Wales, Aberystwyth (UK), 1994

Research

Dr. Mendes' research centers on computer simulation and analysis of biochemical networks. This consists of three main components: development of simulation software (Gepasi and now COPASI), modeling gene expression in the context of metabolic networks, and bioinformatics support for functional genomics.

Dr. Mendes is the author of the popular biochemical simulation software Gepasi, in use in many laboratories worldwide. This software facilitates mathematical modeling of biochemical networks without the need to write the mathematics explicitly. Gepasi is a problem-solving environment that allows biochemists to carry out computer simulation without extra programming. Dr. Mendes is also collaborating with European Media Laboratory (EML) in Heidelberg, Germany, to create a new simulator called COPASI. This program will succeed Gepasi with improved functionality for high-performance computers and a user-friendly interface.

Research in Dr. Mendes' group is active in the area of modeling gene networks together with biochemical pathways. Since gene and biochemical networks are tightly interconnected, it only makes sense to model the two together. An ambitious target of this research has been to uncover these networks from experimental observations of gene expression and proteomics. Progress in this area has proceeded at a steady pace, as reported recently in the journal *Trends in Genetics*. Work is now expanding to microarray and metabolomics data analysis. Efforts in this area benefit from collaborations abroad with Drs. Hans Westerhoff (Free University of Amsterdam) and Jacky Snoep (University of Stellenbosch, South Africa), and in the United States with Drs. Reinhard Laubenbacher (VBI and Mathematics Department), Craig Nessler (Plant Pathology and Physiology), and Ina Hoeschele (Dairy Science).

In order to integrate genetics with metabolism, whole organism-level measurements of thousands of molecules must be made using novel high-throughput technologies, such as microarrays and mass spectrometry. In particular, it is important to measure the levels of small organic molecules (natural products or metabolites). Such metabolite profiling generates large amounts of data and therefore requires extensive bioinformatic support.

Dr. Mendes has many active research projects of this nature together with collaborators in the Samuel Roberts Noble Foundation, namely Drs. Rick Dixon, Lloyd Sumner, and Greg May, and Dr. Vladimir Schulaev at VBI.

Research Group

Dr. Mendes leads a research group of 15 researchers, including 2 Senior Research Associates, 3 Postdoctoral Associates, 4 Research Associates, 2 Ph.D. students and 3 M.S. students. The group is very interdisciplinary, with backgrounds from biochemistry, bioinformatics, computer science, biology, genetics, physics, and mathematics.

Grants

PI. An Integrated Approach to Functional Genomics and Bioinformatics in a Model Legume. National Science Foundation. 8/1/01-7/31/05: $3,587,432.

PI. Reverse Engineering of Biochemical Networks from Whole-Genome Dynamics. National Science Foundation. 10/1/01-9/30/02: $99,999.

PI. Database and Visualization Systems for Metabolomics. Phenomenome Discoveries, Inc. 7/15/01-4/15/02: $228,429.

co-PI. Sun Center of Excellence in Bioinformatics. Sun Microsystems, Inc. 5/1/01-4/30/04: $1,262,486.

co-PI. Metabolic Engineering of Plant Vitamin C Biosynthesis for Improved Nutrition and Health. National Science Foundation. 9/15/01-8/31/04: $99,999.

Selected Publications

Mendes, P. 2002. Emerging bioinformatics for the metabolome. *Briefings in Bioinformatics* 3(2): 134-145.

Mendes, P., A. de la Fuente, and S. Hoops. 2002. Bioinformatics and computational biology for plant functional genomics. *Recent Advances in Phytochemistry* 36: 1-13.

de la Fuente, A., P. Brazhnik, and P. Mendes. 2002. Linking the genes: Inferring gene networks from microarray data. *Trends in Genetics* 18(8): 395-398.

Mendes, P. 2001. Modeling large scale biological systems from functional genomic data: Parameter estimation. In: *Foundations of Systems Biology*, ed. H. Kitano, 163-186. Cambridge: MIT Press.

Mendes, P., and D.B. Kell. 2001. MEG (Model Extender for Gepasi): A program for the modeling of complex, heterogeneous cellular systems. *Bioinformatics* 17: 288-289.

Mendes, P., and D.B. Kell. 1998. Non-linear optimization of biochemical pathways: Applications to metabolic engineering and parameter estimation. *Bioinformatics* 14: 869-883.

