

Introduction

## Proceedings of the 2008 MidSouth Computational Biology and Bioinformatics Society (MCBIOS) Conference

Jonathan D Wren\*<sup>1</sup>, Dawn Wilkins<sup>2</sup>, James C Fuscoe<sup>3</sup>, Susan Bridges<sup>4</sup>, Stephen Winters-Hilt<sup>5</sup> and Yuriy Gusev<sup>6</sup>

Address: <sup>1</sup>Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation; 825 N.E. 13th Street, Oklahoma City, OK 73104-5005, USA, <sup>2</sup>Computer & Information Science Department, The University of Mississippi, University, MS 38677, USA, <sup>3</sup>Center for Functional Genomics, Division of Systems Toxicology; National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, AR 72079, USA, <sup>4</sup>Department of Computer Science and Engineering, Mississippi State University, Box 9637, Mississippi State, MS 39762, USA, <sup>5</sup>Department of Computer Science, University of New Orleans, and The Research Institute for Children, 200 Henry Clay Ave., New Orleans, LA 70118, USA and <sup>6</sup>Department of Surgery, Health Sciences Center, The University of Oklahoma, Oklahoma City, OK 73104, USA

Email: Jonathan D Wren\* - Jonathan.Wren@OMRF.org; Dawn Wilkins - dwilkins@cs.olemiss.edu; James C Fuscoe - james.fuscoe@fda.hhs.gov; Susan Bridges - bridges@cse.msstate.edu; Stephen Winters-Hilt - winters@cs.uno.edu; Yuriy Gusev - Yuriy-Gusev@ouhsc.edu

\* Corresponding author

from Fifth Annual MCBIOS Conference. Systems Biology: Bridging the Omics  
Oklahoma City, OK, USA. 23–24 February 2008

Published: 12 August 2008

BMC Bioinformatics 2008, 9(Suppl 9):S1 doi:10.1186/1471-2105-9-S9-S1

This article is available from: <http://www.biomedcentral.com/1471-2105/9/S9/S1>

© 2008 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 2008 was held February 23–24, 2008 in Oklahoma City, Oklahoma at the Cox Convention Center in Bricktown. It was the best attended in the series of MCBIOS conferences (140 registrants) with the most participation (68 posters submitted). Informative and engaging keynote talks were delivered by Dr. Bruce Roe and Dr. Edward Dougherty. The full agenda is online at <http://www.okbios.org>.

Student poster award winners were: Vinay Ravindrakumar of University of Arkansas for Medical Sciences (1st place), Quan Shi of Little Rock Central High School (2nd) and Brian Roux of the University of New Orleans (UNO) (3rd), with honorary mentions going to Murat Eren of UNO and Prashanti Manda of Mississippi State University (MSU). Student talk winners were: Daniel Quest of the University of Nebraska Medical Center (1st place), Nan Wang of MSU (2nd), and William Sanders of MSU (3rd).

### Proceedings summary

This year, 19 out of 27 submitted papers were accepted for inclusion in the official conference proceedings (70%), similar to the number published from MCBIOS 2007 [1–26]. Each paper was peer-reviewed by at least two reviewers. Our goal in peer-review for the Proceedings is to be inclusive enough to accurately reflect the scope of scientific work presented at the conference yet rigorous enough such that only the highest quality work presented is selected for inclusion in the official proceedings. The general themes of this year's proceedings papers fall into five categories, discussed below.

### Systems biology

Biological systems can be modeled as complex systems with many interactions between the components. One goal of emerging systems biology is to analyze very large complex biological networks such as protein-protein interactions, metabolism, and regulation to identify func-

tional modules and to assign the functions to certain components of the system. Mutlu Mete *et al.* [27] devised a new methodology called SCAN (Structural Clustering Algorithm for Networks) that can efficiently find clusters or functional modules in complex biological networks, as well as hubs and outliers. In addition, nodes can be classified into various roles based on their structures. Interpretations of functional groups found by SCAN showed superior performance over CNM, a well-known modularity-based clustering algorithm.

