

Toxicogenomics through the Eyes of Informatics: Conference Overview and Recommendations

Kenneth Olden,¹ Neysa Call,² Bruno Sobral,² and Robin Oakes²

¹National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina, USA; ²Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

Virginia Bioinformatics Institute, in conjunction with National Institutes of Environmental Health Sciences, hosted a conference, "Toxicogenomics through the Eyes of Informatics," in Bethesda, Maryland, USA, on 12–13 May 2003. Researchers around the world met to discuss how the application of bioinformatics tools, methodologies, and technologies will enhance our understanding of how cells and organisms respond to toxins. Conference topics included statistical methods, quantitative molecular data sets, computational algorithms for data analysis, computational modeling and simulation, challenges and opportunities in computational biology, and information technology infrastructure for data and tool management. This meeting report is a summary of conference presentations, survey results, current toxicogenomics concerns, and future directions of the toxicogenomics community. In conclusion this report discusses toxicogenomics as related to environmental agents, cell–chemical reactions, and gene–environment interactions. **Key words:** bioinformatics, cell–chemical reactions, environmental agents, toxicogenomics, toxins, Virginia Bioinformatics Institute. *Environ Health Perspect* 112:805–807 (2004). doi:10.1289/txg.6799 available via <http://dx.doi.org/> [Online 1 April 2004]

The availability of genome-scale DNA sequence information has radically altered life science research. Moreover, the human genome project has revolutionized the practice of biology and the potential of medicine. The field of toxicology will be greatly augmented by the application of the knowledge of genetics and its sister sciences. Consequently, researchers have entered a new area of specialty—toxicogenomics—the marriage of toxicology and genomics. Data generated by such research will impact many areas including health and environmental sciences. In addition, the science and health care communities will also be better able to identify a wide range of human health risks. However, to realize the potential of toxicogenomics, a global view of gene function and temporal and spatial regulations for all genes at different physiological or pathological states or developmental stages is paramount. The expanding field of bioinformatics offers the convergence of biological, mathematical, and computational platforms to better understand the effects of toxins on biological systems.

Scientists using data developed by the modern advent of biotechnology need novel and efficient bioinformatics tools to monitor gene–environment interactions on a large scale. These bioinformatics platforms (hardware, software, portals, and models, for example) will allow the conversion of data into useful knowledge to prevent disease resulting from environmental exposures.

Conference Overview

The conference "Toxicogenomics through the Eyes of Informatics" was held 12–13

May 2003 in Bethesda, Maryland, USA. The purpose of the conference was to bring together researchers around the world to discuss the application of bioinformatics tools, methodologies, and technologies in order to increase our understanding of how cells and organisms respond to toxins. By understanding the mechanisms from a genomic and a systems view, investigators can more rationally assess and evaluate the risk that toxins pose to humans. The format of this meeting aimed to facilitate input from the range of disciplines that can now provide insight into the field. Table 1 illustrates the conference program with speakers and events. By integrating experimental studies at the structural, molecular, and cellular levels with mathematical and biochemical modeling and informatics, the conference surveyed the integration and merger of advancing technologies. This conference provided a forum to explore the potential applications and implications of the new technologies, centering on toxicogenomics research platforms.

Recommendations and Survey Report

Researchers completed a survey about the future and importance of toxicogenomics (Table 2). The presentations and survey responses are summarized in the following observations:

- The complex characteristics of biological data generated by bioinformatics approaches (genomics, transcriptomics, metabolomics, proteomics) offer an opportunity to explore data analysis and standardization. Furthermore, data

analysis via mathematical and biochemical modeling is paramount in predicting the effects of toxins on biological systems.

- Novel visualization tools will be required for researchers to realize the usefulness of bioinformatics data in toxicogenomics research. For investigators to catalyze forward movement in these new directions, however, baseline (or normal) data for systems, including normal ranges of variation, must be generated and centrally stored in publicly available databases.
- To study the effects of toxins on biological organisms holistically, we must develop more sophisticated and robust computational methods and bioinformatics infrastructures for biological knowledge discovery, including novel [information technology (IT)] and database applications. In addition, predictive modeling will allow for greater understanding and, in turn, knowledge extraction from metabolic and proteomic profiles currently being generated.
- To develop and establish coordinated database and informatics platforms (e.g., portals) for data exchange and analysis, we need to define primary end-user needs and major issues facing relevant subdisciplines. The computational and networking needs for toxicogenomics research should be defined. Moreover, the system must integrate with similar efforts in other health-related disciplines to merge all areas of the life sciences in a scalable, flexible system. This effort will depend on exploring data standards and system interoperability concurrently. The development of toxicogenomics data generation center(s) of excellence linked via bioinformatics center(s) of excellence would allow the compilation, analysis, and sharing of relevant data and tools.

Address correspondence to B. Sobral, Virginia Bioinformatics Institute, Virginia Tech, Bioinformatics Facility 1 (0477), Blacksburg, VA 24061 USA. Telephone: (540) 231-1259. Fax: (540) 231-2606. E-mail: necall@vt.edu

We thank J. Selkirk from the National Institutes of Environmental Health Sciences for his scientific guidance. We also thank C. Phillips from Virginia Bioinformatics Institute for her assistance with conference arrangements and logistics.

The authors declare they have no competing financial interests.

Received 14 October 2003; accepted 31 March 2004.

