

INTRODUCTION

Open Access

Proceedings of the 2010 MidSouth Computational Biology and Bioinformatics Society (MCBIOS) Conference

Jonathan D Wren^{1*}, Doris M Kupfer², Edward J Perkins³, Susan Bridges⁴, Daniel Berleant⁵

From Seventh Annual MCBIOS Conference. Bioinformatics: Systems, Biology, Informatics and Computation Jonesboro, AR, USA. 19-20 February 2010

Introduction

The seventh annual Midsouth Computational Biology and Bioinformatics (MCBIOS) conference took place February 19 and 20, 2010 at Arkansas State University in Jonesboro, AR, presided over by Daniel Berleant, this year's President of MCBIOS. Keynote speakers were Elaine Ostrander of NIH, renowned for her work on dog genomics, Clayton Naeve, CIO of St. Jude Children's Research Hospital, and Robert Cottingham, Leader of the Computational Biology and Bioinformatics Group, Oak Ridge National Laboratory. The record-breaking attendance exceeded 200 participants which necessitated parallel talk sessions for the first time. Student oral presentation award winners were Heidi Pagan (first place) of Mississippi State University (MSU), Juliet Tang (second place) of MSU, and Aleksandra Markovets (third) of Mississippi Valley State University. In addition, a record number of posters were also presented. Due to the number of posters, student poster awards were given in two categories, Biological Focus and Computational Focus. Winners were N. Platt and V. Chaitankar (1st place), G. Cooper and L. Pillai (2nd), and M. Ammari and C. Gearheart (3rd). Reflecting the integration of these foci in the field, winners in one category frequently scored high in the other category as well. MCBIOS is also pleased to have this year acquired legal status as a non-profit organization.

Proceedings summary

A total of twenty eight papers describing research presented at the 2010 conference were accepted for

publication in these proceedings out of a total of forty three papers submitted for consideration (~65%). The number of papers submitted for consideration was the largest submitted since the inception of MCBIOS with a 34% increase over last year's proceedings [1-19], and the highest number of papers submitted to date, reflecting a healthy growth in conference participation. All papers were peer-reviewed by at least two reviewers. Our goal for inclusion of papers was to be inclusive, yet rigorous in the peer-review process such that accepted papers are both high quality and reflective of the work presented at the conference. Papers generally fell into five categories:

Genomic analysis

Bioinformatics was largely born as a field through the need to analyze and understand sequence data. Partly as a result of ultra-high-throughput sequencing technology, the field has seen a strong resurgence recently with new computational methods being developed and applied to downstream applications such as the identification of human mutations [20-25] as well as meta-genomics and whole genome sequencing [26-31]. In these proceedings, several new methods and applications are reported. Hong Fang, *et al.* describe an expansion of the FDA-developed, ArrayTrack™ genomics tool, into a platform that supports microbial data to enable the rapid detection of food borne pathogenic bacteria during outbreak scenarios [32]. Xu *et al.* describe new SNP (single nucleotide polymorphism) and QTL (quantitative trait locus) libraries for the ArrayTrack system which provide users with extensive data including links to other resources, and motivated in part by the need to augment gene association study results with biological context [33]. Christopher Bottoms *et al.* report on the new web-based tool, IView, which graphically displays and makes

* Correspondence: jdwwren@gmail.com

¹Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation; 825 N.E. 13th Street, Oklahoma City, OK 73104-5005, USA
Full list of author information is available at the end of the article

searchable introgression data by marker name, chromosome number, or map position [34].

Quest and colleagues working in the Cottingham laboratory address the problem of brittle genome annotation systems with an OWL-based approach that provides several significant advantages [35]. Work reported by Markovets and colleagues on identifying promoter regions showed that a new neural network-based algorithm outperformed the method of stress-induced duplex destabilization [36]. Smits and Ouverney present a software system that assists users in analyzing sequence data to find phylogenetic trees, while avoiding pitfalls associated with this process as commonly practiced [37].

