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

Gene set association analysis using the coefficient of intrinsic dependence

Chen-An Tsai and Li-Yu Daisy Liu
National Taiwan University, Taiwan

Gene set testing problem has become the focus of microarray data analysis. A gene set is a group of genes that are defined by a priori biological knowledge. Several statistical methods had been proposed to determine whether functional gene sets express differentially (enrichment and/or deletion) in variations of phenotypes. However, little attention has been given to analyze the dependence structure among gene sets. In this study, we have proposed a novel statistical method of gene set association analysis to identify significantly associated gene sets using the coefficient of intrinsic dependence. The simulation studies show that the proposed method outperforms the conventional methods to detect general forms of association in terms of control of type I error and power. The correlation of intrinsic dependence has been applied to a breast cancer microarray dataset to quantify the un-supervised relationship between two sets of genes in the tumor and non-tumor samples. It was observed that the existence of gene-set association altered in different clinical cohorts. In addition, a supervised learning was employed to illustrate that how gene sets in signaling transduction pathways or subnetworks regulated by a set of transcription factors can be discovered using microarray data. In conclusion, the coefficient of intrinsic dependence provides a powerful tool to detect general types of association. Hence, it is applicable to associate gene sets using microarray expression data. Through connecting relevant gene sets, our approach has the potential to reveal underlying associations via drawing a statistically relevant network in a given population, and can be used to complement the conventional gene set analysis.

Biography

Chen-An Tsai earned his Ph.D. degree in statistics from the National Central University, Taiwan in 1999. He started his professional career as a postdoctoral fellow at the National Center for Toxicological Research, FDA, USA. He is now an Associate Professor at the National Taiwan University. His research interests include bioinformatics, biostatistics, statistical machine learning, and the development of statistical approaches for medical data. He has published more than 30 papers in reputed journals and has been serving as an associated editor of BMC Bioinformatics.