Analysis of microarray gene expression data is challenging and may lead to biased or incomplete biological interpretations. To gain a more holistic (i.e., systemic) picture, it is essential to integrate a careful statistical approach with biological knowledge from various sources into the analysis. Mikhail Dozmorov *et al.* [28] present an integrative approach to microarray analysis and demonstrate how the various steps in their process support each other and refine the current model of cell-matrix interaction. With their method, they were able to identify inflammation and G-protein signaling as processes affected by the extracellular matrix.

Metastases are responsible for the majority of cancer fatalities. The molecular mechanisms governing metastasis are poorly understood, hindering early diagnosis and treatment. Unlike most previous studies, a study by Andrey Ptitsyn *et al.* [29] proposes an approach that puts into focus gene interaction networks and molecular pathways rather than separate marker genes. This study indicates that regardless of the tissue of origin, all metastatic tumors share a number of common features related to changes in basic energy metabolism, cell adhesion/cytoskeleton remodeling, antigen presentation and cell cycle regulation.

Circadian rhythm is a crucial factor in orchestration of plant physiology, keeping it in synchrony with the day-light cycle. Previous studies reported approximately 16% of plant genes behaved in a circadian fashion, while studies in mammals suggested circadian baseline oscillation in nearly 100% of genes. Andrey Ptitsyn [30] presents a comprehensive analysis of periodicity in two independent *Arabidopsis thaliana* data sets. This study indicates a more pervasive role of gene expression oscillation in the molecular physiology of plants than previously believed. Application of advanced algorithms identified circadian baseline oscillation in almost all plant genes as well as a complex orchestration of gene expression timing in important biological pathways.

## OMICS

Chromatography coupled to mass spectrometry is a powerful way to resolve and compare the relative abundance

of chemical compounds within heterogeneous biological samples. However the resulting data sets are 2 or 3-dimensional, presenting formidable obstacles to peak alignment – a process required to ensure sample comparison is conducted appropriately. The first dimension of separation is chromatographic elution time, which varies from run to run for each molecular species. To solve this problem, Minh Chae *et al.* [31] developed an iterative block-shifting approach that adjusts for variation in retention time without distorting peak area. They first matched chemically identical peaks based on both retention-time and mass-spectral information. Non-peak regions of each chromatogram were stretched or compressed to align peaks with a reference chromatogram, thus preserving the shapes of matched peaks. Their approach compared favorably to other approaches, and was superior in preservation of peak area.

Also, in the proceedings, Tianxiao Huan *et al.* describe Proteolens, a new tool to navigate and visualize biological networks [32].

## Microarray studies

Microarrays are a powerful technology and an area of active research interest in bioinformatics, with a focus on the development of novel methods for analysis and interpretation of experiments [33-49]. This year's proceedings reflect this area of active research interest with several reports that focus on the development of methods and analysis of microarray data.

Microarray-based molecular signatures have played an increasing role in diagnosis, prognosis and risk/safety assessments, the first step of which is to identify a set of informative genes. Zhenqiang Su *et al.* [50] investigate a new gene selection approach to identify informative genes. The rationale of the approach is that informative genes should consistently be significantly differentially expressed for different variations of sample size. Genes exhibiting significance throughout the iterations are considered a Very Important Pool (VIP) of genes. It was found that the genes identified by the VIP method, but not by the p-value ranking approach, are also related to the disease investigated, and these genes are part of the pathways derived from the common genes shared by both the VIP and p-ranking methods. Moreover, the binary classifiers built from these genes are statistically equivalent to those built from the top 50 p-value ranked genes in distinguishing different types of samples. Therefore, the VIP gene selection approach could identify additional subsets of informative genes that would not always be selected by the p-value ranking method.

The paper by Taewon Lee *et al.* [51] presents a method to test the significance of expression changes within a group

of genes, while considering the correlation structure among genes in each group. This method enables the rapid detection of gene expression changes, indicating altered cell functions or pathways, and facilitates the interpretation of the data. Application of the method to real data shows that it is an improved, practical method to evaluate the effects of treatments on functional classes of genes, such as those based on Gene Ontology descriptors.

Also in the proceedings, Arun Rawat *et al.* report on a method of microarray graph mining to derive co-expressed genes [52], and Leming Shi *et al.* report on an impressively large study of the reproducibility of gene lists for microarray experiments, and conclude with recommendations for detecting significant differential expression [53].

