Editorial

Zhongming Zhao

Departments of Biomedical Informatics,
Psychiatry, and Cancer Biology,
Vanderbilt University Medical Center,
Nashville, Tennessee 37232, USA
E-mail: zhongming.zhao@vanderbilt.edu

Antonio Sanfilippo

Computational and Statistical Analytics Division,
Pacific Northwest National Laboratory,
Richland, WA 99352, USA
E-mail: antonio.sanfilippo@pnl.gov

Kun Huang

Department of Biomedical Informatics,
The OSUCCC Biomedical Informatics Shared Resource,
Ohio State University Medical Center,
Columbus, OH 43210, USA
E-mail: kun.huang@osumc.edu

Biographical notes: Zhongming Zhao received his PhD in Human and Molecular Genetics from the University of Texas Health Science Center at Houston and MD Anderson Cancer Center, Houston, Texas, in 2000. Currently, he is an Associate Professor in the Departments of Biomedical Informatics, Psychiatry, and Cancer Biology at the Vanderbilt University Medical Center. His research interests include bioinformatics and systems biology approaches to studying complex diseases, genome-wide or large-scale analysis of genetic variation and methylation patterns, next-generation sequencing data analysis, comparative genomics and biomedical informatics.

Antonio Sanfilippo is a Chief Scientist in the Computational and Statistical Analytics Division at the Pacific Northwest National Laboratory. His research focus is on computational linguistics, content analysis, knowledge technologies and predictive analytics with reference to cognitive, social, behavioural and biomedical sciences. He holds a Laurea Degree in Foreign Modern Languages awarded *cum laude* from the University of Palermo in Italy, an MA and an MPhil in Anthropological Linguistics from Columbia University, and a PhD in Cognitive Science from the University of Edinburgh (UK).

Kun Huang received his PhD in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign in 2004. Currently, he is an Associate Professor in the Department of Biomedical Informatics and the co-director of the OSU Comprehensive Cancer Center Biomedical Informatics Shared Resource at the Ohio State University Medical Center. His research

interests include computational and systems biology, bioimage informatics, computer vision and machine learning.

The past decade has witnessed an exponential growth of biological data including genomic sequences, gene annotations, expression and regulation and Protein–Protein Interactions (PPIs). A key aim in the post-genome era is to systematically catalogue gene networks and pathways in a dynamic living cell and apply them to study diseases and phenotypes. To promote the research in systems biology and its application to disease studies, we organised a workshop focusing on the reconstruction and analysis of gene networks and pathways in any organisms from high-throughput data collected through techniques such as microarray analysis and RNA-Seq. The workshop was held on 2 August 2010 in Niagara Falls, New York, USA, in conjunction with the 2010 ACM International Conference on Bioinformatics and Computational Biology (ACM-BCB'2010). We invited the researchers to submit their original work in the following areas:

- large-scale or cross-species data integration for the reconstruction of networks and pathways
- module and feature analysis of networks and pathways and applications to diseases and other phenotypes
- quantitative understanding of dynamics of regulatory, signalling, interaction and metabolic networks through modelling and simulation techniques
- enhancement and enablement of pathway discovery through prior biological and biomedical knowledge available in databases and ontologies are extracted from the scientific literature using text-mining techniques
- development of network visualisation and analysis tools.

Our call for papers received excellent responses including a good number of high-quality manuscripts submitted. In consultation with our programme committee, we selected nine papers to present in our workshop at ACM-BCB'2010. We also organised a panel discussion, which was led by Dr. Kun Huang. The panel discussion included three topics:

- data integration from multiple sources to identify gene networks and pathways for complex diseases
- network and pathway analysis using next-generation sequencing data: challenges and opportunities
- identification of tissue-/cell-specific gene networks and pathways.

We invited the authors of those nine papers to expand the scope of their workshop contributions and submit revised versions with new significant results to this special issue. All the papers went through rigorous peer review and five papers were included in this special issue.

