A Network Approach for Integrative Analysis of Genomic Data in Ovarian Cancer

Andrew Quitadamo and Xinghua Shi

College of Computing and Informatics, University of North Carolina at Charlotte

Background

Ovarian cancer causes 5% of cancer deaths in women. While the 5 year relative survival is above 90% when ovarian cancer is diagnosed in the earliest stage, the overall 5 year relative survival rate rapidly decreases to 44%. Thus, a better understanding of the biological mechanisms of ovarian cancer is needed to for earlier diagnosis and more effective treatment.

Methods

As shown in Fig. 1., We downloaded CpG methylation and gene expression data for 268 samples of serous ovarian carcinoma from The Cancer Genome Atlas (TCGA) [1]. We applied two orthologous methods to obtain networks that capture the relationship between DNA methylation and gene expression. 1) We first used a pair-wise approach, Matrix eQTL [2], to find highly correlated cis Expression Quantitative Trait Loci (meQTLs). The cis meQTLs genereated a list of CpG genes and the perturbed genes. These genes were treated as seed nodes respectively, and two networks were expended using DAPPLE [3], which utilizes protein-protein interactions to build networks from seed genes. 2)We then used a two-graph guided multi-task lasso model (2GLasso) [4] that simultaneously identify meQTLs and build networks. Finally, we merged the 2GLasso network with DAPPLE expended meQTL networks by collapsing common nodes (Fig. 2).

Fig. 1: Workflow

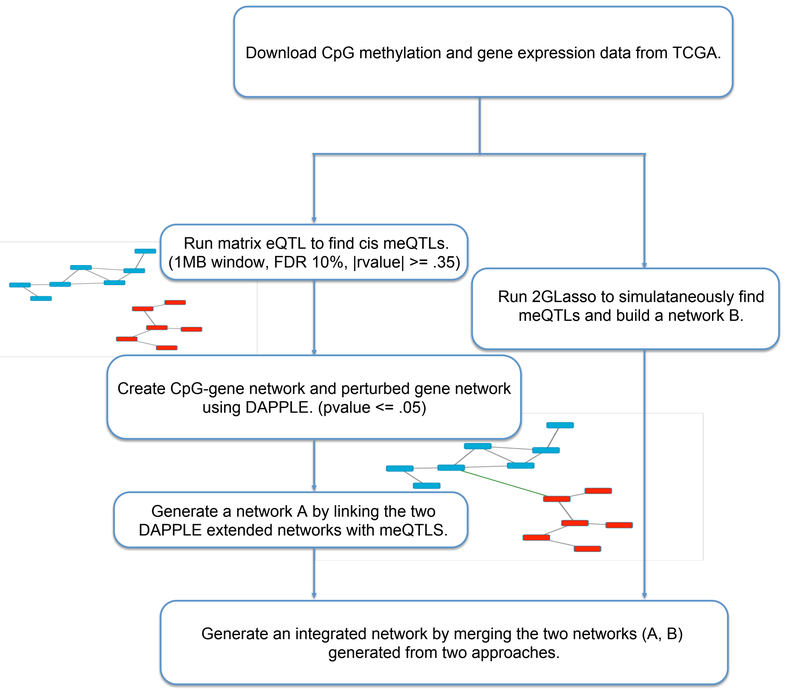
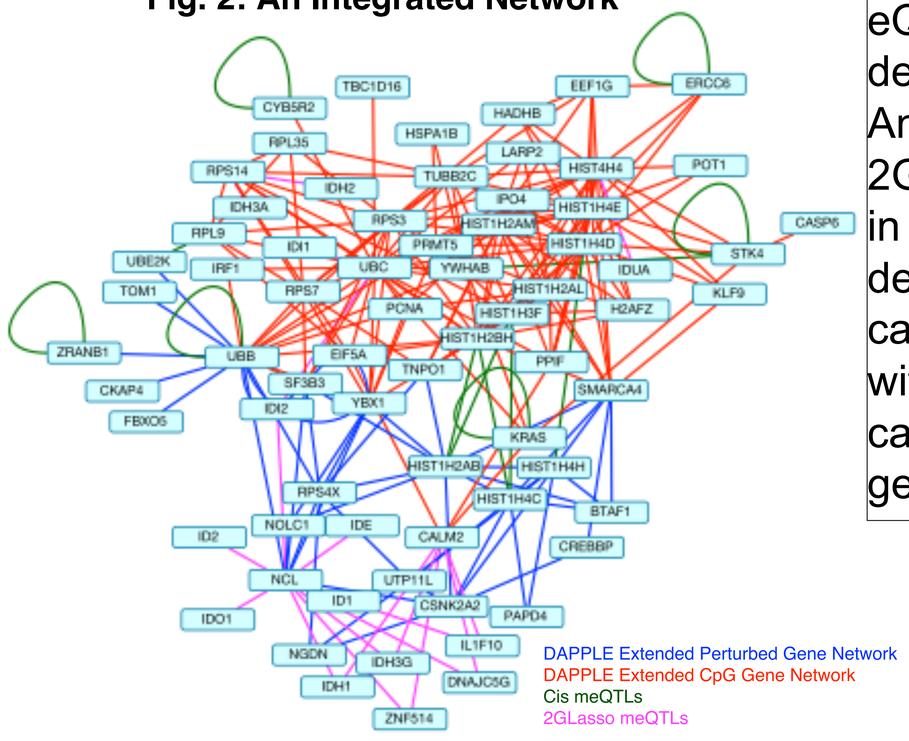


