Identifying analytes associated with poor prognosis in colorectal adenocarcinoma

Christi L. French[1] and Alissa M. Weaver[1]

Short Abstract — **A key question in cancer systems biology is how to use molecular data to predict the biological behavior of tumors from individual patients. While genomics has been heavily studied, protein signaling data is more directly connected to biological phenotype and might predict cancer cell and tumor phenotypes such as invasion, metastasis, and patient survival. In this study, we mined publicly available data for colorectal adenocarcinoma from the Cancer Genome Atlas (TCGA) and identified protein expression and signaling changes that are statistically associated with patient outcome. This approach appears promising as a discovery method for identification of prognostic and therapeutic markers.**

Keywords — **Bioinformatics, big data, reverse phase protein array, TCGA, colorectal adenocarcinoma, prognosis**

I. Purpose

Cancer patients with undetectable micrometastases at the time of diagnosis are in danger of subsequent metastatic growth that worsens their outcome. In colorectal carcinoma (CRC), 20-30% of stage II patients will have a recurrence at a distant site after removal of the primary tumor [1]. Treatment of these patients is variable, with many patients receiving no treatment other than surgery and others receiving additional systemic therapy. Identification of patients with latent metastases could reduce morbidity and mortality by helping physicians to choose whether to treat early stage patients with an adjuvant therapy. The goal of our study was to identify molecular markers or signatures that can identify patients that are "at risk" of tumor recurrence at a distant site.

II. Results

To identify molecular markers that are predictive of poor patient outcome, we analyzed publicly available data from the Cancer Genome Atlas (TCGA) for colorectal adenocarcinoma tumors [2]. We specifically focused on reverse phase protein array (RPPA) data as the most proximal data set to signaling changes that might drive aggressive tumor behavior. We analyzed the RPPA data by peforming statistical analyses to identify molecular analytes that were predictive of poor patient prognosis, as defined by tumor recurrence, patient death, or the presence of positive lymph nodes or distant metastases (tumor stage) at the time of diagnosis. Cox and Cox-LASSO methods were used to identify single analyte predictors whereas Wilma and LEAPS algorithms were used to identify groups of analytes associated with poor prognosis. Using these methods, survival and recurrence were superior to tumor stage with respect to identifying analytes that could segregate patients according to outcome. Furthermore, a number of analytes were chosen across all methods used, giving confidence in their potential predictive power. Interestingly, a number of predictive analytes appeared to be related to the presence of tumor stromal and immune cells in the samples. We are currently performing followup staining experiments of human CRC tumor samples to identify location and positivity of potential predictive markers.

III. Conclusions

The use of patient followup information combined with bioinformatics analysis of tumor signaling appears promising as a discovery tool to identify meaningful predictors of patient prognosis from large datasets. This general approach may serve as a useful method to identify prognostic indicators and therapeutic targets in a variety of tumor types.

References

[1] Shibayama M., et al. (2011) Prediction of metastasis and recurrence in colorectal carcinoma based on gene expression analysis: ready for the clinic? *Cancers* **3**, 2858-2869.

[2] The Cancer Genome Atlas Network. (2012) Comprehensive molecular characterization of colon and rectal carcinoma. *Nature* **487**, 330-337.

Acknowledgements: This work was funded by a GI SPORE pilot project grant to AW. The PI of the GI SPORE is Robert Coffey.

[1]Department of Cancer Biology, Vanderbilt University Medical Center. E-mail: Alissa.weaver@vanderbilt.edu