An examination of substitution rates over evolutionary time by Ulrich Melcher compared viral and non-viral sequences to reveal that the rate of nucleotide interchanges in plant viruses is not constant, impacting phylogeny studies. The evidence suggests that different evolutionary rules may apply to recent divergences than to those linked to distant speciation events [38].

An in depth analysis of the two forms of Bovine Viral Diarrhea, a worldwide cattle disease, was completed by Mais Ammari *et al.* using a proteomics approach to evaluate the host protein expression differences. Two gene sets were found, each specific to one disease form which showed significant functional differences [39]. Arun Rawat *et al.* address the need for additional genomic information for Northern bobwhite, an avian species critical to ecosystem maintenance. Using next generation sequencing of cDNAs from a multi-tissue library, they generated ESTs which were pipelined into a uni-gene database and made publicly available at a searchable website [40].

Cyriac Kandoth *et al* [41] mined the *Arabidopsis thaliana* genome using a new method to scan genome sequences for partially symmetric inverted repeats that might represent miRNA precursors. Chao Di *et al* [42] carried out a comparative genome analysis of the Prohibitin gene, important functions from diseases to development, finding that the gene family is conserved across different kingdoms falling into five different clades and indicating that different duplication events were involved for gene family expansion, especially the segmental duplications in *Arabidopsis*. The conserved gene evolution indicated that the study in the model species can be translational to human disease studies. Rowena Kelly *et al* [43] have constructed a database with a web-based interface that integrates large datasets on *Aspergillus flavus* resistance in Maize such as gene expression, proteomic, QTL genetic mapping studies and sequence data from the literature to help identify and select candidate *A. flavus* resistance genes.

3D/structural analysis

Biological information may be encoded in a one-dimensional sequence, but its manifestation takes place in three-dimensional space in the form of proteins, cells and tissues among other structures. Computational analysis and prediction of these structures is challenging, to say the least, and an active area of recent bioinformatics interest [44-48]. Computational simulations are critical for testing proposed solutions, as well as in rationally engineering more specific binding/activity for drug candidates.

To better understand and improve of the creation of anti-inflammatory drugs through inhibition of lipoxigenase, an essential enzyme in the inflammatory pathway, Shuju Bai *et al.* report on a method for modeling interactions between 8R-lipoxigenase and its substrate [49]. A mixed reality surgical simulator with a VICON motion tracking system, developed for the rasping procedure in artificial cervical disc replacement surgeries is described by Tansel Halic *et al.*, which can serve as an important training and practice tool for surgeons [50]. The potential of using hydrogen deuterium exchange (HDX) mass spectrometry for analyzing enzymes that degrade cell walls is explored by Uzuner *et al.*[51]. They report that the HDX mass spectrometry can be a powerful tool for exploring the molecular mechanisms of enzyme functions.

Sinan Kockara *et al* [52] and Mutlu Mete *et al* [53] examined ways to automatically quantitate and define the extent of malignant melanomas on skin while offering speed and consistency in detection of the lesions borders, elements especially useful in areas without dermatologists. Kockara *et al* compared density based clustering and Fuzzy C-Means clustering algorithms for border detection in dermoscopy images finding that density based clustering performed best with a low border error, high precision and recall; however, the Fuzzy C-Means clustering algorithm had poor performance especially in border detection. Mete *et al* proposed two new methods, boundary driven density based clustering based algorithms which performed better delineation with noisy images and an active contour model that gave the best detection with optimum parameters. Denise Koessler *et al* [54] developed a tool to predict secondary RNA structures based on the novel approach of using graph-theoretic values as input for a neural network and computes the probability of secondary RNA structure.