### Genomic analysis

As more and more genomes become fully sequenced in the coming years, gene identification is still a limiting factor to scientific discovery. Since a significant proportion of genes exist as members of families of genes with related functions, Ronald Frank *et al.* [54] have employed a strategy to identify these gene family members using patterns indicating negative selection pressure on the coding region. The authors tested the strategy on several well-characterized gene families from *Arabidopsis thaliana* and report their success in correctly identifying several members of each gene family starting with one known member and using only EST data.

Highly accurate and reproducible genotype calling are paramount for genome-wide association studies (GWAS), since errors introduced by calling algorithms can lead to inflation of false associations between genotype and phenotype. Most genotype calling algorithms currently used for GWAS are based on multiple arrays, consisting of many samples. Huixiao Hong *et al.* [55] observed that batch size and composition affect the genotype calling results in GWAS using the algorithm BRLMM. The larger the differences in batch sizes, the larger the effect. The more homogenous the samples in the batches, the more consistent the genotype calls. The inconsistency propagates to the lists of significantly associated single nucleotide polymorphisms identified in downstream association analysis. Thus, uniform and large batch sizes should be used to make genotype calls for GWAS. In addition, samples of high homogeneity should be placed into the same batch.

The cellular machinery by which genes are expressed is both complex and an active area of recent bioinformatics research [56-66]. A first step in understanding this process is to locate the binding positions of transcription factors over the chromosome. Since the search space is large,

advanced computational tools play a central role in solving this problem. Despite the development of nearly two hundred tools to elucidate transcription factor binding sites, much controversy still remains on how to build methods with high sensitivity and specificity. Central in this debate is determining the factors that will improve the quality of computational predictions. The paper by Daniel Quest *et al.* [67], presents a novel benchmarking strategy to automate and evaluate methods designed to detect transcription factor binding sites. The strategy allows researchers, for the first time, to evaluate transcription factor detection methods on the genome scale. In particular, researchers can vary the data, algorithms, parameters and transcription factor binding site representations to determine the method best suited to their problem of interest. The proposed platform allows for rapid evaluation of deficits in current models and paves the way to develop new tools to overcome these problems.

Also, the Garner Lab extends their work on predicting the impact of single nucleotide polymorphisms (SNPs) in a paper by Vinayak Kulkarni *et al.* [68], and Jerzy Zielinski *et al.* report on a method of analyzing genomic sequences by a time-dependent autoregressive moving average [69].

### Miscellaneous

Text-mining is an area of bioinformatics whereby identification and analysis of trends in text is done computationally [70-78]. To this end, Cory Giles and Jonathan Wren developed a method of identifying directional relationships within text (e.g., chemical X increases heart rate, or gene Y elevates inflammation) using natural language processing (NLP) [79]. Their goals were also to make their system scalable to large bodies of text (e.g. MEDLINE has 18 million records and counting), as well as understanding how much apparent contradiction takes place when attempting to extract isolated facts from within a greater context from these huge bodies of text.

Christopher Bottoms and Dong Xu study atom-naming conventions in the Protein Data Bank and find that some names are assigned ad hoc, resulting in duplicate names and creating problems for standardization and data-mining [80].

In [81], Roux and Winters-Hilt describe Hybrid SVM/HMM structural sensors for use in analysis of stochastic sequential data. They begin with a novel approach to classification using Support Vector Machines and Markov Models with application to detecting Intron-Exon and Exon-Intron (5' and 3') splice sites. The approach also includes the application of Shannon Entropy based analysis of the stochastic datasets to detect minimal data components for feature extraction. Results are presented for a variety of eukaryotic species.

In the Winters-Hilt group, work continues on developing nanopore detector signal analysis via machine learning methods for classification and knowledge discovery. In [82], Churbanov and Winters-Hilt describe the application of a distributed Mixture of Hidden Markov Models (MHMMs) to the problem of channel current blockade clustering and associated analyte classification. The distributed MHMM provides a feature extraction that is equivalent to that of the sequential HMM with a speedup factor approximately equal to the number of independent CPUs operating on the data.