The first paper (Huang et al., 2011) investigates the role of *hepatitis C virus* (HCV) in the dynamic protein interaction networks of hepatocellular cirrhosis and carcinoma,

aiming to understand the mechanisms of HCV-induced transformation at the systems level. The authors identify dysfunctional protein networks using gene expression profiles of normal, cirrhotic and Hepatocellular Carcinoma (HCC) liver tissues. Their findings indicate that many cancer genes begin to change during the cirrhosis stage. They also report that HCV protein CORE regulates human nuclear factor Yin Yang 1, which subsequently amplifies and transmits the CORE signal to the insulin, Jak/STAT and TGF-beta pathways. This mechanism can at least partially explain HCV-induced HCC transformation.

Guo et al.'s paper (2011) describes a novel network model that utilises genome-wide co-expression networks and signalling pathways to identify prognostic gene signature of cancer recurrence. The authors specifically analysed lung cancer datasets and identified a 13-gene lung cancer prognostic signature. The authors stated that the 13-gene signature is an accurate prognostic factor compared with clinical covariates, and is involved in cancer oncogenesis in functional pathway analysis.

Zhou et al. (2011) provide a characterisation of the nascent proteome using neuronal cell cultures. The nascent proteome is the proteome that appears early when brain cells receive an ischaemic insult and show the proteomic changes of newly synthesised proteins. Identification of effectors of this phase of response to ischaemia bears the best promise of identifying therapeutic targets for treating acute stroke in emergency clinical situation. The authors compare these nascent proteomes across different ischaemic conditions using bioinformatics tools.

Taylor et al. (2011) explore a novel approach for computer-assisted enrichment of transcriptional regulatory networks inferred from gene expression data. GO-based gene similarity is first tuned to an initial network augmented with gene links mined from PubMed and then used to drive network construction using a bootstrapping algorithm. The authors describe two applications of this approach and discuss its added value in terms of corroboration, annotation and expansion of manually constructed and reversed engineered networks.

The last paper (Zhao et al., 2011) reports an extensive investigation of virus interactions with human proteins at the signal transduction pathways, aiming to understand the functional significance of the interactions between pathogens and their hosts in the context of signal transduction pathways that may be disrupted or altered by PPIs involving the pathogens. After overlaying human-virus PPIs on human signalling pathways, the authors found evidence supporting the hypothesis that viruses often interact with different proteins depending on the targeted pathway. Numerous proteins in virus-targeted pathways are known drug targets, suggesting that these might be exploited as potential new approaches to treatments against multiple viruses.

This workshop and special issue was not possible without the support of ACM-BCB conference organisers or the efforts of many experts who provided valuable advice and critical review comments. We are especially grateful to the anonymous reviewers who helped improve the quality of the manuscripts.

Workshop: Gene Network and Pathway Generation and Analysis

(2 August 2010, Niagara Falls, New York, USA, http://bioinfo.mc.vanderbilt.edu/acmbcb/gnpga.html)

Organisers: Zhongming Zhao, Antonio Sanfilippo, Kun Huang

Program Committee: Zhongming Zhao (co-chair, Vanderbilt University), Antonio Sanfilippo (co-chair, Pacific Northwest National Laboratory), Kun Huang (co-chair, Ohio State University), Judith Blake (The Jackson Laboratory), Olivier Bodenreider (National Library of Medicine, National Institutes of Health), Danail Bonchev (Virginia Commonwealth University), Brent Cochran (Tufts University School of Medicine), Francisco Couto (University of Lisbon), Jan Freudenberg (Feinstein Institute for Medical Research), Anil Jegga (Cincinnati Children's Hospital Medical Center), Cliff Joslyn (Pacific Northwest National Laboratory), Jason McDermott (Pacific Northwest National Laboratory), Karin Rodland (Pacific Northwest National Laboratory), Andrey Rzhetsky (University of Chicago), Roger Simon (Legacy Clinical Research & Technology Center), Mary Stenzel-Poore (Oregon Health & Sciences University), Susan Stevens (Oregon Health & Sciences University), Ron Taylor (Pacific Northwest National Laboratory), Karin Verspoor (University of Colorado), Momiao Xiong (University of Texas Health Science Center at Houston) and Ping Liang (Brock University).