2025 – 02

Gene signature for response prediction to immunotherapy in mUC and RCC

Peter Langfelder1, 2, 3, En-Tni Lin1, Yi-Ta T sai4,
Tai-Lung Cha5, Grace S. Shieh1, 6, 7, 8*

May 28, 2025

Gene signature for response prediction to immunotherapy in mUC and RCC

Peter Langfelder[1, 2, 3], En-Tni Lin[1], Yi-Ta Tsai[4], Tai-Lung Cha[5], Grace S. Shieh[1, 6, 7, 8*]

[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, University of California, Los Angeles, Los Angeles, United States

[3] Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, United States

[4] Graduate Institute of Life Sciences, National Defense Medical Center, Taipei, Taiwan

[5] National Institute of Cancer Research, National Health Research Institutes, Miaoli, Taiwan

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

[7] Data Science Degree Program, Academia Sinica and National Taiwan University, Taipei, Taiwan

[8] Genome and Systems Biology Degree Program, Academia Sinica and National Taiwan University, Taipei, Taiwan

*** Correspondence:**

Grace S. Shieh

gshieh@stat.sinica.edu.tw

Abstract

To date, 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 20-25% in urothelial carcinoma (UC) and renal cell carcinoma (RCC), 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 and IMmotion150 datasets, which includes RNA-seq and clinical data from ~298 and 77 patients with metastatic UC (mUC) and metastatic RCC (mRCC), respectively. Notably, our predictor LogitDA, using the identified 49 (27)-gene signature, achieved a prediction AUC of 0.75 (0.72) in independent datasets, PCD4989g(mUC and mRCC). Moreover, our signature outperformed six state-of-the-art signatures, PD-L1 IHC, and five tumor microenvironment signatures, including IFN-γ, T-effector, and T-cell exhaustion signatures. Our signature for mRCC was second to T exhaust in prediction AUC but surpassed the six known signatures in prediction accuracy.

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