Introduction

Proceedings of the 2009 MidSouth Computational Biology and Bioinformatics Society (MCBIOS) Conference

Jonathan D Wren*1, Yuriy Gusev2, Raphael D Isokpehi3, Daniel Berleant4, Ulisses Braga-Neto5, Dawn Wilkins6 and Susan Bridges7

Address: 1Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation; 825 N.E. 13th Street, Oklahoma City, OK 73104-5005, USA, 2Department of Surgery, Health Sciences Center, The University of Oklahoma, Oklahoma City, OK 73104, USA, 3College of Science, Engineering and Technology, Jackson State University, Jackson, MS, USA, 4Department of Information Science, University of Arkansas at Little Rock, Little Rock, AR 72204, USA, 5Department of Electrical and Computer Engineering, Texas A&M University, College Station, 77843-3128, USA, 6Computer & Information Science Department, The University of Mississippi, University, MS 38677, USA and 7Department of Computer Science and Engineering, Mississippi State University, Box 9637, Mississippi State, MS 39762, USA

E-mail: Jonathan D Wren* - jdwren@gmail.com; Yuriy Gusev - Yuriy-Gusev@ouhsc.edu; Raphael D Isokpehi - raphael.isokpehi@jsums.edu; Daniel Berleant - berleant@gmail.com; Ulisses Braga-Neto - ulisses@ece.tamu.edu; Dawn Wilkins - dwilkins@cs.olemiss.edu; Susan Bridges - bridges@cse.msstate.edu

*Corresponding author from Sixth Annual MCBIOS Conference. Transformational Bioinformatics: Delivering Value from Genomes Starkville, MS, USA 20–21 February 2009

Published: 08 October 2009

BMC Bioinformatics 2009, 10(Suppl 11):S1 doi: 10.1186/1471-2105-10-S11-S1

This article is available from: http://www.biomedcentral.com/1471-2105/10/S11/S1

© 2009 Wren et al; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

MCBIOS 2009 was held February 20–21, 2009 at the Hunter Henry Center in Starkville, Mississippi on the Mississippi State University campus. The conference was hosted by the four research universities in Mississippi comprising the Mississippi Computational Biology Consortium – Jackson State University (JSU), Mississippi State University (MSU), the University of Mississippi (UM), and the University of Southern Mississippi (USM). Dr. Dawn Wilkins of UM and Dr. Susan Bridges of MSU served as conference co-chairs. Keynote speakers were Dr. Laura Elnitski, Head of the Genomic Functional Analysis Section at the National Human Genome Research Institute (NHGRI), Howard Cash, President and CEO of Gene Codes Corporation, and Dr. Cathy Wu, Director of the Protein Information Resource (PIR).

Dr. Raphael Isokpehi of JSU organized a panel discussion entitled Careers in Computational Biology and Bioinformatics: Perspective from Employers and Students. Panelists were Dr. Ed Perkins of the U.S. Army Engineering Research and Development Center, Robert Cottingham of Oak Ridge National Laboratory (ORNL), Dr. Doris M. Kupfer, Federal Aviation Administration, James C. Fuscoe, FDA National Center for Toxicological Research, and Cynthia Jeffries, student intern at ORNL. Over 40 student and faculty attendees also participated in the Speed Networking event organized by Dr. Isokpehi and featured in a recent Science Careers article [1]. Each person in the event spent 3 minutes talking to twenty other participants in the two hour session providing students an opportunity to interact with employers and with students from other universities in the region.

Dr. Bindu Nanduri and Dr. Andy Perkins, both of MSU managed the setup, judging, and scoring of 80 posters. Monetary awards were provided by Dr. Ed Perkins. Student poster award winners were: first place – Eric Morales of the University of New Orleans (UNO), second place Teresia Buza of MSU, third place – Prashanti Manda of MSU and honorable mention to
Amanda Alba of the UNO, Surya Saha of MSU and Sudhir Chowbina of the Indiana University School of Informatics. Student winners for oral presentations were: first place – Enis Afgan of the University of Alabama at Birmingham, second place – Lipi R. Archarya of UNO, and third place – Anastasia Chueva of Mississippi Valley State University.

Proceedings summary
Thirty-two papers from the 2009 conference were submitted to be considered for inclusion in this supplement, and of them a total of 20 were accepted (62.5% accept rate), making this year’s Proceedings the most stringent to date. The number of submitted papers exceeded that of the 2008 MCBIOS Proceedings [2-20]. As in prior conferences, we strove to be inclusive yet rigorous in the peer-review process, so that the end result is both high quality and reflective of the work presented. Papers generally fell into five categories:

Systems biology
A means of better understanding how systems work and interconnect as opposed to the traditional approach of isolating subcomponents and focusing on them is called Systems Biology and has become an increasingly active area of study in bioinformatics [21-34]. Andrey Ptitsyn offers a systems biology approach to the analysis of microarray data [35], partitioning gene expression space into a multi-dimensional taxonomy. Differences between “taxa” are then studied using biological pathway analysis software to associate gene expression patterns with specific phenotypes. For example, the application identifies a link between glycine metabolism aberrations and metabolic biomarkers for metastatic prostate cancer.

