Proteomics-2013
OMICS Group Conferences
Accelerating Scientific Discovery
3rd International Conference on
Proteomics & Bioinformatics
July 15-17, 2013 Courtyard by Marriott Philadelphia Downtown, USA

A bayesian feature selection approach for biomarker discovery from integrated omic studies

Bobbie-Jo Webb-Robertson
Pacific Northwest National Laboratory, USA

Abstract

High-throughput technologies currently have the capability to capture information at both global and targeted scales for the transcriptome, proteome and metabolome, as well as determining functional aspects of these biomolecules. The promise of data integration is that by utilizing these disparate data streams a more complete or accurate estimate of system behavior can be obtained. In the case of biomarker discovery to better diagnose and predict outcomes of disease, one goal is to identify the best subset of molecules that can separate specific phenotypes of interest. However, in a space of tens of thousands of variables (e.g., genes, proteins), feature selection approaches often yield over-trained models with poor predictive power. Moreover, feature selection algorithms are typically focused on single sources of information and do not evaluate the effect on downstream statistical integration models. Bayesian statistics has been shown to be an effective approach for statistical integration across multiple data streams using standard meta-learning approaches. We present a sampling strategy to evaluate potential variables spaces and using the Bayesian framework obtain probabilistic measures associated with one more potential variable and dataset collections to evaluate the confidence in potential solutions. In addition, using standard Bayesian statistical methods a marginal probability associated with individual features can be obtained to aid both in biomarker selection and in downstream analyses, such as pathway modeling. We utilize the Bayesian sampling feature selection approach on several disease-based proteomics datasets (diabetes and cancer) to demonstrate the flexibility and robust nature of the Bayesian sampling approach.

Biography

Bobbie-Jo Webb-Robertson completed her M.E. in Statistics and Operations Research and her Ph.D in Decision Sciences and Engineering Systems from Rensselaer Polytechnic Institute. She joined Pacific Northwest National Laboratory in 2002 and is currently working as a Senior Research Scientist in the Computational Biology and Bioinformatics group. Her research is primarily focused in the fields of proteomics and statistical data integration of high-throughput data. She is currently an author on over 55 peer-reviewed journal articles and in an executive editor of the Journal of Proteomics and Genomics Research.