Introduction

Selected proceedings of the 2009 Summit on Translational Bioinformatics

Yves A Lussier*[1] and Indra Neil Sarkar*[2]

Address: [1]Center for Biomedical Informatics and Section of Genetic Medicine of the Department of Medicine, Institute for Translational Medicine, Institute for Genomics and Systems Biology, Computational Institute, and Ludwig Center for Metastasis Research; Cancer Research Center, University of Chicago, 5841 South Maryland Ave, Room N-660B, Mail Code 6091, Chicago, IL 60637, USA and [2]Center for Clinical and Translational Science, University of Vermont, Burlington, VT 05401, USA

E-mail: Yves A Lussier* - lussier@uchicago.edu; Indra Neil Sarkar* - neil.sarkar@uvm.edu
*Corresponding author

This article is available from: http://www.biomedcentral.com/1471-2105/10/S9/I1

Background

"It is the responsibility of those of us involved in today's biomedical research enterprise to translate the remarkable scientific innovations we are witnessing into health gains for the nation [1]." The First Summit on Translational Bioinformatics (held in 2008) was organized in response to the call for translational bioinformaticians that was issued by the Director of the NIH, at a time there was no national annual conference or symposium for the presentation and discussion of research work in Translational Bioinformatics.

This Second Summit built on the success of the first Summit on Translational Bioinformatics, highlighting the multidisciplinary nature of this rapidly maturing research field and providing the opportunity to forge new transdisciplinary collaborations as the finest minds of the academia, industry, government and non-profit sector are brought together. Reflecting on the unique position of translational bioinformatics in medical research and healthcare management, the 2009 Scientific program Committee designed a comprehensive program comprised of the four original tracks (#1–4) and with two new ones (#5,6):

1 – Informatics methods for the analysis of molecular and clinical measurements,
2 – Computational approaches to finding mechanisms and therapies for diseases,
3 – Relating and representing phenotypes and diseases,
4 – Dissecting disease through the study of organisms, evolution and taxonomy,
5 – Informatics concepts, tools, and techniques to enable integrative translational research,
6 – Informatics methods in genetics discoveries and clinical practice.

Workshop program

Three hundred and fifty attendees, from 201 distinct institutions or companies covering 36 states in the US and 17 other countries, met over three days in San Francisco, CA. Sixty-three papers and thirty-seven abstracts were submitted for oral presentation to the Summit, of which 39 and 28 were respectively accepted for oral presentation by the Scientific Program Committee (Additional file 1). In addition to these papers, 3 tutorials, 64 posters and 8 panels were presented.

The opening session keynote was by Dr. Lynn H. Vogel, Vice President and Chief Information Officer at The University of Texas M.D. Anderson Cancer Center in Houston. Dr. Vogel presented the concepts and investments underway at the M.D. Anderson Cancer Center for "Building an IT Architecture to Facilitate Personalized Medicine". The second keynote was Dr. Andrea Califano, Professor of Biomedical Informatics at Columbia University, where he is currently Director of the Center for the Multiscale Analysis of Genetic Networks (MAGNet – one of seven NIH-funded National Centers for Biomedical Computing), Associate Director for Bioinformatics of the Herbert Irving Comprehensive Cancer Center (HICCC), and co-Director of the Center for Computational Biology and Bioinformatics (C2B2). He presented on "Using Cancer Interactome for therapeutic target and associated biomarker discovery". The closing session keynote presentation was delivered by Dr. Russ B. Altman, Professor and Chair of Bioengineering, Professor of Medicine, Genetics, and Computer Science (by courtesy), Stanford University, and Principal Investigator of Simbios, one of the National Centers for Biomedical Computing. His presentation highlighted a year-in-review of the literature in Translational Bioinformatics.

Slides from most tutorials, papers, and panels are publicly available at http://summit2009.amia.org/presentation-slides

Summary of the selected contributions

The seventeen papers selected for this supplement to *BMC Bioinformatics* are extended and improved versions of the best papers accepted to the 2009 Summit of Translational Bioinformatics. In the following paragraphs, we briefly describe them.

Six of these papers deal with the science of network analyses applied to the Translational Bioinformatics field. Chang and Ramoni show how Bayesian Networks (BN) can improve on current microarray analyses by generating a predictor that does not require the biased adjustment of parameters of the current state if the art. An unbiased BN predictor of aortic aneurysm diagnosis is provided as proof of concept [2]. Lee *et al.* show how protein interactions and gene expression can be combined for predicting novel joint networks ignored by array analyses alone[3]. Bhavnani *et al.* provide a molecular classification of renal diseases showing gene expression similarities rather than morphological similarities, perhaps illustrative of shared expression underpinning organ-specific phenotypic networks [4]. Butte's group illustrates how non-trivial correlations between two physiological measurements in intensive care unit can be more reliable for prognosis than single

measurements or laboratory values suggesting an unbiased and systematic approach to discovery of novel complex prognosis markers [5]. Dean *et al.* demonstrated how evolutionary conserved binding sites of transcriptional factors used jointly with expression profiling can be used to subsequently construct cis-regulatory expression modules in osteocytes [6]. Yang and Lussier propose a novel and robust algorithm to identify epigenetic regulation derived from co-expression networks and epigenetic targets in the literature, with a specific evaluation in terms of significance and accuracy in the context of Acute Lymphobalstic Leukemia [7].