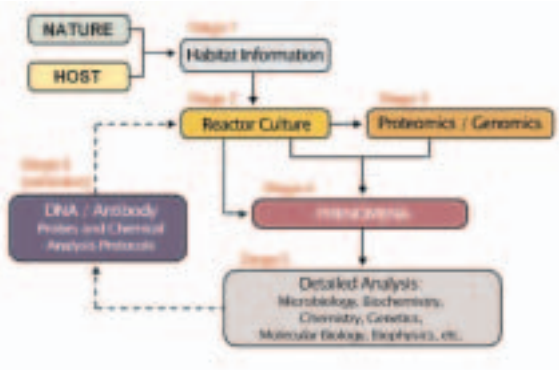
Biswarup Mukhopadhyay
Research Assistant Professor
Adjunct Assistant Professor of Biochemistry and Biology

Ph.D., Microbiology
University of Iowa,1993

Postdoctoral Training in the Laboratory of Prof. Ralph S. Wolfe, University of Illinois at Urbana-Champaign

Research

Dr. Mukhopadhyay's research falls into two fields: experimental functional genomics and mechanistic biochemistry. His work also emphasizes understanding of evolutionary processes. The figure below summarizes the steps comprising the experimental functional genomics approach. This approach is used to discover environmental stimuli-directed novel behaviors of target microorganisms and to elucidate their underlying molecular mechanisms. This complete approach makes use of advances in genomics and chemistry, and helps to prioritize and provide *in vivo* relevance to *in vitro* biological research. It has aided discovery of certain microbial behaviors that were previously logically unpredictable.



Model organisms for experimental functional genomics projects are *Methanococcus jannaschii*, an extremophile that lives in submarine hydrothermal vents, and *Mycobacteria*, major human and animal pathogens as well as inhabitants of the soil.

The mechanistic biochemistry projects deal with the following systems and involve structure-function analysis and physiological studies: oxaloacetate biosynthesis in methanogenic archaea; anaplerotic and gluconeogenesis reactions in *Mycobacteria* that also serve as models for studying certain aspects of non-insulin-dependent diabetes in vertebrates; evolution of biotin-dependent enzymes; hydrogen sensing, detoxification and volatilization of selenium, and protein modification in *M. jannaschii*.

The work covers enzyme purification and characterization (structural and kinetic properties), site-directed mutagenesis and design of chimeric proteins (including the development of genetic screens for the mutant or designed proteins), and the design of inhibitors (drugs), etc. The overall goal is to study the ecophysiology of a target microorganism and the biochemical basis of microbial diversity.

Research Group

Biswarup's research group consists of Haufang "Lilly" Lai, Senior Lab Specialist; and Jessica Kraszewski and Kristin Boswell, both Undergraduate Research Students.

Selected Publications

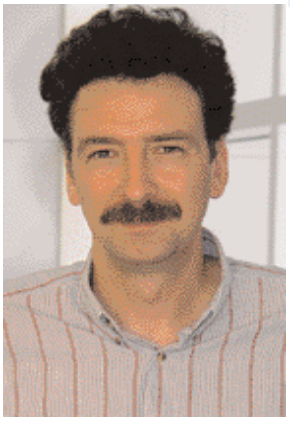
Galagan, J. E., C. Nusbaum, A. Roy, M. G. Endrizzi, P. Macdonald, W. FitzHugh, S. Calvo, R. Engels, S. Smirnov, D. Atnoor, A. Brown, N. Allen, J. Naylor, N. Stange-Thomann, K. DeArellano, R. Johnson, L. Linton, P. McEwan, K. McKernan, J. Talamas, A. Tirrell, W. Ye, A. Zimmer, R. D. Barber, I. Cann, D. E. Graham, D. A. Grahame, A. M. Guss, R. Hedderich, C. Ingram-Smith, H. C. Kuettner, J. A. Krzycki, J. A. Leigh, W. Li, J. Liu, B. Mukhopadhyay, J. N. Reeve, K. Smith, T. A. Springer, L. A. Umayam, O. White, R. H. White, E. Conway de Macario, J. G. Ferry, K. F. Jarrell, H. Jing, A.J. Macario, I. Paulsen, M. Pritchett, K. R. Sowers, R. V. Swanson, S. H. Zinder, E. Lander, W. W. Metcalf, and B. Birren. 2002. The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res.* 12: 532-542. http://www.genome.org/cgi/content/full/12/4/532.

Mukhopadhyay, B., E. M. Concar, and R. S. Wolfe. 2001. A GTP-dependent vertebrate-type phosphoenolpyruvate carboxykinase from *Mycobacterium smegmatis*. *J. Biol. Chem.* 276: 16137-16145. http://www.jbc.org/cgi/content/full/276/19/16137.

Mukhopadhyay, B., E. Purwantini, C. L. Kreder, and R. S. Wolfe. 2001. Oxaloacetate synthesis in the methanarchaeon *Methanosarcina barkeri*: Pyruvate carboxylase genes and a putative *Escherichia coli*-type bifunctional botin protein ligase gene (bpl/birA) exhibit a unique gene organization. *J. Bacteriol.* 183: 3804-3810. http://jb.asm.org/cgi/content/full/183/12/3804.