Systems biology

There are many areas within biology that are amenable to computational analysis, each of which is usually approached separately. Somewhat recently, the term “systems biology” has emerged to describe

interconnected analyses, particularly those that help reduce complexity in these systems, as defined by the number of interconnections between component parts, into a more functional understanding (e.g., input and output). The term “systems biology” is often used broadly and not always consistently, but is an emergent area of high interest in bioinformatics [55-57].

A novel method for identifying subnetwork markers in a human protein-protein interaction network is described by Junjie Su *et al.*[58]. They report the identification of subnetwork markers that improve upon current gene-based and pathway-based markers in their discriminative power, reproducibility and classification. Elina Tjioe *et al.* report on the development of a Web-based gene-discovery bioinformatics tool, FAUN (Feature Annotation Using Nonnegative matrix factorization), which identifies implied gene relationships from biomedical literature, enabling hypotheses on functional characterization [59].

A paper by Vijender Chaitankar, *et al.*, explores the important problem of inferring gene regulatory networks from time sliced microarray data. They claim that mutual information (MI) based approaches have inherent limits to their capabilities in this context which they were able to improve upon with algorithms based on modifications to the mutual information metric [60].

Kumar and Nanduri present a downloadable database, HPIDB, which serves as a resource of host-pathogen protein-protein interactions integrated non-redundantly from public databases. The authors report on the flexibility of the database for querying and an output format which allows portage to downstream applications [61].

Microarray studies

Microarrays have long been a subject of bioinformatics analysis not only for better understanding transcription, a process affected by both genetic and epigenetic factors [62,63], but also as a model for large-scale analysis of genetic behavior.

A new non-stationary Dynamic Bayesian Network with a flexible lag choosing mechanism that detects potential genetic regulators and has improved structure prediction is reported by Yi Jia and Jun Huan [64].

Zhining Wen *et al.* address two interesting QC questions of whether expired Affymetrix GeneChip microarrays, an expensive resource, are still useful up to four years after expiration and if RNA stored at -80°C for the same period was suitable as template source. In spite of a decrease in sensitivity, the authors found that these arrays generated data consistent with that from unexpired arrays and report favorably on the stability of the RNA [65].

Miscellaneous

There are a few other papers published in the proceedings but, for space, are not summarized here: [66-69].

Future meetings

MCBIOS 2011 will be held in College Station, Texas. The 2010-2011 President of MCBIOS is Ulisses Braganeto of the University of Texas A&M, and Susan Bridges of Mississippi State University is now the President-elect. MCBIOS is a regional affiliate of the International Society for Computational Biology (<http://www.ISCB.org>). For information regarding MCBIOS and our future meetings, see <http://www.MCBIOS.org>.

Acknowledgements

We would like to thank the many anonymous peer reviewers who helped to ensure the quality of these Proceedings, and Dr. Kristi Koelsch for help with manuscript development. EP was supported by the U.S. Army Environmental Quality Technology Research Program. Permission to publish this information was granted by the Chief of Engineers.

This article has been published as part of *BMC Bioinformatics* Volume 11 Supplement 6, 2010: Proceedings of the Seventh Annual MCBIOS Conference. Bioinformatics: Systems, Biology, Informatics and Computation. The full contents of the supplement are available online at <http://www.biomedcentral.com/1471-2105/11?issue=S6>.

Author details

¹Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation; 825 N.E. 13th Street, Oklahoma City, OK 73104-5005, USA .
²Aerospace Medical Institute, Federal Aviation Administration, Oklahoma City, OK 73169, USA .
³US Army Engineer Research and Development Center, 3909 Halls Ferry Road, Vicksburg, MS 39180, USA .
⁴Department of Computer Science and Engineering, Mississippi State University, Box 9637, Mississippi State, MS 3976, USA .
⁵Department of Information Science, University of Arkansas at Little Rock, Little Rock, AR 72204, USA.

Author contributions

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

Competing interests

The authors have no competing interests to declare.