Fig. 2: An Integrated Network



Results

The two networks (Fig. 2) constructed using two different methods (e.g. Matrix eQTL and DAPPLE, 2GLasso), albeit with some common genes, capture different underlying mechanisms regarding how CpGs affect gene expression in ovarian cancer. Thus, our integrative network that consolidates these differently constructed networks provide a comprehensive view of DNA methylation affecting gene expression. Such an integration provides a list of known cancer genes (Table 1). For example, *YBX1* found in the matrix eQTL analysis has been associated with decreased ovarian cancer survival [6]. Another gene *IDO1,* found from the 2GLasso analysis, has been implicated in increased paclitaxel resistance and decreased survival in serous ovarian cancer [7]. Those genes that interact with known cancer genes in the network can harbor potential candidate cancer genes for future investigation.

Table 1. Cancer Genes in the Networks

	DAPPLE Extended CpG-Gene Network	DAPPLE Extended Gene Network	2GLasso Network	Integrated Network
Genes	40	24	22	77
Cancer Related Genes [8]	6	3	1	8
Ovarian Cancer Genes [9]	4	1	1 [7]	5

Summary

We develop a network approach that integrates DNA methylation and gene expression data in TCGA ovarian cancer. We apply two methods to obtain networks by mining the associations between DNA methylation and gene expression. We then build an integrated network by consolidating these networks constructed from different methods. Our further analysis of the integrated network points to a network view of how epigenetic signature, particularly DNA methylation, perturbs gene regulation and leads to tumorigenesis and cancer progression in ovarian cancer.

References:

[1] The Cancer Genome Atlas Research Network (2011). Integrated genomic analyses of ovarian carcinoma. Nature. 474(7353):609-15.
[2] Rossin EJ, et al. (2011) Proteins Encoded in Genomic Regions Associated with Immune-Mediated Disease Physically Interact and Suggest Underlying Biology. PLoS Genetics. 7(1): e1001273.
[3] Shabalin AA. (2012) Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. Bioinformatics. 28(10):1353-1358.
[4] Chen X, et al. (2012). A Two-Graph Guided Multi-task Lasso Approach for eQTL Mapping. Journal of Machline Learning Research (JMLR) W&CP 22: 208-217.
[5] Smoot M, et al. (2012). Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics. 27(3): 431–432.
[6] Panupinthu N, Yu S, Zhang D, Zhang F, Gagea M, Lu Y, Grandis JR, Dunn SE, Lee HY, Mills GB. Self-reinforcing loop of amphiregulin and Y-box binding protein-1 contributes to poor outcomes in ovarian cancer. Oncogene. 2014 May 29;33(22):2846-56.
[7] Aikou Okamoto, Takashi Nikaido, Kazunori Ochiai, et al. Indoleamine 2,3-Dioxygenase Serves as a Marker of Poor Prognosis in Gene Expression Profiles of Serous Ovarian Cancer Cells. Clin Cancer Res 2005;11:6030-6039.
[8] Atlas of Genetics and Cytogenetics in Oncology and Haematology. URL http://AtlasGeneticsOncology.org
[9] Wall DP, Pivovarov R, Tong M, Jung JY, Fusaro VA, DeLuca TF, Tonellato PJ: Genotator: a disease-agnostic tool for genetic annotation of disease. BMC Med Genomics 2010, 3:50.