Mukhopadhyay, B., E. F. Johnson, and R. S. Wolfe. 2000. A novel pH2 control on the expression of flagella in the hyperthermophilic strictly hydrogenotrophic methanarchaeon *Methanococcus jannaschii*. *Proc. Natl. Acad. Sci. USA* 97: 11522-11527. http://www.pnas.org/cgi/content/full/97/21/11522. http://www.elsevier.nl/gejng/10/11/32/53/30/27/abstract.html.

Mukhopadhyay, B., V. J. Patel, and R. S. Wolfe. 2000. A stable archaeal pyruvate carboxylase from the hyperthermophile *Methanococcus jannaschii*. *Arch. Microbiol.* 174: 406-414. http://link.springerny.com/link/service/journals/00203/contents/00/00225/index.html.



Vladimir Shulaev
Research Associate
Professor

Ph.D., Biological Sciences
Ukraine Academy of
Sciences, 1987

Ph.D., Plant Biology
Rutgers University, 1995

Research

Dr. Shulaev's research focuses on studying the biochemical, cellular, and molecular mechanisms of plant responses to biotic and abiotic stress. Susceptibility of crops to stresses—diseases, drought, high salinity, temperature, and others—significantly reduces global agriculture productivity. Efforts to improve plant stress tolerance through traditional breeding or genetic engineering have had limited success, due in part to a poor understanding of the basic mechanisms underlying plant adaptive responses to stress. A more mechanistic understanding of the underlying plant responses to environmental stresses is essential in formulating future breeding and engineering strategies aimed at reducing crop losses.

Current research in Dr. Shulaev's lab aims at comprehensive metabolic profiling of plant-derived chemicals involved in stress response in order to identify novel protective metabolites and genes involved in their biosynthesis and regulation. Environmental stress response in plants is an extremely complex trait controlled by multiple genes and affected by numerous external factors. Plants respond to stress by dramatically altering both primary and secondary metabolism. These metabolic changes result in biosynthesis of many specific chemicals designed to protect cells and organs from extreme environmental conditions and pathogens. Numerous plant metabolites are synthesized only in response to a specific stress factor. Many stress-related plant compounds are also traditionally used in pharmaceutical, health, food, cosmetic, agrochemical, and other industries as nutritional supplements, antioxidants, pharmaceuticals, insecticides, and other agents.

In the area of abiotic stress tolerance, the major research focus is on understanding the basic mechanisms of plant adaptation to osmotic and oxidative stress using *Arabidopsis thaliana* as a model system. Another research area includes investigation of early events in fungal pathogenesis using *Arabidopsis* interaction with various biotrophic and necrotrophic fungal pathogens. Early events in pathogen recognition are very important in determining the specificity of responses to different pathogens.

For analysis of plant metabolites, two major analytical platforms are utilized — gas chromatography coupled with mass spectrometry (GC-MS) and high performance liquid chromatography linked to mass spectrometry (LC-MS). Implementation of this dual platform allows simultaneous analysis of a large number of individual compounds with different size, polarity, and other physicochemical properties. Use of *Arabidopsis* as a model system permits rapid identification of genes involved in the biosynthesis and regulation of various classes of compounds. Knowledge of biosynthetic genes and global pathway regulators will improve our ability to manipulate metabolic fluxes toward production of specific products or even generate novel compounds with desired properties.

The ultimate goal of Dr. Shulaev's research is the development of new stress-tolerant crops and improving methods of disease control. Biotechnological manipulation of stress metabolites can provide novel ways of engineering plants with desired agronomic traits or increased nutritional value. In addition, finding novel ways to regulate the levels and yield of specific phytochemicals, or identification of novel compounds for human use, will be of great benefit and significant economic importance.

Selected Publications

Mittler, R., and V. Shulaev. 2002. Cell death in plant, yeast, and bacterium. In: *Essentials of Apoptosis: A Guide for Basic and Clinical Research*. Totowa, NJ: Humana. In Press.

Mittler, R., E. Lam, V. Shulaev, and M. Cohen. 1999. Signals controlling the expression of cytosolic ascorbate peroxidase during pathogen-induced programmed cell death in tobacco. *Plant Mol. Biol.* 39: 1025-1035.

Ribnicky, D.M., V. Shulaev, and I. Raskin. 1998. Intermediates of salicylic acid biosynthesis in tobacco. *Plant Physiology* 118: 565-72.

Shulaev, V., P. Silverman, and I. Raskin. 1997. Airborne signaling by methyl salicylate in plant pathogen resistance. *Nature* 385: 718-721.

Mittler, R., V. Shulaev, M. Seskar, I. Raskin, and E. Lam. 1996. Inhibition of programmed cell death in tobacco plants during a pathogen-induced hypersensitive response at low oxygen pressure. *The Plant Cell* 8: 1991-2001.

