Journal of
Data Mining in Genomics & Proteomics

ISSN: 2153-0602

Editorial
Open Access

Highlights of Recent Articles on Data Mining in Genomics and Proteomics

Natarajan S*

Department of Agriculture, Soybean Genomics and Improvement Laboratory, Beltsville, USA

Advancements in 'Omics' technologies continue to enhance research to solve the world's medical and agricultural problems. The 2012 issues of Data Mining in Genomics & Proteomics solicited contributions from leading authors covering various aspects of data mining technologies for genomics and proteomics, and the application of these technologies to biological sciences.

Zhou et al. [1] developed a computational platform called OmicsMiner to facilitate the design of customized pipelines to process multiple biological datasets. The platform is adaptable and can easily be expanded by integrating new algorithms and protocols. OmicsMiner has a well-designed graphical interface for data processing that utilizes Java protocols, which do not require additional sub-protocol installations.

Gomase and Chitlange [2] describe a microbial proteomic approach to predict the binding ability of antigen protein for selecting nonamers for use in rational vaccine design and increase understanding of the roles of the immune system in infectious diseases. They demonstrated the utility of this approach in a project to examine the tapeworm parasite of sheep, *Taenia ovis,* which causes Sheep Measles. The authors used position specific scoring matrices (PSSM) for antigen design. Prediction of antigen binding ability is important in vaccine development for Sheep Measles and was obtained by using the support vector machine (SVM) based method. A proteomic approach employing 2-dimensional polyacrylamide gel electrophoresis (2D-PAGE) was used by Yang et al. [3] to separate the membrane proteins of the Chinese *Brucella melitensis* M5 strain that causes livestock and human disease. Characterization and identification of immunogenic proteins were accomplished using protein blots, mass spectrometry and bioinformatics. Using this approach, they were able to identify immune reactive proteins that show promise as vaccine candidates. In another protein related article, Sharma and Bhalla [4] reported on the computational analysis of nitrilase enzymes to define unique motifs, and correlated physiochemical and phylogenetic analyses. The results can be applied to the prediction and differentiation of these enzymes based on their substrate specificity and to find new sources of nitrilase across various databases.

Several approaches to disease control and prevention were discussed in articles addressing the association of increased B-cell activating factor (BAFF) expression with disease markers in patients with Lupus Nephritis (LN), the potential advantages of retinoids treatment in LN, and the role of glycoprotein M64 (GPM6A) in human thyroid cancer. Eilertsen et al. [5] utilized quantitative RT-PCR to analyze gene expression in systemic lupus erythematosus patients and further investigated their reaction with clinical, histological and acute phase protein. They found increased BAFF mRNA expression and concluded that the increase is not reflective of histological disease severity. A review article by Khalid et al. [6] presents an overview of therapeutic use of retinoids, a group of natural and synthetic derivatives of vitamin A for LN. The review discusses the details of LN pathogenesis and severity in humans, and current and new agents for the treatment.

Computational analysis of microarray data and serial analysis of gene expression enabled Khalid et al. [6] to show that the protein GPM6A plays a major role in cell migration and is a major contributor in tumor formation. The results could lead to earlier diagnosis and better treatment for thyroid cancer patients. Sequence alignment and protein structural analysis were used by Holmes and Cox [7] to study the evolution of vertebrate pancreatic lipase related proteins which perform a regulatory role in lipolysis.

Muller and Pfaffl [8] demonstrate that principal component analysis (PCA) can identify off-target effects of RNAi-knockdown experiments. RNA interference is used for gene silencing and it can be applied with therapeutic aims or for identification of drug targets. A major problem of RNAi knockdown application is regulation of non-specific targets in the treated cell culture. The authors conclude that PCA derived plots more clearly visualize off-target effects compared to the commonly used heat maps. Garcia et al. [9] studied the genetic and population variability of virus infecting maize using interactive network exploration and analysis of network topological properties.

References

1. Zhou Z, Netzer M, Lee IH, Handler M, Manickam VA, et al. (2012) OmicsMiner: A Biological Data Mining Framework. J Data Mining in Genom Proteomics 3: 115.

2. Gomase VS, Chitlange NR (2012) Microbial Proteomics Approach for Sensitive Quantitative Predictions of MHC Binding Peptide from *Taenia ovis*. J Data Mining Genomics Proteomics 3: 121.

3. Yang Y, Yin J, Wang X, Wang X, Wang F, et al. (2012) Immunoproteomic Analysis of Chinese *Brucella* Vaccine Strain M5 and New Immunogenic Proteins. J Data Mining Genomics Proteomics 3: 117.

4. Sharma N, Bhalla TC (2012) Motif Design for Nitrilases. J Data Mining Genomics Proteomics 3: 119.

5. Eilertsen G, Ghelue MV, Nossent JC (2012) BAFF Expression is Increased in Patients with Lupus Nephritis and Associated with Antinucleosome Antibodies, C1 Inhibitor, A-1-Acid-Glycoprotein and Endothelial Activation Markers. J Data Mining Genomics Proteomics 3: 113.

6. Khalid Z, Sameen S, Malik SI, Shehzad S (2012) Computational Analysis on the Role of GPM6A in Human Thyroid Cancer. J Data Mining Genom Proteomics 3: 114.

7. Holmes RS, Cox LA (2012) Bioinformatics and Evolution of Vertebrate Pancreatic Lipase and Related Proteins and Genes. J Data Mining Genom Proteomics 3: 111.

8. Müller J, Pfaffl MW (2012) RNA Interference Off-target Screening using Principal Component Analysis. J Data Mining Genomics Proteomics 3: 116.

9. Garcia MA, de la Paz Gimenez Pecci M, Cabral JB, Castillo AN, Laguna IG (2012) Interactive Network Exploration in the Kdd Process, Contributions in the Study of Population Variability of a Corn Fijivirus. J Data Mining Genomics Proteomics 3: 120.

*Corresponding author: Natarajan S, Department of Agriculture, Soybean Genomics and Improvement Laboratory, Beltsville, USA, E-mail: ssavithiry@yahoo.com, rajannatarajan@yahoo.com