### Future meetings

The Sixth annual MCBIOS Conference will be held in Starkville, Mississippi in early spring, 2009. See <http://www.MCBIOS.org> for further information on MCBIOS and future meetings. MCBIOS and OKBIOS are both regional affiliates 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. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Food and Drug Administration.

### Acknowledgements

We thank the Conference Committee, Program Committee, student volunteers and sponsors for their help in organizing MCBIOS 2008. We would like to especially thank Jim Mason and Frank Waxman for their sponsorship of MCBIOS 2008 as well as OKBIOS 2004 and OKBIOS 2005. We also thank our peer-reviewers for their quality efforts to review submitted manuscripts.

This article has been published as part of *BMC Bioinformatics* Volume 9 Supplement 9, 2008: Proceedings of the Fifth Annual MCBIOS Conference. Systems Biology: Bridging the Omics. The full contents of the supplement are available online at <http://www.biomedcentral.com/1471-2105/9?issue=S9>

### References

1. **Computational frontiers in biomedicine. Proceedings of the Fourth Annual Conference of the MidSouth Computational Biology and Bioinformatics Society. February 1-3, 2007. New Orleans, Louisiana, USA.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S1-25.
2. Winters-Hilt S: **The alpha-hemolysin nanopore transduction detector – single-molecule binding studies and immunological screening of antibodies and aptamers.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S9.
3. Ding Y, Dang X, Peng H, Wilkins D: **Robust clustering in high dimensional data using statistical depths.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S8.
4. Pirooznia M, Gong P, Guan X, Inouye LS, Yang K, Perkins EJ, Deng Y: **Cloning, analysis and functional annotation of expressed sequence tags from the Earthworm *Eisenia fetida*.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S7.
5. Yuan JS, Burris J, Stewart NR, Mentewab A, Stewart CN Jr: **Statistical tools for transgene copy number estimation based on real-time PCR.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S6.
6. Nagarajan V, Elasri MO: **Structure and function predictions of the Msa protein in Staphylococcus aureus.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S5.
7. Mei N, Guo L, Liu R, Fuscoe JC, Chen T: **Gene expression changes induced by the tumorigenic pyrrolizidine alkaloid riddelliine in liver of Big Blue rats.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S4.
8. Schnackenberg LK, Sun J, Espandiani P, Holland RD, Hanig J, Beger RD: **Metabonomics evaluations of age-related changes in the urinary compositions of male Sprague Dawley rats and effects of data normalization methods on statistical and quantitative analysis.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S3.
9. Loganantharaj R, Atwi M: **Towards validating the hypothesis of phylogenetic profiling.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S25.
10. Bridges SM, Magee GB, Wang N, Williams WP, Burgess SC, Nanduri B: **ProtQuant: a tool for the label-free quantification of MudPIT proteomics data.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S24.
11. Sanders WS, Bridges SM, McCarthy FM, Nanduri B, Burgess SC: **Prediction of peptides observable by mass spectrometry applied at the experimental set level.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S23.
12. Guo L, Mei N, Dial S, Fuscoe J, Chen T: **Comparison of gene expression profiles altered by comfrey and riddelliine in rat liver.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S22.
13. Das MK, Dai HK: **A survey of DNA motif finding algorithms.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S21.
14. Winters-Hilt S, Morales E, Amin I, Stoyanov A: **Nanopore-based kinetics analysis of individual antibody-channel and antibody-antigen interactions.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S20.
15. Dozmorov MG, Kyker KD, Saban R, Shankar N, Baghdayan AS, Centola MB, Hurst RE: **Systems biology approach for mapping the response of human urothelial cells to infection by Enterococcus faecalis.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S2.
16. Winters-Hilt S, Baribault C: **A novel, fast, HMM-with-Duration implementation – for application with a new, pattern recognition informed, nanopore detector.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S19.
17. Winters-Hilt S, Merat S: **SVM clustering.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S18.
18. Mete M, Xu X, Fan CY, Shafirstein G: **Automatic delineation of malignancy in histopathological head and neck slides.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S17.
19. Gusev Y, Schmittgen TD, Lerner M, Postier R, Brackett D: **Computational analysis of biological functions and pathways collectively targeted by co-expressed microRNAs in cancer.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S16.
20. Ptitsyn AA, Gimble JM: **Analysis of circadian pattern reveals tissue-specific alternative transcription in leptin signaling pathway.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S15.
21. Churbanov A, Baribault C, Winters-Hilt S: **Duration learning for analysis of nanopore ionic current blockades.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S14.
22. Li P, Zhang C, Perkins EJ, Gong P, Deng Y: **Comparison of probabilistic Boolean network and dynamic Bayesian network approaches for inferring gene regulatory networks.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S13.
23. Landry M, Winters-Hilt S: **Analysis of nanopore detector measurements using Machine-Learning methods, with application to single-molecule kinetic analysis.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S12.
24. Thomson K, Amin I, Morales E, Winters-Hilt S: **Preliminary nanopore cheminformatics analysis of aptamer-target binding strength.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S11.
25. Winters-Hilt S, Davis A, Amin I, Morales E: **Nanopore current transduction analysis of protein binding to non-terminal and terminal DNA regions: analysis of transcription factor binding, retroviral DNA terminus dynamics, and retroviral integrase-DNA binding.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S10.
26. Wilkins D, Gusev Y, Loganantharaj R, Bridges S, Winters-Hilt S, Wren JD: **Proceedings of the Fourth Annual Conference of the Mid-South Computational Biology and Bioinformatics Society.** *BMC Bioinformatics* 2007, **8(Suppl 7)**:S1.