Shulaev, V., J. Leon, and I. Raskin. 1995. Is salicylic acid a translocated signal of systemic acquired resistance in tobacco? *The Plant Cell* 7: 1691-1701.

Leon, J., V. Shulaev, N. Yalpani, M.A. Lawton, and I. Raskin. 1995. Benzoic acid 2-hydroxylase, a soluble monooxygenase from tobacco, catalyzes salicylic acid biosynthesis. *Proc. Natl. Acad. Sci. USA* 92: 10413-10417.

Mittler, R., V. Shulaev, and E. Lam. 1995. Coordinated activation of programmed cell death and defense mechanisms in transgenic tobacco plants expressing a bacterial proton pump. *The Plant Cell* 7: 29-42.

Bruno Sobral
Director, Virginia Bioinformatics Institute
Professor of Plant Pathology, Physiology, and Weed Science

Ph.D., Genetics
Iowa State University, 1989

Research

Dr. Sobral's group focuses on two broad areas of biology. One is computationally oriented, while the other centers around understanding and reverse engineering host-pathogen-environment interactions. The computationally oriented effort aims at building, deploying, and evolving information systems that support biological data collection, integration, analysis, and discovery, especially for diverse communities of specialists with disparate needs and interests. In this context, the research focuses on information technology (IT) and the engineering of information systems that handle biological data for diverse user communities. The group's IT research and development deals directly with high-performance computational infrastructure for the biological research community as well as others who consume and apply biological knowledge.

Host-pathogen-environment interactions, sometimes called "the disease triangle", provide interesting biological features for understanding evolutionary and developmental processes of a diverse array of organisms. When pathogens and hosts interact in specific environments, they display complex and interwoven communication strategies and, through this, perturb one another's genetic, regulatory, and biochemical networks. Dissecting and understanding such interactions requires an integration of reductionist molecular data and approaches as well as an integrated and comparative view of the interaction (especially important to discover and understand network behavior and constraints). The nitrogen-fixing legume-rhizobial symbioses, especially the *Medicago truncatula – Sinorhizobium meliloti* system (since the genomes are either complete or in progress), are important to the group's continued work in this area. Various hypotheses exist regarding the evolutionary history of this association, but it is clear that rhizobia have successfully interacted with legumes to control the developmental processes that enable the symbiosis, and evidence is accumulating that this is accomplished through manipulation of the genes that are very similar to plant disease resistance genes. In some ways, symbiosis is a special case of a host-pathogen-environment interaction; one in which the host and pathogen have decided that collaboration is preferable to escalation of the arms race, as long as both sides are mutually satisfied. In addition, Dr. Sobral's group also focuses on *Phytophthora*-plant systems and is particularly interested in the interactions of *Phytophthora* with model plants, such as *Arabidopsis thaliana* and *Medicago trunculata*.

Grants

PI. Shared University Research Grant. IBM Corporation. 2001: $893,220.

PI. Collaborative Development of an EST Database and Analysis Pipeline. University of Nevada-Reno and Samuel Roberts Noble Foundation. 12/1/00-11/30/02. $323,016.

PI. PathPort: A Common Asset for Biological Security. U.S. Department of Defense. 4/1/02-3/31/07: $4,000,000.

PI. Sun Center of Excellence in Bioinformatics. Sun Microsystems, Inc. 5/1/01-4/30/04: $1,262,486.

PI. Johns Hopkins University Collaboration. Johns Hopkins University Bloomberg School of Public Health. 2/1/02-1/30/07: $10,000,000.

PI. Analysis of Plant Genome Duplication Events and Their Functional Relevance. U.S. Department of Agriculture - Cooperative State, Research, Education, and Extension Service. 4/1/01-3/31/02: $444,230.

co-PI. Communication, Training, and Resources for *Phytophthora* Research. National Science Foundation. 4/1/02-3/30/07: $499,633.

co-PI. Genome Sequence of *Phytophthora sojae*. USDA CSREES. 10/1/02-9/30/04: $2,529,384.

co-PI. Collaboration with the College of William & Mary and INCOGEN, Inc. Commonwealth Technology Research Fund. 1/1/02-12/31/04: $3,251,901.

co-PI. Collaborative Research in Bioinformatics. Commonwealth Technology Research Fund. 7/1/01 – 6/30/02: $2,500,201.

co-PI. Virginia Bioinformatics Consortium. Commonwealth Technology Research Fund. 7/1/01-6/30/02: $1,500,000.

Selected Publications

Huala, Eva, Allan W. Dickerman, Margarita Garcia-Hernandez, Danforth Weems, Leonore Reiser, Frank LaFond, David Hanley, Donald Kiphart, Mingzhe Zhuang,Wen Huang, Lukas A. Mueller, Debika Bhattacharyya, Devaki Bhaya, Bruno W. Sobral, William Beavis, David W. Meinke, Christopher D. Town, Chris Somerville, and Seung Yon Rhee. 2001. The *Arabidopsis* Information Resource (TAIR): A comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. *Nucl. Acids Res.* 29: 102-105.

