An Integrative Approach to Constructing MicroRNA-Gene Networks in Ovarian Cancer

Andrew Quitadamo, Benika Hall, and Xinghua Shi
Department of Bioinformatics and Genomics
University of North Carolina at Charlotte
9201 University City Blvd., Charlotte, NC 28223.
aquitada,bjohn157,x.shi@uncc.edu

August 10, 2015

Abstract

Network integration is critical in understanding the underlying mechanisms of human health and diseases. Changes in miRNA and mRNA expression are known to be involved in both ovarian cancer development and progression. Pin-pointing the exact changes and the relationships that occur between them could lead to advances in how ovarian cancer is treated and diagnosed. Creating an integrated network involving eQTLs, miRNA targets, protein-protein interactions and correlation graphs is one way to explore these relationships. Integrating multiple data sources can thus allow us to create a wider and more holistic view of the interactions in ovarian cancer. Therefore, we developed a new method of constructing an integrated network by combining the strength of association study and network analysis. Applied to ovarian cancer, our integrated analysis replicated known cancer related miRNAs and genes, in addition to providing new candidate markers.

1 Introduction

Ovarian cancer is a deadly female reproductive cancer. Understanding the biological mechanisms underlying ovarian cancer could lead to quicker and more accurate diagnosis and more effective treatments. Both changes in microRNA(miRNA) expression and miRNA/mRNA dysregulation have been associated with ovarian cancer. With the availability of whole-genome miRNA and mRNA sequencing we now have new potentials to study these associations. Although various efforts have focused on identifying miRNAs with direct effect on gene expression. However, it is still unclear to what extent that changes in miRNA expression affect gene expression, and bring about fluctuations in biological pathways and networks in ovarian cancer. In order to understand such

complex effects of miRNA expression on gene expression, we develop an integrated network approach that combines eQTL analysis with network analysis using various types of relationships and interactions among miRNAs and genes.

2 Methods

The overall workflow of our integrative approach includes the following five phases, including data preprocessing, classical eQTL analysis, sparse learning based network association, statistical network expansion and network integration. Prior to analysis, we performed data pre-processing including data normalization. Next, we performed classical eQTL analysis between miRNA and gene expression, where pairs of miRNA and genes were tested using a linear regression model followed by multi-test correction. We then used a two-graph guided multi-task LASSO model to identify network associations considering the correlations among miRNA and gene expressions respectively. We consequently incorporated the known target genes of those miRNA eQTLs. We applied a statistical method that extended the network utilizing protein protein interactions. By combining all the associations and interactions described above, we finally constructed an integrative network incorporating various data types of relationships among miRNA and gene expressions.

3 Results

Our integrative approach constructed an integrative network that illustrates the complex interplay among miRNA and gene expression from a systems perspective. This integrated network took account of miRNA eQTL associations, miRNAs and their targets, protein-protein interactions, co-expressions among miRNAs and genes respectively. Applied to the ovarian cancer data set from The Cancer Genome Atlas (TCGA), we created an integrated network with 167 nodes containing 108 miRNA-target interactions and 145 from protein-protein interactions, starting from 44 initial eQTLs. This integrated network encompassed 26 genes and 14 miRNAs associated with cancer. In particular, 11 genes and 12 miRNAs in the integrated network are associated with ovarian cancer.

4 Conclusions

We demonstrated an integrated network approach that integrates multiple data sources at a systems level. We applied this approach to the TCGA ovarian cancer dataset, and constructed an integrated network that provided a more inclusive view of miRNA and gene expression in ovarian cancer. In the future, we will extend our approach by incorporating other types of network information such as regulatory and signaling pathways.