27. Mete M, Tang F, Xu X, Yuruk N: **A Structural Approach for Finding Functional Modules from Large Biological Networks.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S19.
28. Dozmorov MG, Kyker KD, Hauser PJ, Saban R, Buethe DD, Dozmorov I, Centola MB, Culkun DJ, 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.
29. Ptitsyn AA, Weil MM, Thamm DH: **Systems biology approach to identification of biomarkers for metastatic progression in cancer.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S8.
30. Ptitsyn AA: **Comprehensive analysis of circadian periodic pattern in plant transcriptome.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S18.
31. Chae M, Shmookler Reis RJ, Thaden JJ: **An iterative block-shifting approach to retention time alignment that preserves the shape and area of gas chromatography-mass spectrometry peaks.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S15.
32. Huan T, Sivachenko AY, Harrison SH, Chen JY: **ProteoLens: a Visual Analytic Tool for Multi-scale Database-driven Biological Network Data Mining.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S5.
33. Bovelstad HM, Nygard S, Storvold HL, Aldrin M, Borgano O, Frigessi A, Lingjaerde OC: **Predicting survival from microarray data—a comparative study.** *Bioinformatics* 2007, **23(16)**:2080-2087.
34. Burkart MF, Wren JD, Herschkowitz JI, Perou CM, Garner HR: **Clustering microarray-derived gene lists through implicit literature relationships.** *Bioinformatics* 2007, **23(15)**:1995-2003.
35. Chen X, Hughes TR, Morris Q: **RankMotif+: a motif-search algorithm that accounts for relative ranks of K-mers in binding transcription factors.** *Bioinformatics* 2007, **23(13)**:i72-79.
36. Dyer MD, Murali TM, Sobral BV: **Computational prediction of host-pathogen protein-protein interactions.** *Bioinformatics* 2007, **23(13)**:i159-166.
37. Elo LL, Jarvenpaa H, Oresic M, Laheesmaa R, Aittokallio T: **Systematic construction of gene coexpression networks with applications to human T helper cell differentiation process.** *Bioinformatics* 2007, **23(16)**:2096-2103.
38. Fishel I, Kaufman A, Ruppin E: **Meta-analysis of gene expression data: a predictor-based approach.** *Bioinformatics* 2007, **23(13)**:i599-1606.
39. Fujita A, Sato JR, Garay-Malpartida HM, Morettin PA, Sogayar MC, Ferreira CE: **Time-varying modeling of gene expression regulatory networks using the wavelet dynamic vector autoregressive method.** *Bioinformatics* 2007, **23(13)**:1623-1630.
40. Gao S, Wang X: **TAPPA: topological analysis of pathway phenotype association.** *Bioinformatics* 2007, **23(22)**:3100-3102.
41. Gyenesei A, Wagner U, Barkow-Oesterreicher S, Stolte E, Schlapbach R: **Mining co-regulated gene profiles for the detection of functional associations in gene expression data.** *Bioinformatics* 2007, **23(15)**:1927-1935.
42. Hibbs MA, Hess DC, Myers CL, Huttenhower C, Li K, Troyanskaya OG: **Exploring the functional landscape of gene expression: directed search of large microarray compendia.** *Bioinformatics* 2007, **23(20)**:2692-2699.
43. Lim WK, Wang K, Lefebvre C, Califano A: **Comparative analysis of microarray normalization procedures: effects on reverse engineering gene networks.** *Bioinformatics* 2007, **23(13)**:i282-288.
44. Martin S, Zhang Z, Martino A, Faulon JL: **Boolean dynamics of genetic regulatory networks inferred from microarray time series data.** *Bioinformatics* 2007, **23(7)**:866-874.
45. Medina I, Montaner D, Tarraga J, Dopazo J: **Prophet, a web-based tool for class prediction using microarray data.** *Bioinformatics* 2007, **23(3)**:390-391.
46. Mukhopadhyay ND, Chatterjee S: **Causality and pathway search in microarray time series experiment.** *Bioinformatics* 2007, **23(4)**:442-449.
47. Schreiber AW, Baumann U: **A framework for gene expression analysis.** *Bioinformatics* 2007, **23(2)**:191-197.
48. Sean D, Meltzer PS: **GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor.** *Bioinformatics* 2007, **23(14)**:1846-1847.
49. Yu Z, Wong HS, Wang H: **Graph-based consensus clustering for class discovery from gene expression data.** *Bioinformatics* 2007, **23(21)**:2888-2896.