Sobral, B.W.S., H. Mangalam, A. Siepel, P. Mendes, R. Pecherer, and G. MacLaren. 2001. Bioinformatics for rice resources. In: *Rice Biotechnology: Improving Yield, Stress Tolerance, and Grain Quality*. Wiley: Chichester (Novartis Foundation Symposium 236). 59-84.

Sobral, B.W.S. 2001. The role of bioinformatics in germplasm conservation and use. In: *Managing Plant Diversity*. eds. J.M.M. Engels, V. Ramanatha Rao, A.H.D. Brown, M. Jackson. CABI Publishing.

Sobral, B.W.S., M.E. Waugh, and W.D. Beavis. 2001. Systems approaches to support discovery in agricultural genomics. Chapter 8 In: *DNA-based Markers in Plants*, 2nd ed. eds. I. Vasil, R.L. Phillips.

Brett Tyler
Research Professor
Professor of Plant Pathology, Physiology, and Weed Science

Ph.D., Molecular Biology
University of Melbourne, Australia, 1981

Research

Dr. Tyler's research pinpoints genes and molecules that mediate interactions between plants and microbes. These interactions can be beneficial, resulting in increased plant performance, or negative, resulting in plant disease. In either case, the end result of the interaction is not just the result of two organisms interacting, but is the result of a complex web of signals and responses exchanged among the vast diversity of microbes, microfauna, predators, and competing plants that comprise the environment of a plant.

The long-term goal of Dr. Tyler's research is to understand the operation of this signaling web sufficiently to enable the design of more sustainable agricultural systems. This will require broad scale identification of the organisms participating in the communities, the genes they possess, how they use those genes, and the signals they transmit among each other. Dr. Tyler is also interested in the application of similar approaches to the interactions of microbes with animals and humans.

Dr Tyler's current research centers on identifying and characterizing the signals exchanged between plant pathogens called *Phytophthora* and the plant species they attack, especially soybean. *Phytophthora* pathogens are fungus-like organisms called oomycetes that include the organisms responsible for the Irish potato famine in the 19th century. His research includes a focus on individual genes and signals involved in *Phytophthora*-plant interactions, and whole-genome approaches that include characterizing all the genes in a *Phytophthora* species and determining how they contribute to signaling and pathogenesis.

Brett Tyler and Xuemin Zhang examine soybean plants for disease resistance.

Research Group

Dr. Tyler's group, which includes Felipe Arredondo, Xuemin Zhang, Sucheta Tripathy, Lecong Zhou, Dianjing Guo, Trudy Torto, Brian Smith, and Dao Dou, has begun research into bioinformatic approaches that are needed to unravel the functioning of complex communities of micro-organisms and their interactions with macro-organisms such as plants and animals.

Grants

PI. Communication, Training, and Resources for *Phytophthora* Research. National Science Foundation. 4/1/02-3/30/07: $499,633.

PI. Dissecting Soybean Resistance to *Phytophthora* by QTL Analysis of Host and Pathogen Expression Profiles. 10/1/02-9/30/07: $6,764,465.

PI. Function of Avirulence Genes in *Phytophthora sojae* Infection of Soybean. USDA Biology of Microbe-Plant Interactions. 9/1/02-8/31/04: $230,000.

PI. Genome Sequence of *Phytophthora sojae*. USDA CSREES. 10/1/02-9/30/04: $2,529,384.

co-PI. Bioinformatics Prediction of Functions of Unculturable Microbes in Ecosystems. National Science Foundation. 1/1/02-12/31/02: $100,000.

Patents

Tyler, Brett M., Lloyd Yu, Christopher Mau, and Elizabeth Doyle. Nucleic Acid Encoding Plant Elicitin Receptor And Methods Of Use. Pending.

Selected Publications

Tyler, B.M. 2002. Molecular basis of recognition between *Phytophthora* species and their hosts. *Annual Reviews of Phytopathology* 40. In press.

Chamnanpunt, J., W-X Shan and B.M. Tyler. 2001. High frequency mitotic gene conversion in genetic hybrids of the oomycete *Phytophthora sojae*. *PNAS* 98(25): 14530-14535.

Tyler, Brett M. 2001. Genetics and genomics of the *Phytophthora*-host interface. *Trends in Genetics* 17(11): 611-614.

Morris, P.F., E. Bone and B.M. Tyler. 1998. Chemotropic and contact responses of *Phytophthora sojae* hyphae to soybean isoflavonoids and artificial substrates. *Plant Physiol.* 117(4): 1171-1178.