Published: 7 October 2010

References

1. Bright LA, *et al*: Structural and functional-annotation of an equine whole genome oligoarray. *BMC Bioinformatics* 2009, **10**(Suppl 11):S8.
2. Buza TJ, *et al*: Facilitating functional annotation of chicken microarray data. *BMC Bioinformatics* 2009, **10**(Suppl 11):S2.
3. Chavan SS, *et al*: NATbox: a network analysis toolbox in R. *BMC Bioinformatics* 2009, **10**(Suppl 11):S14.
4. Chen B, Johnson M: Protein local 3D structure prediction by Super Granule Support Vector Machines (Super GSVM). *BMC Bioinformatics* 2009, **10**(Suppl 11):S15.
5. Chowbina SR, *et al*: HPD: an online integrated human pathway database enabling systems biology studies. *BMC Bioinformatics* 2009, **10**(Suppl 11):S5.
6. Gao C, *et al*: Graph ranking for exploratory gene data analysis. *BMC Bioinformatics* 2009, **10**(Suppl 11):S19.
7. Garcia-Reyero N, *et al*: 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.

8. Jankun-Kelly TJ, Lindeman AD, Bridges SM: **Exploratory visual analysis of conserved domains on multiple sequence alignments.** *BMC Bioinformatics* 2009, **10**(Suppl 11):S7.
9. Li Z, et al: **Microarray platform consistency is revealed by biologically functional analysis of gene expression profiles.** *BMC Bioinformatics* 2009, **10**(Suppl 11):S12.
10. Malone BM, Perkins AD, Bridges SM: **Integrating phenotype and gene expression data for predicting gene function.** *BMC Bioinformatics* 2009, **10**(Suppl 11):S20.
11. Mete M, et al: **Automatic identification of angiogenesis in double stained images of liver tissue.** *BMC Bioinformatics* 2009, **10**(Suppl 11):S13.
12. Pendarvis K, et al: **An automated proteomic data analysis workflow for mass spectrometry.** *BMC Bioinformatics* 2009, **10**(Suppl 11):S17.
13. Perkins AD, Langston MA: **Threshold selection in gene co-expression networks using spectral graph theory techniques.** *BMC Bioinformatics* 2009, **10**(Suppl 11):S4.
14. Ptitsyn A: **Computational analysis of gene expression space associated with metastatic cancer.** *BMC Bioinformatics* 2009, **10**(Suppl 11):S6.
15. van den Berg BH, et al: **Comparing gene annotation enrichment tools for functional modeling of agricultural microarray data.** *BMC Bioinformatics* 2009, **10**(Suppl 11):S9.
16. Xu Z, et al: **Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom.** *BMC Bioinformatics* 2009, **10**(Suppl 11):S3.
17. Zhang L, et al: **PathBinder—text empirics and automatic extraction of biomolecular interactions.** *BMC Bioinformatics* 2009, **10**(Suppl 11):S18.
18. Zhou X, et al: **Novel software package for cross-platform transcriptome analysis (CPTRA).** *BMC Bioinformatics* 2009, **10**(Suppl 11):S16.
19. Zollanvari A, et al: **Analysis and modeling of time-course gene-expression profiles from nanomaterial-exposed primary human epidermal keratinocytes.** *BMC Bioinformatics* 2009, **10**(Suppl 11):S10.
20. Holden M, et al: **GSEA-SNP: applying gene set enrichment analysis to SNP data from genome-wide association studies.** *Bioinformatic* 2008, **24**(23):2784-5.
21. Homer N, et al: **Multimarker analysis and imputation of multiple platform pooling-based genome-wide association studies.** *Bioinformatics* 2008, **24**(17):1896-902.
22. Pirinen M: **Estimating population haplotype frequencies from pooled SNP data using incomplete database information.** *Bioinformatics* 2009, **25**(24):3296-302.
23. Wegrzyn JL, et al: **PineSAP—sequence alignment and SNP identification pipeline.** *Bioinformatics* 2009, **25**(19):2609-10.
24. O'Dushlaine C, et al: **The SNP ratio test: pathway analysis of genome-wide association datasets.** *Bioinformatics* 2009, **25**(20):2762-3.
25. Manske HM, Kwiatkowski DP: **SNP-o-matic.** *Bioinformatics* 2009, **25**(18):2434-5.
26. Monzoorul Haque M, et al: **Sort-ITEMS: Sequence orthology based approach for improved taxonomic estimation of metagenomic sequences.** *Bioinformatics* 2009, **25**(14):1722-30.
27. Hu GQ, et al: **MetaTISA: Metagenomic Translation Initiation Site Annotator for improving gene start prediction.** *Bioinformatics* 2009, **25**(14):1843-5.
28. Huang Y, Gilna P, Li W: **Identification of ribosomal RNA genes in metagenomic fragments.** *Bioinformatics* 2009, **25**(10):1338-40.
29. Kristiansson E, Hugenholtz P, Dalevi D: **ShotgunFunctionalizer: an R-package for functional comparison of metagenomes.** *Bioinformatics* 2009, **25**(20):2737-8.
30. Kim JH, et al: **HAPLOWSER: a whole-genome haplotype browser for personal genome and metagenome.** *Bioinformatics* 2009, **25**(18):2430-1.
31. Mitra S, Klar B, Huson DH: **Visual and statistical comparison of metagenomes.** *Bioinformatics* 2009, **25**(15):1849-55.
32. Fang H, et al: **An FDA Bioinformatics Tool for Microbial Genomics Research on Molecular Characterization of Bacterial Foodborne Pathogens Using Microarrays.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S4.
