Breast Cancer Gene Expression Modules: A Platform for Companion Diagnostics Development

Scott A. Tomlins1, SethSadis2, Daniel R. Rhodes2,3

1Michigan Center for Translational Pathology, Ann Arbor, MI
2Compendia Bioscience, Ann Arbor, MI.

Abstract:
Background: Gene expression patterns in breast cancer are capable of stratifying patients into significantly different risk categories. Disease behaviors and responses to therapy are more accurately predicted using this baseline gene expression data. Subtypes are also better defined, and new applications for breast cancer diagnosis and treatment can be implemented. Despite the development of the breast cancer diagnosis, only limited studies have been performed using expression profiling data for clinical trials and treating patient outcomes.

Methods: We performed a co-expression analysis and identified highly co-expressed sets of genes (i.e. modules) across multiple breast cancer microarray datasets. Candidate gene modules were generated and prioritized based on disease response and outcome data. Each module was scored across all breast cancer datasets described in Table 1 (see Methods). Data relationships were characterized for all pair wise comparisons by linear regression analysis. Module scoring derived a module score as the average or median of the representative genes. Module scores can be used to define patient subtypes for novel diagnostic tests.

Results:
A total of 10 modules were derived with significant correlation across the breast cancer datasets. Nine modules demonstrated association with recurrence, including the proliferation module which was positively associated with recurrence, consistent with prior knowledge. The proliferation module is highly interlinked with the cell cycle, histone, and cell growth pathways. The Basal module was associated with non-metastatic breast cancer progression, consistent with prior knowledge. The Basal and proliferation modules were associated with non-metastatic disease progression. The Basal and cell cycle modules were associated with metastatic disease progression. The Basal and proliferation modules were associated with non-metastatic disease progression. The Basal and cell cycle modules were associated with metastatic disease progression.

Conclusions: The breast cancer segmentation panel used in this study was derived using a supervised learning approach. This panel can be used to derive a module score as the average or median of the representative genes. The module scores can be used to define patient subtypes for novel diagnostic tests. The breast cancer segmentation panel can be used to derive a module score as the average or median of the representative genes. The module scores can be used to define patient subtypes for novel diagnostic tests.