Claassen,V.P., R.J. Zasoski and B.M. Tyler. 1996. A method for direct soil extraction and PCR amplification of endomycorrhizal fungal DNA. *Mycorrhiza* 6(5): 447-450.

Tyler, B.M., Wu, M-H., Wang, J-M., Cheung, W.W.S. and Morris, P.F. 1996. Chemotactic preferences and strain variation in the response of *Phytophthora sojae* zoospores to host isoflavones. *Appl. Environ. Microbiol.* 62(8): 2811-2817.

Tyler, B.M., Förster, H. and Coffey, M.D. 1995. Inheritance of avirulence factors and RFLP markers in outcrosses of the oomycete *Phytophthora sojae*. *Mol. Plant-Microbe Inst.* 8(4): 515-523.

Lauren Coble
Associate Director of
Administration and Finance

Dave Sebring
Associate Director of
Government and Corporate
Relations

Administration

VBI's Administrative Staff ensures that the internal administration of the Institute runs smoothly. Their many responsibilities include coordinating high volume recruitment and hiring; maintaining a multi-faceted financial management process including procurement, accounting, and financial reporting; and managing a wide array of facilities and construction needs. This team also oversees the internal grant proposal process, from pre-proposal to funded project. Their enthusiasm and dedication are essential to the Institute's continued success.

Administration Team Members

From the back, then left to right, the members of the Administration Team are:

Matthew Knefel
Program Support Technician

Shannon Worringham
Executive Assistant to the Director

Tom Smith
Undergraduate Facilities Assistant

Laura Lowe
Undergraduate Administrative Assistant

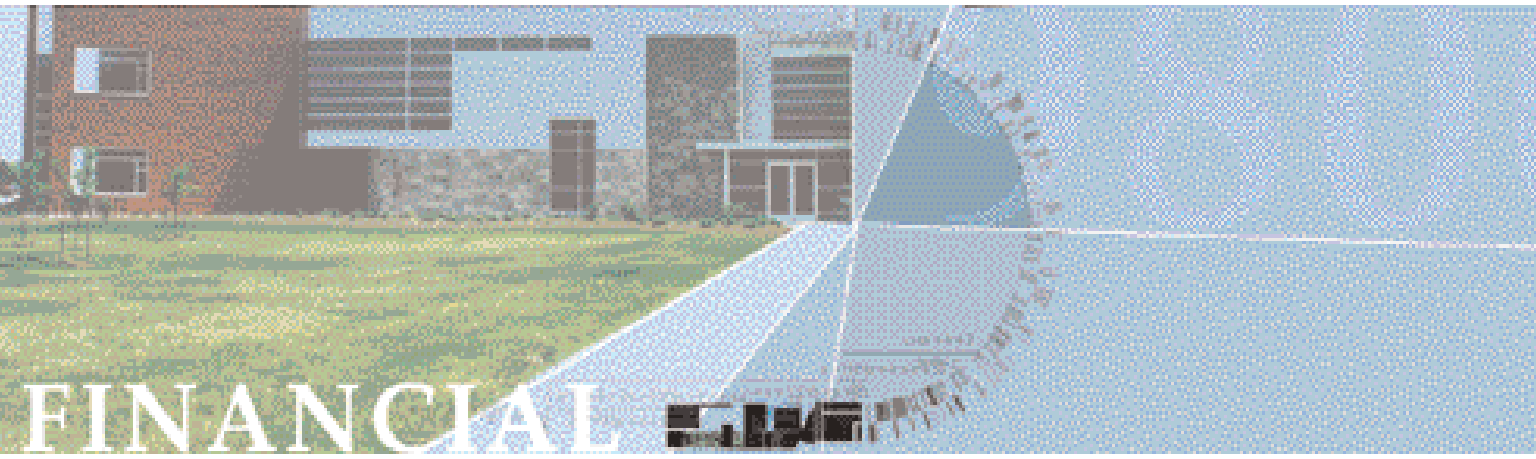
Stacey Lyons
Senior Fiscal Technician

Dawn Maxey
Facilties Manager

Lynn Hudson
Human Resources Coordinator

Debi Darnell
Human Resources Recruiter

Lauren Coble
Associate Director of Administration and Finance



VBI's research portfolio has spurred tremendous growth for the Institute since its inception. We have garnered support from a variety of sectors, from private foundations to corporations to government agencies. Specifically, VBI has received financial support from the Virginia Tobacco Indemnification and Community Revitalization Commission to conduct important research and outreach objectives, addressing regional, national, and international issues through bioinformatics.

Special thanks to our collaborators:

In collaborations with VBI, external researchers and institutions have received in excess of $10 million.

