CLIA-certified cancer gene panel-based machine learning method to predict sensitivity of anticancer drugs for precision oncology

CLIA certified molecular/genetic panel testing of formalin-fixed, paraffin embedded (FFPE) material including studies of small biopsies offers the potential to identify individualized treatments that target specific genetic alterations such as EGFR mutation. However, molecularly-guided therapy is only available for the minority of lung cancer patients carrying such alterations for targeted drugs (e.g., ~15% of lung adenocarcinoma); thus the selection of chemotherapy or other treatment for the majority of non-small-cell lung cancer (NSCLC) patients without such alterations is still limited. In addition, despite the early success of targeted therapy in NSCLC patient care, patients treated with targeted drugs often developed resistance to these treatments. Thus, it is critical to build a predictive model based on information of genetic panel testing to predict sensitivity/resistance of drug for individualized treatment stratification. To tackle this challenge, we develop a novel machine learning approach called Robust Bayesian Matrix Factorization (RBMF) to integrate genetic information on a large panel of non-small cell lung cancer (NSCLC) lines (e.g., Single Nucleotide Variants (SNVs) found on targeted gene panel or whole exome sequences) with large-scale drug/chemical compound screening profiles on these same NSCLC lines to (a) discover a genetic variation-based predictive biomarker(s) and (b) use this to predict response of drugs in other NSCLC lines (and ultimately in patients). The RBMF method leverages information across multiple related drug/chemical compound screening profiles that have similar mechanisms of actions/targets as well as samples (e.g., NSCLC lines) with similar genetic variant profiles (i.e., exploring clusters of drugs and samples), thus can be robust against noise from each data/drug screening experiment and more accurate to predict sensitivity of drugs.

In experiments with our institutional drug/chemical screening profiles and SNVs present in known cancer-related genes and/or a commercially available genetic panel such as *FoundationOne* in NSCLC cell lines, the RBMF method showed better prediction performance compared to the state-of-the-art methods. Moreover, the RBMF method identified novel mutation-drug sensitive/resistant associations that can serve as a predictive biomarker to stratify patients. Independent validations with Genomics of Drug Sensitivity in Cancer and Cancer Cell Line Encyclopedia datasets demonstrated that the RBMF consistently outperformed current state-of-the-art methods for sensitivity prediction for well-known cancer drugs.

Taken together, our proposed method demonstrated the clinical utility of the use of genetic panels to predict drug response in NSCLC lines. Furthermore, the novel mutation-drug sensitive/resistant association discovered by the RBMF method could provide unprecedented opportunities to develop a clinical assay as a predictive biomarker, which could individualize treatments based on the genetic information of cancer patients.