Two papers are in the area of decision making relying on genomic or complex multiscale analyses. The PAPAyA framework integrates the complexity of multiscale analyses inclusive of clinical prognosis, histopathological morphology and genomic measurements to facilitate multiscale biomarker development, hypothesis generation and decision support [8]. Overby provides a framework for question answering with genomic and clinical knowledge from the Unified Medical Language System [9].

Three papers were highlighted in the area of large-scale repositories in support of genomic medicine. Ohno-Machado's group create a comprehensive and consistent annotation system for the context in which samples are annotated in the NIH Gene Expression Omnibus database (GEO), inclusive of tissue, cell line, histopathological type, and subject characteristics such as demographics, treatment, or survival. They report an implementation of over 6,000 samples extracted from GEO [10] and inter-rater agreement of 92% [11]. Tarczy-Hornoch's group created SNPit, a database that provides some guidance on the variability of SNPs as an indicator of an adequate measure for clinical relevance [12].

Three papers focused on phenotypic representation, their ontologies and their coding in narratives. South *et al.* provided a practical approach to text processing of phenotypic information in clinical records [13], while Wang, Hripcsak and Friedman provide an algorithm based on mutual information to disclose interdependencies that may not be linear between environmental (medication) and phenotypic data [14]. Musen's group provides a novel algorithm for term-based ontology annotation that performs better than the NIH gold standard MetaMap[15].

Finally, three papers were chosen in the unrelated areas of fMRI, structural biology and rule learning. Jin *et al.* demonstrated a feature-learning algorithm to identify patterns of voxels in fMRI associate to phenotypic classification [16]. Lustgard used an ontology-anchored

set of known biomarkers in order to improve machine learning algorithms designed to discover novel ones in a proteomic datasets [17]. Shivakumar and Krauthammer use structural similarity to identify substitute sensitivity profiles for medications that are not included in cancer screens, aiming to potentially increase the relevance of existing anti-cancer drug screens [18].

The next Summit on Translational Bioinformatics will be held in 2010 in San Francisco.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Both authors wrote and approved the manuscript. Lussier was the Chair of Scientific Program Committee for the 2009 AMIA Summit on Translational Bioinformatics; Sarkar was the Vice-Chair of the Scientific Program Committee and the Chair of the Poster Program Committee for the 2009 AMIA Summit on Translational Bioinformatics.

Acknowledgements

References

1. Zerhouni EA: **Translational and clinical science – time for a new vision.** *N Engl J Med* 2005, **353(15):**1621–1623, Epub 2005 Oct 12.
2. Chang H and Ramoni MF: **Transcriptional network classifiers.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S1.
3. Lee E, Jung H, Radivojac P, Kim J and Lee D: **Analysis of AML genes in dysregulated molecular networks.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S2.
4. Bhavnani SK, Eichinger F, Martini S, Saxman P, Jagadish HV and Kretzler M: **Network analysis of genes regulated in renal diseases: implications for a molecular-based classification.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S3.
5. Grossman A, Cohen MJ, Manley GT and Butte AJ: **Infection in the intensive care unit alters physiological networks.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S4.
6. Dean AK, Harris SE, Kalajzic I and Ruan J: **A systems biology approach to the identification and analysis of transcriptional regulatory networks in osteocytes.** *BMC Bioinformatics* 2009, **10 (Suppl 9):**S5.
7. Yang X, Huang Y, Chen JL, Xie J, Sun X and Lussier YA: **PGnet: Mechanism-anchored profiling from epigenetic networks predicts outcome in acute lymphoblastic leukemia.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S6.
8. Janevski A, Kamalakaran S, Banerjee N, Varadan V and Dimitrova N: **PAPAyA: A Platform for breast cancer biomarker signature discovery and evaluation and assessment.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S7.
9. Overby C, Tarczy-Hornoch P and Demner-Fushman D: **The potential for automated question answering in the context of genomic medicine: an assessment of existing resources and properties of answers.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S8.
10. Pitzer E, Lacson R, Hinske C, Kim J, Galante PAF and Ohno-Machado L: **Towards large-Scale sample annotation in gene expression repositories.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S9.
11. Lacson R, Pitzer E, Hinske C, Galante P and Ohno-Machado L: **Evaluation of a Large-Scale Biomedical Data Annotation Initiative.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S10.
12. Shen TH, Carlson CS and Tarczy-Hornoch P: **Evaluating the accuracy of a functional SNP annotation system.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S11.
13. South BR, Shen S, Jones M, Garvin J, Samore M, Chapman W and Gundlapalli A: **Developing a manually annotated clinical document corpus to identify phenotypic information for inflammatory bowel disease.** *BMC Bioinformatics* 2009, **10 (Suppl 9):**S12.
14. Wang X, Hripcsak G and Friedman C: **Characterizing environmental and phenotypic associations using information theory and electronic health records.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S13.
15. Shah NH, Bhatia N, Jonquet C, Rubin D, Chiang AP and Musen MA: **Comparison of concept recognizers for building the open biomedical annotator.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S14.
16. Jin B, Strasburger A, Laken SJ, Kozel FA, Johnson KA, George MS and Lu X: **Feature selection for fMRI-based deception detection.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S15.
17. Lustgarten JL, Visweswaran S, Bowser RP, Hogan WR and Gopalakrishnan V: **Knowledge-based variable selection for learning rules from proteomic data.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S16.
18. Shivakumar P and Krauthammer M: **Structural similarity assessment for drug sensitivity prediction in cancer.** *BMC Bioinformatics* 2009, **10(Suppl 9):**S17.