33. Xu J, et al: **Two New ArrayTrack Libraries for Personalized Biomedical Research.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S6.
34. Bottoms CA, Flint-Garcia S, McMullen MD: **View: Introgression library visualization and query tool.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S28.
35. Quest DJ, et al: **Next Generation Models for Storage and Representation of Microbial Biological Annotation.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S15.
36. Bland C, Newsome AS, Markovets AA: **Promoter prediction in E. coli based on SIDD profiles and Artificial Neural Networks.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S17.
37. Smits SA, Ouverney CC: **Phylometrics: a pipeline for inferring phylogenetic trees from a sequence relationship network perspective.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S18.
38. Melcher U: **Assessing constancy of substitution rates in viruses over evolutionary time.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S3.
39. Ammarī M, et al: **Analysis of Bovine Viral Diarrhea Viruses-infected monocytes: identification of cytopathic and non-cytopathic biotype differences.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S9.
40. Rawat A, et al: **Quail Genomics: a knowledgebase for Northern bobwhite.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S13.
41. Kandoth C, Ercal F, Frank RL: **A framework for automated enrichment of functionally significant inverted repeats in whole genomes.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S20.
42. Di C, et al: **Comparative genome analysis of PHB gene family reveals deep evolutionary origins and diverse gene function.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S22.
43. Kelley RY, et al: **Integrated Database for Identifying Candidate Genes for Aspergillus flavus Resistance in Maize.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S25.
44. La D, et al: **3D-SURFER: software for high-throughput protein surface comparison and analysis.** *Bioinformatics* 2009, **25**(21):2843-4.
45. Lee S, Blundell TL: **BIPA: a database for protein-nucleic acid interaction in 3D structures.** *Bioinformatics* 2009, **25**(12):1559-60.
46. Preibisch S, Saalfeld S, Tomancak P: **Globally optimal stitching of tiled 3D microscopic image acquisitions.** *Bioinformatics* 2009, **25**(11):1463-5.
47. Mann M, et al: **CPSP-web-tools: a server for 3D lattice protein studies.** *Bioinformatics* 2009, **25**(5):676-7.
48. Stratmann D, van Heijenoort C, Guittet E: **NOEnet—use of NOE networks for NMR resonance assignment of proteins with known 3D structure.** *Bioinformatics* 2009, **25**(4):474-81.
49. Bai S, Du T, Khosravi E: **Applying internal coordinate mechanics to model the interactions between 8R-lipoxygenase and its substrate.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S2.
50. Halic T, et al: **Mixed reality simulation of rasping procedure in artificial cervical disc replacement (ACDR) surgery.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S11.
51. Uzuner U, et al: **Enzyme Structure Dynamics of Xylanase I from Trichoderma longibrachiatum.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S12.
52. Kockara S, et al: **Analysis of density based and fuzzy c-means clustering methods on lesion border extraction in dermoscopy images.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S26.
53. Mete M, Sirakov NM: **Lesion detection in dermoscopy images with novel density-based and active contour approaches.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S23.
54. Koessler D, et al: **A predictive model for secondary RNA structure using graph theory and a neural network.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S21.
55. Bois FY: **GNU MCSim: Bayesian statistical inference for SBML-coded systems biology models.** *Bioinformatics* 2009, **25**(11):1453-4.
56. Reyes-Palomares A, et al: **Systems biology metabolic modeling assistant: an ontology-based tool for the integration of metabolic data in kinetic modeling.** *Bioinformatics* 2009, **25**(6):834-5.
57. Mirschel S, et al: **PROMOT: modular modeling for systems biology.** *Bioinformatics* 2009, **25**(5):687-9.
58. Su J, Yoon B, Dougherty ER: **Identification of diagnostic subnetwork markers for cancer in human protein-protein interaction network.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S8.
59. Tjioe E, Berry MW, Homayouni R: **Discovering gene functional relationships using FAUN (Feature Annotation Using Nonnegative matrix factorization).** *BMC Bioinformatics* 2010, **11**(Suppl 6):S14.
60. Chaitankar V, et al: **Time lagged information theoretic approaches to the reverse engineering of gene regulatory networks.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S19.
61. Kumar R, Nanduri B: **HPIDB - a unified resource for host-pathogen interactions.** *BMC Bioinformatics* 2010, **11**(Suppl 6):S16.
62. Hughes T, et al: **DNA methylome in human CD4+ T cells identifies transcriptionally repressive and non-repressive methylation peaks.** *Genes Immun* 2010.

