A Novel Multi-omic Approach to Identifying Cancer Recurrence-Associated Gene Signatures

Sheng, Kevin (School: Roanoke Valley Governor's School for Science and Technology)

Cancer biomarkers are measurable variables indicative of disease presence or progression. However, inadequate methods of biomarker identification combined with the underlying complexity of cancer genetics has led to variability in the predictive power of current molecular biomarkers, limiting their clinical efficacy. Here, a novel biomarker discovery pipeline is reported that integrates data from global methylation and transcriptomic profiling and data from genome-wide RNA interference screens enriched with genes essential for cancer cell survival. By applying this pipeline to a large retrospective colon adenocarcinoma (COAD) patient cohort compiled from four independent datasets, novel biomarkers were identified and designated as a COAD recurrence gene signature (COAD-RGS). Compared to existing biomarkers and clinically-available predictive tests, COAD-RGS was more accurate in predicting patient prognosis based upon area under the receiver operating curve values. Further analyses revealed that COAD-RGS accurately stratified patients with high- or low-risk of recurrent disease since high-risk patients displayed worse prognostic outcomes, increased frequency of tumor recurrence, and poor responses to chemotherapy. The accuracy in predictions of tumor recurrence and patient prognosis by COAD-RGS was robustly validated in both an internal validation cohort and a test cohort compiled from three additional independent datasets. Taken together, these results demonstrate the feasibility of the novel pipeline developed in this study in identifying novel and reliable cancer biomarkers.