College of William and Mary (Williamsburg, Virginia)

Department of Energy Joint Genome Institute (Walnut Creek, California)

European Media Laboratory GMbH (Schloss-Wolfsbrunnenweg, Heidelberg, Germany)

IBM Watson Research Center (Yorktown Heights, New York)

INCOGEN, Inc. (Williamsburg, Virginia)

Johns Hopkins University Bloomberg School of Public Health (Baltimore, Maryland)

Laboratoire Bordelais de Recherche Informatique (LaBRI) (Bordelais, France)

Ohio State University (Wooster, Ohio)

Oklahoma State University (Stillwater, OK)

Phenomenome Discoveries, Inc. (Saskatoon, Saskatchewan, Canada)

Samuel Roberts Noble Foundation (Ardmore, Oklahoma)

Sun Microsystems, Inc. (Santa Clara, California)

University of California-Berkeley (Berkeley, California)

University of Nevada-Reno (Reno, Nevada)

Virginia Bioinformatics Consortium (University of Virginia, George Mason University, Virginia Commonwealth University)

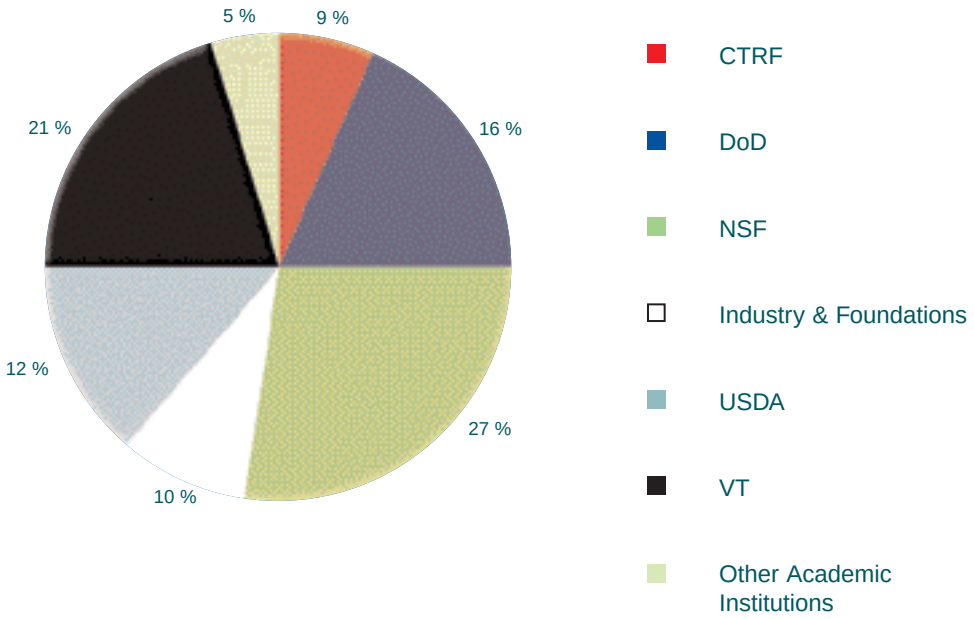
Departments within Virginia Tech:

Virginia Tech departments have garnered $5.7 million in funding through collaboration with VBI. Currently, VBI has partnerships with the following departments:

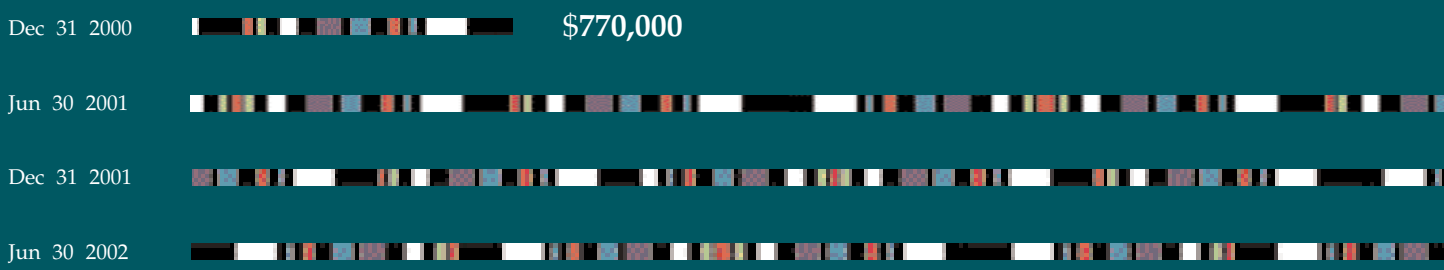
- **Biology**
- **Computer Science**
- **Crop and Soil Environmental Sciences**
- **Dairy Science**
- **Electrical and Computer Engineering**
- **Fisheries and Wildlife Sciences**
- **Horticulture**
- **Mathematics**
- **Plant Pathology, Physiology, and Weed Science**
- **Statistics**
- **Virginia-Maryland Regional College of Veterinary Medicine**

Financial Summary

VBI Research Portfolio by Funding Source



Research and Award Proposal Activity • December 31, 2000 to June 30, 2002



- VBI currently has 28 funded research projects totaling over $24 million. The dollar amounts on these two pages reflect awards attributable to VBI only.
- Research funding increased 30-fold from December 31, 2000 to June 30, 2002.
- VBI gives special thanks to the Virginia Tobacco Indemnification and Community Revitalization Commission for $11,558,314 in financial support (FY 00-02).

