[BIOINFO 2025]

Towards Precision Immunotherapy: A Multimodal AI Approach Combining Microbiome Taxonomy and Genomic Data for Drug Response Prediction

Jae Woo Baek[1], Suyeon Lee[1], Hyeong Joon Ahn[1], Choong Hwan Choi[2], Sunjae Lee[1], and Junghyun Namkung[2]

[1]School of Life Sciences, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea

[2]SK AI R&D Center, SK Telecom, Seongnam, Republic of Korea

Background: The gut microbiome significantly influences immune checkpoint inhibitor (ICI) response rates by modulating anti-tumor immunity through bacterial metabolites and direct microbial interactions with immune cells. Existing prediction models demonstrate limited accuracy and poor generalizability, primarily relying on taxonomic information alone while rarely exploring functional genomic features for response prediction.

Objectives: We developed a multimodal AI framework integrating taxonomic and functional genomic profiles to enhance prediction accuracy and generalizability for ICI response prediction.

Methods: We analyzed 114 samples from the PRIMM study dataset, applying data augmentation techniques to address class imbalance and enhance inter-cohort variability. To preserve hierarchical taxonomic structure and leverage functional relationships, we employed image representation approaches for taxonomy data visualization. Gene functions were mapped to Enzyme Commission (EC) numbers, and convolutional neural networks (CNNs) were applied to both taxonomic and functional profile matrices. The resulting embedded features from both modalities will be integrated through fully connected layers to generate fused representations.

Results: The CNN-based models achieved AUCs of 0.66 when using only taxonomy features and 0.69 when using only gene functions. These results represent substantial improvements of 27% and 11% respectively, over corresponding logistic regression models, highlighting the effectiveness of deep learning approaches in capturing complex data patterns. The ensemble integration of both CNN-based models further enhanced predictive performance, achieving an optimal AUC of 0.73.

Conclusion: We successfully demonstrated that the CNN-based models effectively capture distinct predictive signals from both taxonomic and functional microbiome data. Building upon this, we focus on fusing these complementary data streams within a multimodal framework. This integrative approach represents a promising advancement toward more accurate and generalizable ICI response prediction.

Keywords: SMOTE sampling, Microbiome, Drug response prediction, Multimodal AI, Immune checkpoint inhibitors, PRIMM dataset