50. Su Z, Hong H, Fang H, Shi L, Perkins R, Tong W: **Very Important Pool (VIP) Genes – An Application for Microarray-Based Molecular Signatures.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S9.
51. Lee T, Desai VG, Velasco C, Shmookler Reis RJ, Delongchamp RR: **Testing for treatment effects on gene ontology.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S20.
52. Rawat A, Deng Y: **Novel implementation of conditional co-regulation by graph theory to derive co-expressed genes from microarray data.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S7.
53. Shi L, Jones WD, Jensen RV, Harris SC, Perkins RG, Goodsaid FM, Guo L, Croner LJ, Boysen C, Fang H, et al.: **The Balance of Reproducibility, Sensitivity, and Specificity of Lists of Differentially Expressed Genes in Microarray Studies.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S10.
54. Frank RL, Kandoth C, Ercal F: **Validation of an NSP-based (negative selection pattern) gene family identification strategy.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S2.
55. Hong H, Su Z, Ge W, Shi L, Perkins R, Fang H, Xu J, Chen JJ, Han T, Kaput J, et al.: **Assessing Batch Effects of Genotype Calling Algorithm BRLMM for the Affymetrix GeneChip Human Mapping 500 K Array Set Using 270 HapMap Samples.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S17.
56. Gamberoni G, Lamma E, Lodo G, Marchesini J, Mascellani N, Rossi S, Storari S, Tagliavini L, Volinia S: **Fun&Co: identification of key functional differences in transcriptomes.** *Bioinformatics* 2007, **23(20)**:2725-2732.
57. Hertzberg L, Izraeli S, Domany E: **STOP: searching for transcription factor motifs using gene expression.** *Bioinformatics* 2007, **23(14)**:1737-1743.
58. Hijikata A, Kitamura H, Kimura Y, Yokoyama R, Aiba Y, Bao Y, Fujita S, Hase K, Hori S, Ishii Y, et al.: **Construction of an open-access database that integrates cross-reference information from the transcriptome and proteome of immune cells.** *Bioinformatics* 2007, **23(21)**:2934-2941.
59. Meng H, Banerjee A, Zhou L: **BLISS 2.0: a web-based tool for predicting conserved regulatory modules in distantly-related orthologous sequences.** *Bioinformatics* 2007, **23(23)**:3249-3250.
60. Wang RS, Wang Y, Zhang XS, Chen L: **Inferring transcriptional regulatory networks from high-throughput data.** *Bioinformatics* 2007, **23(22)**:3056-3064.
61. Yuan S, Li KC: **Context-dependent clustering for dynamic cellular state modeling of microarray gene expression.** *Bioinformatics* 2007, **23(22)**:3039-3047.
62. Zhu QH, Guo AY, Gao G, Zhong YF, Xu M, Huang M, Luo J: **DPTF: a database of poplar transcription factors.** *Bioinformatics* 2007, **23(10)**:1307-1308.
63. Montgomery SB, Griffith OL, Sleumer MC, Bergman CM, Bilenky M, Pleasance ED, Prychyna Y, Zhang X, Jones SJ: **ORegAnno: an open access database and curation system for literature-derived promoters, transcription factor binding sites and regulatory variation.** *Bioinformatics* 2006, **22(5)**:637-640.
64. Narlikar L, Gordan R, Ohler U, Hartemink AJ: **Informative priors based on transcription factor structural class improve de novo motif discovery.** *Bioinformatics* 2006, **22(14)**:e384-392.
65. Shipra A, Chetan K, Rao MR: **CREMOFAC—a database of chromatin remodeling factors.** *Bioinformatics* 2006, **22(23)**:2940-2944.
66. Sonnenburg S, Zien A, Ratsch G: **ARTS: accurate recognition of transcription starts in human.** *Bioinformatics* 2006, **22(14)**:e472-480.
67. Quest D, Dempsey K, Shafiqullah M, Bastola D, Ali H: **MTAP: The Motif Tool Assessment Platform.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S6.
68. Kulkarni V, Errami M, Barber R, Garner HR: **Exhaustive prediction of disease susceptibility to coding base changes in the human genome.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S3.
69. Zielinski JS, Bouaynaya N, Schonfeld D, O'Neil WV: **Time-dependent ARMA modeling of genomic sequences.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S14.
70. Minguez P, Al-Shahrour F, Montaner D, Dopazo J: **Functional profiling of microarray experiments using text-mining derived bioentities.** *Bioinformatics* 2007, **23(22)**:3098-3099.
71. Torvik VI, Smalheiser NR: **A quantitative model for linking two disparate sets of articles in MEDLINE.** *Bioinformatics* 2007, **23(13)**:1658-1665.