Gao et al. [36] describe an approach to better integrate data from disparate sources to reach better conclusions about biological response to stimulus than could be obtained from any of the component sources alone. Their study combined the Gene Ontology (GO) with microarray data using a bigraph data structure and associated algorithms.

Chowbina et al. address data integration from a different perspective[37], describing a technical resource for regulatory, signaling and biochemical reactions in human biomolecular pathways. Their resource, Human Pathway Database (HPD), integrates several existing resources to provide a web interface to access approximately one thousand pathways.

Zhang et al. present text empirics in the context of the PathBinder system [38]. Text empirics refers to properties of texts that are derived by examining the texts themselves. It differs from the machine learning approach in that derived properties are manually curated and presented as facts that system builders can use when they are designing their systems.

OMICS
“OMICS” is a broad category, based on the fact that fields of study for biological entities often end with the suffix “omics” (e.g., transcriptomics, proteomics) [39-45]. Pendarvis et al. [46] describe the creation of a workflow system to analyze mass spectrometry replicate datasets to generate a list of identified proteins and expression changes and link them to biological knowledge. Because it is compliant with proteomic data submission guidelines, users can readily publish data to public proteomic data repositories.

Mete et al. [47] present a novel digital image analysis approach to the automated evaluation of angiogenesis in Liver Cancer. The method identifies regional, morphological and fractal features and helps measure micro-vessel density on digitized images of liver tumor sections. Their results show agreement between automated calculations and manual counting of micro-vessels.

Microarray studies
Microarrays have long been a subject of bioinformatics analysis not only for better understanding transcription, but also as a model for large-scale analysis of genetic behavior [48-56]. Van Den Berg et al. compared gene annotation enrichment tools for functional modeling of agricultural microarray data [57], a category that is underrepresented among current tools. Chavan et al. describe Network Analysis Toolbox in R (NATBox), which provides a menu-driven GUI for modeling and analysis of functional relationships inferred from gene expression profiles. It is suited for interdisciplinary researchers and biologists with minimal programming experience who would like to use systems biology approaches without delving into the algorithmic aspects, and can be a useful demonstration/teaching tool.

Perkins and Langston [59] identify a means of effectively selecting thresholds for gene expression on microarrays by identifying “cliques” or subgroups of co-expressed genes. Roughly, the idea being that individual measurements are prone to error, but when a gene is seen co-expressed with other genes, then this reinforces the validity of the expression being biologically significant.

Li et al. [60] investigate inter- and intra-platform consistency of microarray measurements. Using pathway
data as a guide for identifying biologically consistent results, the authors demonstrate that highly consistent biological information can be generated from different microarray platforms.

**Genomic analysis**

T.J. Jankun-Kelly and colleagues present a method to visually explore conserved domains on Multiple Sequence Alignments [61], simultaneously communicating the relative similarity of proteins across species and the differences in how function is expressed via conserved domains. It quickly identifies conserved domains and allows both macro (sequence-long) and micro (small amino-acid neighborhood) views.

Xu et al. [62] report the results of an extensive comparative genomics analysis of plant lignin biosynthesis gene families. Lignin is a major component of plant cell walls and appears to be the major cause of cell wall recalcitrance in energy production. Their analysis confirms that, among the species analyzed, the first complete lignin biosynthesis pathway appeared in moss. The major expansion of biosynthesis families occurred after the divergence of monocots and dicots and duplicated genes had many different fates. They conclude that transgenic lignin modification strategies to bioenergy feedstock may only be successful between closely related species.

T. Buza et al. report a method of automated functional annotation of the chicken genome, improving both the breadth and quality of annotation of genes [63]. The method can facilitate functional annotation of other microarrays via an Array GO Mapper (AGOM) tool to help researchers quickly retrieve corresponding functional information.

Zhou et al. [64] develop CPTRA (Cross-Platform Transcriptome Analysis) to analyze Next Generation Sequencing based on Transcriptome profiling in species with limited genome information.

**Miscellaneous**

Garcia-Reyero et al. [65] study gene expression in fathead minnows (*Pimephales promelas*) in well-characterized sites adjacent to sewage treatment plants and find short-term exposure to effluents were able to induce a site-specific gene expression pattern in the fathead minnow gonad and liver and to affect fish sexual behavior.

There are several other papers published in the supplement but, for space, are not summarized here: Chen and Johnson [66], Bright et al. [67], Zollanvari et al. [68] and Malone et al. [69].

**Future meetings**

MCBIOS 2010 will be held in Jonesboro, Arkansas, and hosted by Arkansas State University and the Arkansas Biosciences Institute. Daniel Berleant of University of Arkansas at Little Rock is the 2009–2010 President of MCBIOS and will be chair of MCBIOS 2010. Ulisses Braga-Neto of Texas A&M University is the new President-elect of the society. See http://www.MCBIOS.org for information regarding MCBIOS and future meetings. MCBIOS is a regional affiliate of the International Society for Computational Biology http://www.ISCB.org.

**Competing interests**

The authors declare that they have no competing interests.

**Authors’ contributions**

All authors served as co-editors for these proceedings, with JDW serving as Senior Editor. All authors helped write this editorial.

