Google goes cancer: improving outcome prediction for cancer patients by network-based ranking of marker genes.

Predicting the clinical outcome of cancer patients based on the expression of marker genes in their tumors has received increasing interest in the past decade. Accurate predictors of outcome and response to therapy could be used to personalize and thereby improve therapy. However, state of the art methods used so far often found marker genes with limited prediction accuracy, limited reproducibility, and unclear biological relevance. To address this problem, we developed a novel computational approach to identify genes prognostic for outcome that couples gene expression measurements from primary tumor samples with a network of known relationships between the genes. Our approach ranks genes according to their prognostic relevance using both expression and network information in a manner similar to Google’s PageRank. We applied this method to gene expression profiles which we obtained from 30 patients with pancreatic cancer, and identified seven candidate marker genes prognostic for outcome. Compared to genes found with state of the art methods, such as Pearson correlation of gene expression with survival time, we improve the prediction accuracy by
up to 7%. Accuracies were assessed using support vector machine classifiers and Monte Carlo cross-validation. We then validated the prognostic value of our seven candidate markers using immunohistochemistry on an independent set of 412 pancreatic cancer samples. Notably, signatures derived from our candidate markers were independently predictive of outcome and superior to established clinical prognostic factors such as grade, tumor size, and nodal status. As the amount of genomic data of individual tumors grows rapidly, our algorithm meets the need for powerful computational approaches that are key to exploit these data for personalized cancer therapies in clinical practice.