72. Xu H, Fan JW, Hripcsak G, Mendonca EA, Markatou M, Friedman C: **Gene symbol disambiguation using knowledge-based profiles.** *Bioinformatics* 2007, **23(8)**:1015-1022.
73. Chen ES, Hripcsak G, Xu H, Markatou M, Friedman C: **Automated acquisition of disease drug knowledge from biomedical and clinical documents: an initial study.** *J Am Med Inform Assoc* 2008, **15(1)**:87-98.
74. Lin Y, Li W, Chen K, Liu Y: **A document clustering and ranking system for exploring MEDLINE citations.** *J Am Med Inform Assoc* 2007, **14(5)**:651-661.
75. Liu H, Hu ZZ, Torii M, Wu C, Friedman C: **Quantitative assessment of dictionary-based protein named entity tagging.** *J Am Med Inform Assoc* 2006, **13(5)**:497-507.
76. Fan JW, Friedman C: **Semantic classification of biomedical concepts using distributional similarity.** *J Am Med Inform Assoc* 2007, **14(4)**:467-477.
77. Huang Y, Lowe HJ: **A novel hybrid approach to automated negation detection in clinical radiology reports.** *J Am Med Inform Assoc* 2007, **14(3)**:304-311.
78. Uzuner O, Luo Y, Szolovits P: **Evaluating the state-of-the-art in automatic de-identification.** *J Am Med Inform Assoc* 2007, **14(5)**:550-563.
79. Giles CB, Wren JD: **Large-scale Directional Relationship Extraction and Resolution.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S11.
80. Bottoms CA, 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.
81. Roux B, Merat S, Winters-Hilt S: **Hybrid SVM/HMM Structural Sensors for Stochastic Sequential Data.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S12.
82. Churbanov A, Winters-Hilt S: **Clustering ionic flow blockade toggles with a Mixture of HMMs.** *BMC Bioinformatics* 2008, **9(Suppl 9)**:S13.

Publish with **BioMed Central** and every scientist can read your work free of charge

*"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."*

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