**Acknowledgements**

We would like to thank the Mississippi NSF EPSCoR for providing support for keynote speakers and the hard work of the peer-reviewers to help produce this supplement.

This article has been published as part of BMC Bioinformatics Volume 10 Supplement 11, 2009: Proceedings of the Sixth Annual MCBIOS Conference. Transformational Bioinformatics: Delivering Value from Genomes. The full contents of the supplement are available online at http://www.biomedcentral.com/1471-2105/10?issue=S11.

**References**

1. Holmes L: Perspective: Speed Networking for Scientists. Science Careers 2009.
2. Bottoms CA and Xu D: Wanted: unique names for unique atom positions. PDB-wide analysis of diastereotopic atom names of small molecules containing diphosphate. BMC Bioinformatics 2008, 9(Suppl 9):S16.
3. Chae M, Shmookler Reis RJ and Thaden JJ: An iterative blockshifting approach to retention time alignment that preserves the shape and area of gas chromatography-mass spectrometry peaks. BMC Bioinformatics 2008, 9(Suppl 9):S15.
4. Churbanov A and Winters-Hilt S: Clustering ionic flow blockade toggles with a mixture of HMMs. BMC Bioinformatics 2008, 9(Suppl 9):S13.
5. Dozmorov MG, Kyker KD, Hauser PJ, Saban R, Bueche DD, Dozmorov I, Centola MB, Culkin DJ and Hurst RE: From microarray to biology: an integrated experimental, statistical and in silico analysis of how the extracellular matrix modulates the phenotype of cancer cells. BMC Bioinformatics 2008, 9(Suppl 9):S4.
6. Frank RL, Kandoth C and Erical F: Validation of an NSP-based (negative selection pattern) gene family identification strategy. BMC Bioinformatics 2008, 9(Suppl 9):S2.
7. Giles CB and Wren JD: Large-scale directional relationship extraction and resolution. BMC Bioinformatics 2008, 9(Suppl 9):S11.
8. Hong H, Su Z, Ge W, Shi L, Perkins R, Fang H, Xu J, Chen JJ, Han T and Kaput J, et al: Assessing batch effects of genotype calling
58. Chavan SS, Bauer MA, Scutari M and Nagarajan R: **NATbox: A Network Analysis Toolbox in R.** *BMC Bioinformatics* 2009, 10(Suppl 11):S14.
59. Perkins AD and Langston MA: **Threshold selection in gene co-expression networks using spectral graph theory techniques.** *BMC Bioinformatics* 2009, 10(Suppl 11):S4.
60. Li Z, Su Z, Wen Z, Shi L and Chen T: **Microarray Platform Consistency Is Revealed by Biologically Functional Analysis of Gene Expression Profiles.** *BMC Bioinformatics* 2009, 10(Suppl 11):S12.
61. Jankun-Kelly TJ, Lindeman AD and Bridges SM: **Exploratory Visual Analysis of Conserved Domains on Multiple Sequence Alignments.** *BMC Bioinformatics* 2009, 10(Suppl 11):S7.
62. Xu Z, Zhang D, Hu J, Zhou X, Ye X, Reichel KL, Stewart NR, Syrenne RD, Yang X and Gao P, et al: **Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom.** *BMC Bioinformatics* 2009, 10(Suppl 11):S3.
63. Buza TJ, Kumar R, Gresham CR, Burgess SC and McCarthy FM: **Facilitating Functional Annotation of Chicken Microarray Data.** *BMC Bioinformatics* 2009, 10(Suppl 11):S2.
64. Zhou X, Su Z, Sammons RD, Peng Y, Tranel PJ, Stewart N Jr and Yuan JS: **Novel Software Package for Cross-Platform Transcriptome Analysis (CPTRA).** *BMC Bioinformatics* 2009, 10(Suppl 11):S16.
65. Garcia-Reyero N, Adelman IR, Martinović D, Liu L and Denslow ND: **Site-specific impacts on gene expression and behavior in fathead minnows (Pimephales promelas) exposed in situ to streams adjacent to sewage treatment plants.** *BMC Bioinformatics* 2009, 10(Suppl 11):S11.
66. Chen B and Johnson M: **Protein Local 3D Structure Prediction by Super Granule Support Vector Machines (Super GSVM).** *BMC Bioinformatics* 2009, 10(Suppl 11):S15.
67. Bright L, Swiderski C, Burgess S, Chowdhary B and McCarthy FM: **Structural and Functional-annotation of an Equine Whole Genome Oligoarray.** *BMC Bioinformatics* 2009, 10(Suppl 11):S8.
68. Zollanvari A, Cunningham MJ, Braga-Neto U and Dougherty ER: **Analysis and Modeling of Time-Course Gene-Expression Profiles from Nanomaterial-Exposed Primary Human Epi-dermal Keratinocytes.** *BMC Bioinformatics* 2009, 10(Suppl 11):S10.
69. Malone BM, Perkins AD and Bridges SM: **Integrating phenotype and gene expression data for predicting gene function.** *BMC Bioinformatics* 2009, 10(Suppl 11):S21.