63. Sawalha AH: Epigenetics and T-cell immunity. *Autoimmunity* 2008, **41**(4):245-52.
64. Jia Y, Huan J: Constructing Non-Stationary Dynamic Bayesian Networks with a Flexible Lag Choosing Mechanism. *BMC Bioinformatics* 2010, **11**(Suppl 6):S27.
65. Wen Z, *et al*: Evaluation of gene expression data generated from expired Affymetrix GeneChip® microarrays using MAQC reference RNA samples. *BMC Bioinformatics* 2010, **11**(Suppl 6):S10.
66. Ding D, *et al*: The EDKB: an Established Knowledge Base for Endocrine Disrupting Chemicals. *BMC Bioinformatics* 2010, **11**(Suppl 6):S5.
67. Ghosh P, *et al*: Dynamics of Protofibril Elongation and Association involved in A β 42 Peptide Aggregation in Alzheimer's Disease. *BMC Bioinformatics* 2010, **11**(Suppl 6):S24.
68. Manda P, *et al*: GOModeler- A tool for hypothesis-testing of functional genomics datasets. *BMC Bioinformatics* 2010, **11**(Suppl 6):S29.
69. Zhou A, Zhang F, Chen J: PEPPI: A Peptidomic Database of Human Protein Isoforms for Proteomics Experiments. *BMC Bioinformatics* 2010, **11**(Suppl 6):S7.

doi:10.1186/1471-2105-11-S6-S1

Cite this article as: Wren *et al.*: Proceedings of the 2010 MidSouth Computational Biology and Bioinformatics Society (MCBIOS) Conference. *BMC Bioinformatics* 2010 **11**(Suppl 6):S1.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

