Prediction of Response to Immunotherapy in Cancers via Transfer Learning

Peter Langfelder[1,2], En-Tni Lin[1], Grace S. Shieh[1,3,4,5*]

[1]Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

[2]Center for Neurobehavioral Genetics, The Jane and Terry Semel Institute for Neuroscience and Human Behavior, UCLA, Los Angeles, U.S.A.

[3]Bioinformatics Program, Taiwan International Graduate Program, Academia Sinica, Taipei, Taiwan

[4]Data Science Degree Program, Academia Sinica and NTU, Taipei, Taiwan

[5]Genome and Systems Biology Degree Program, Academia Sinica and NTU, Taipei, Taiwan

Abstract

To date, immunotherapies such as immune checkpoint inhibitors (ICIs) have emerged as a leading treatment for metastatic cancer, significantly improving patient survival while causing relatively few side effects. However, the objective response rate for ICIs remains low, approximately 25% for metastatic urothelial carcinoma (mUC) and renal cell carcinoma (mRCC), and ~40% for melanoma, underscoring the urgent need for predictive response biomarkers. Several state-of-the-art signatures have been revealed in top-tier journals, highlighting the importance of this field. As the number of genes (~20,000) far exceeds the sample sizes of typical training sets (generally ≤ 300), we first developed feature selection procedures to reduce the number of features to a few hundred. We then trained multiple machine learning classifiers using the selected genes and the IMvigor210, IMmotion150, and Gide cohorts, which includes RNA-seq and clinical data from 298, 77, and 91 patients with mUC, mRCC, and melanoma, respectively. Notably, our predictor LogitDA using the identified gene signatures achieved a prediction AUC of 0.75, 0.83, and 0.71~0.75 in independent cohorts, PCD4989g (mUC, mRCC), and three melanoma cohorts, respectively. Moreover, our signatures outperformed (most of) six state-of-the-art signatures, PD-L1 IHC, and five tumor microenvironment signatures, including IFN-γ, T-effector, and T-cell exhaustion signatures in mUC (mRCC and melanoma). From our signatures, we identified key prognostic biomarkers in mUC, mRCC, and melanoma, respectively.

Keywords: biomarker, cancer, machine learning, regression, prediction