VBI Research Awards by Sponsor

Sponsor	Amount	Percentage
CTRF (Commonwealth Technology Research Fund)	$2,153,277	9 %
DoD (Department of Defense)	$4,000,000	16 %
NSF (National Science Foundation)	$6,621,841	27 %
Industry & Foundations	$2,429,135	10 %
USDA (US Department of Agriculture)	$2,973,614	12 %
VT (Virginia Tech)	$5,017,537	21 %
Other Academic Institutions	$1,118,391	5 %

$9,421,007

$19,945,290

$24,376,795

VBI Funded Grants

Sponsor	Investigator	Amount*	Project
Commonwealth Technology Research Fund	Bruno Sobral Dennis Kafura	$2,500,201	Collaborative Research in Bioinformatics
Commonwealth Technology Research Fund	Bruno Sobral Murray Black William Pearson Jeffrey Plank Gregory Buck	$1,500,000	Virginia Bioinformatics Consortium
Commonwealth Technology Research Fund	Dennis Manos, et al. Maciek Sasinowski Bruno Sobral	$3,251,901	Collaboration with the College of William and Mary and INCOGEN, Inc.
Department of Defense	Bruno Sobral	$4,000,000	PathPort: A Common Asset for Biological Security
IBM	Bruno Sobral	$893,220	Shared University Research
National Institutes of Health	Zhijian Tu Chunhong Mao	$1,357,113	Characterization and Organization of Transposable Elements
National Science Foundation	Allan Dickerman Brett Tyler	$100,000	Bioinformatics Prediction of Functions of Unculturable Microbes in Ecosystems
National Science Foundation	David Meinke, et al. Allan Dickerman David Patton, et al.	$852,207	Essential Gene Functions in *Arabidopsis* Seed Development
National Science Foundation	Reinhard Laubenbacher	$122,475	Algebraic Algorithms for Cell Complexes
National Science Foundation	Pedro Mendes	$99,999	Reverse Engineering of Biochemical Networks from Whole-Genome Dynamics
National Science Foundation	Pedro Mendes Richard Dixon	$3,587,432	An Integrated Approach to Functional Genomics and Bioinformatics in a Model Legume
National Science Foundation	Craig Nessler Boris Chevone Pedro Mendes	$99,999	Metabolic Engineering of Plant Vitamin C Biosynthesis for Improved Nutrition and Health
National Science Foundation	Brett Tyler Bruno Sobral	$499,633	Communication, Training, and Resources for *Phytophthora* Research
National Science Foundation	Brett Tyler M.A. Saghai Maroof Glenn Buss Ina Hoeschele Anne Dorrance Steve St. Martin	$6,764,465	Dissecting Soybean Resistance to *Phytophthora* by QTL Analysis of Host and Pathogen Expression Profiles
Phenomenome Discoveries, Inc.	Pedro Mendes	$228,429	Database and Visualization Systems for Metabolomics
Sun Microsystems, Inc.	Bruno Sobral Pedro Mendes	$1,262,486	Sun Center of Excellence in Bioinformatics
United States Department of Agriculture	Bruno Sobral Allan Dickerman	$444,230	Analysis of Plant Genome Duplication Events and Their Functional Relevance
United States Department of Agriculture	Brett Tyler Bruno Sobral	$2,529,384	Genome Sequence of *Phytophthora sojae*
United States Department of Agriculture	Brett Tyler	$230,000	Function of Avirulence Genes in *Phytophthora sojae* Infection of Soybean
Virginia Tech Diversity Initiative	Reinhard Laubenbacher	$1,000	Recruitment and Educational Development of Minority Mathematics Students
Virginia Tech/Johns Hopkins University	Bruno Sobral et al. Diane Griffin Alan Scott Nirbhay Kumar David Sullivan Ying Zhang Richard Markham	$10,000,000	Collaboration with John Hopkins Univ., Bloomberg School of Public Health
University of Nevada-Reno/Samuel Roberts Noble Foundation	Bruno Sobral John Cushman Greg May	$323,016	Collaborative Development of an EST Database and Analysis Pipeline

*Amounts reflect award totals for VBI, including internal and external collaborators.

The VBI 2002 Annual Report was created and designed by the team at  • the Visual Design Studio for Education, Outreach, Research, & Exhibition at Virginia Tech.

Tiffany Trent, Candace Baracat, and Dr. Neysa Call led the development of the 2002 Annual Report in collaboration with researchers at VBI.