- Toxicogenomics research programs should be developed to facilitate interoperability among the varying activities under way. Each individual research project will be more successful, ultimately, if it is part of a larger coordinated and synergistic approach.
- Mathematical modeling will continue to reveal qualitative patterns and define crucial parameters of biological systems. These efforts will facilitate the design of experiments and increase our understanding of complex biological processes. The process is spiral in that models suggest new hypotheses to be tested in new experiments, but the latter will provide data to improve the models. At each turn knowledge increases and uncertainty decreases.
- Several reports concur that the demand for qualified bioinformaticians far outstrips the current number of qualified graduates. The development of multidisciplinary educational and training programs in bioinformatics, and training for toxicogenomics research in particular, must be in step with scientific discovery.

- Programs to support additional toxicogenomics conferences and workshops would allow communities to foster communication and coordinate efforts.

Conclusion

Bioinformatics tools for disease research will catalyze our quest to understand how individuals differ in their susceptibility to environmental agents and how these susceptibilities change over time. As the inherent goal of systems biology continues to intertwine experimentation, mathematical modeling, and computation, enhancing the investment in each of these efforts will

hasten our understanding of cell–chemical interactions. At this level, functional, structural, and comparative analysis help provide insights into the environmental influence predisposing plants, animals, and humans to disease. Recent contributions of bioinformatics research have yielded molecular data on expressed genes (RNA), proteins, and metabolites. Unlike genomic sequencing (DNA), proteomic, metabolomic, and transcriptomic data provide “state” information that helps form a dynamic description of how the organism interacts with its environment. Age, developmental stage, general health, and

Table 2. Conference participant questionnaire.

In your view, what are critical areas of research—that if enhanced via additional funding—would enable the integration and use of mathematics, modeling, simulation, and IT infrastructure to enhance our understanding of cell–chemical interactions?

How do you think experimental approaches will be altered or enhanced via advances in mathematics, modeling, simulation, and IT infrastructure? For example, can experimental designs be altered to maximize data set utility?

How might mathematics, modeling, simulation, and IT infrastructure alter laboratory, macrostructures, and “omics” data generation and vice versa?

What do experimentalists require most of mathematics, modeling, simulation, and IT infrastructure to remove key bottlenecks?

Table 1. Toxicogenomics through the Eyes of Informatics: conference program and presentations.

Session/speaker	Affiliation	Title of presentation
IT infrastructure Barbara Eckman	IBM Life Sciences	Data integration and management for molecular and cell biology
Rolf Apweiler	European Bioinformatics Institute	Integration and standardization: driving forces in the development of database resources
Shankar Subramaniam	University of California San Diego	Bioinformatics challenges in the post-genomic era
Computational modeling and simulation Zoltan Oltvai Pedro Mendes	Northwestern University Virginia Bioinformatics Institute	Organization of cellular networks Inferring biochemical networks from functional genomic data
Hiroaki Kitano Reinhard Laubenbacher Herbert Sauro	Virginia Bioinformatics Institute Keck Graduate Institute	Luncheon address Reverse-engineering in computational immunology Data exchange and programmatic resource sharing: the Systems Biology Workbench (SBW), bioSPICE, and the Systems Biology Markup language
John Tyson Charles deLisi	Virginia Polytechnic Institute and State University Boston University	From molecular regulatory networks to cell physiology Inferring the function of genes and networks
Statistical methods and quantitative molecular data sets Bruce Weir Lloyd Sumner	North Carolina State University Noble Foundation	Association patterns in the human genome Proteomics and metabolomics: temporal snapshots of the biochemistry of <i>Medicago truncatula</i>
Carolyn Mattingly	Mount Desert Island Biological Laboratory	The Comparative Toxicogenomics Database (CTD): new perspectives through data integration and curation
John Quackenbush Ute Roessner-Tunali	The Institute for Genomic Research Max-Planck-Institute for Molecular Plant Physiology	Interpreting expression through biology Contrasting and combining information derived from the systems biology approaches of metabolite and transcript profiling
Computational algorithms for data analysis and visualization Heike Hofmann Lior Pachter Tarif Awad	Iowa State University University of California Berkeley Affymetrix, Inc.	Graphical opportunities in exploring microarray data Comparative genomics by phylogenetic shadowing Gene ontology graphs as mining and display tools for gene expression data
Keith Allen	Paradigm Genetics	Integrating metabolomic and gene expression data
Computational biology: looking forward Stefan Unger Congressman Virgil H. Goode Jr.	Sun Microsystems Fifth District of Virginia	Computational biology: challenges and opportunities Passing the hat

nutritional status have significant influence on disease outbreaks, whether caused by environmental pollution, toxic agents, or pathogens. An integrated approach to exposure assessment, epidemiology, and data collection, for example, would allow quicker identification of susceptibility genes and characterization of their functions. If bioinformatics data relating to human disease, regardless of cause, could be integrated and manipulated via bioinformatics platforms available to all research stakeholders, the scientific community

could achieve an integrated view of gene–environment interaction at the level of whole organisms.

This resulting comprehensive catalog of human gene variation and response would increase our understanding of chronic disease. Subsequently, our capability to control environmental toxin levels and pathogens that cause disease to plants, animals, and humans would increase. However, many of the clinical, scientific, and pathogen data generated today are not put into a meaningful and integrated

context that would enable problem solving via bioinformatics applications. Limited bioinformatics platforms allow biological annotations to be easily integrated with related host-baseline or broader environment or ecosystem data when available. A centrally funded program in bioinformatics research dedicated to the synthesis and informatics to support toxicogenomics research would provide for the realization of unprecedented opportunities to elucidate biological responses to environmental toxicants and stressors at the genomic level.
