An integrative multi-omics and machine learning framework for diagnostic biomarker identification in patients with liver cirrhosis

Contact information
Ki Tae Suk, MD, PhD
ktsuk@hallym.ac.kr

I G Park[1], S J Yoon[1] , S M won[1] , K K oh[1] , U J Lee[3] , Y L HAM[2] , K T Suk[1]
[1] Institute for Liver and Digestive Diseases, Hallym University, Chuncheon, Korea, Rep. of South
[2] Hallym University, Department of Electronic Engineering, Chuncheon, Korea, Rep. of South
[3] Daewon University College, Department of Nursing, Jaecheon, Korea, Rep. of South

EASL CONGRESS
7-10 May 2025
Amsterdam, the Netherlands

Introduction

Cirrhosis is the final stage of chronic liver disease and presents substantial diagnostic challenges due to its multifactorial complexity. Traditional diagnostic methods often lack accuracy and sensitivity for early detection. Recent advances in multi-omics integration have shown promise in identifying novel biomarkers. In particular, the gut microbiome plays a key role in liver disease pathogenesis and exhibits distinct alterations in cirrhotic patients.

This study aims to integrate microbial, functional gene, and metabolite data derived from fecal shotgun metagenomics to develop machine learning-based models for cirrhosis diagnosis and identify key biomarkers across multiple omics layers.

Method

This study included 34 patients with cirrhosis and 49 non-cirrhotic controls. Shotgun metagenomics data were processed to extract microbial species profiles using MetaPhlAn and KEGG Orthology (KO) gene profiles using HUMAnN. These profiles were subsequently integrated with fecal metabolite data.

Feature selection was independently performed for each omics dataset (species, KO genes, metabolites) using the Boruta algorithm. Selected features from each omics layer were used to train machine learning models including Support Vector Machine, Random Forest, 1D Convolutional Neural Network, and Multi-Layer Perceptron. Model performance was evaluated using 10-fold cross-validation with AUROC as the evaluation metric.

Finally, selected features from all omics types were integrated to build an optimal multi-omics model, which was also evaluated using the same machine learning framework.

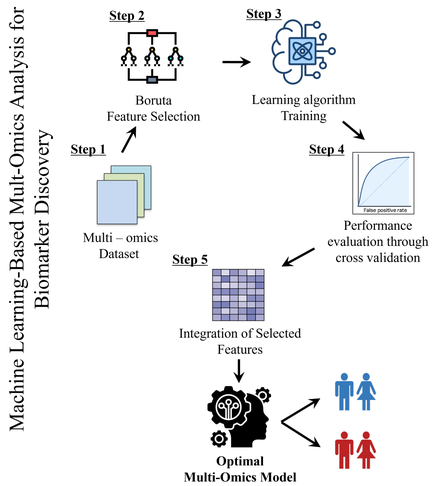

Fig 1. Workflow of machine learning-based multi-omics biomarker discovery

Conclusions

This study highlights the contribution of multi-omics data integration in improving diagnostic accuracy and underscores the potential of machine learning approaches in identifying key biomarkers associated with cirrhosis. Our findings suggest that multi-omics analysis provides new directions for liver disease diagnosis and personalized treatment strategies, offering promising avenues for future clinical applications and research in liver disease management.

Results

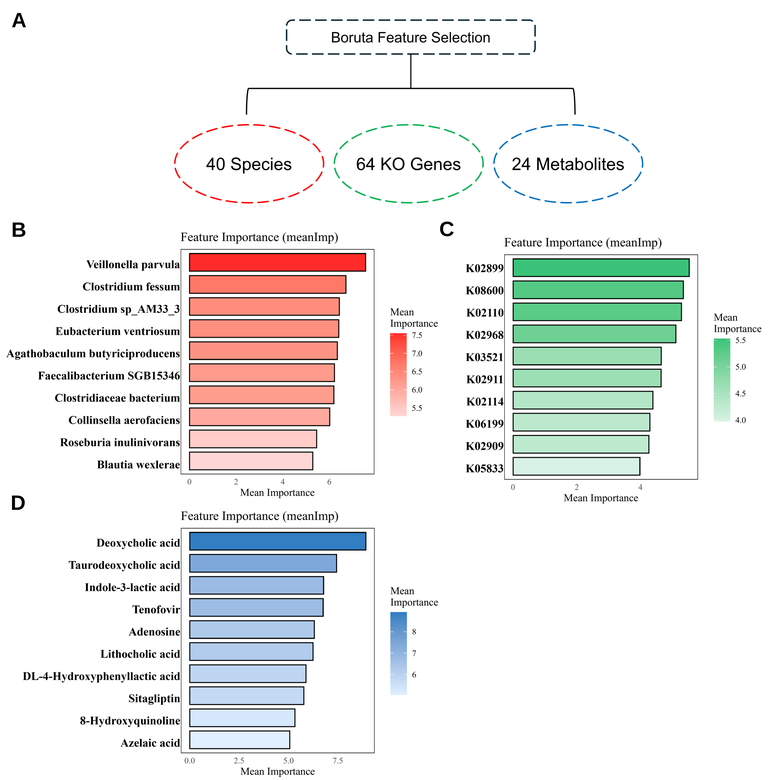

Fig 2. ML-based Boruta Feature Selection for Multi-Omics Biomarkers

(A) Summary diagram of selected features from Boruta algorithm, including 40 microbial species, 64 KEGG Orthology (KO) genes, and 24 metabolites. (B) Top 10 microbial species selected based on feature importance. (C) Top 10 KO genes selected based on feature importance. (D) Top 10 metabolites selected based on feature importance.

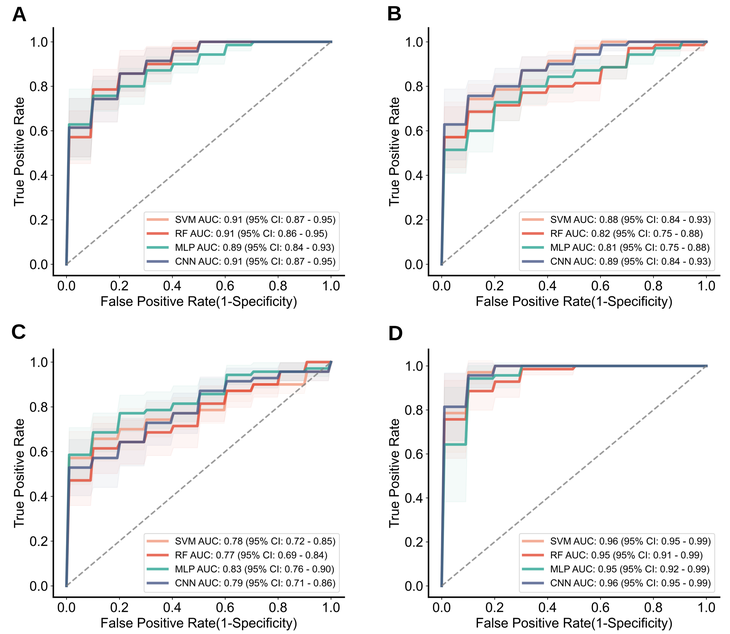

Fig 3. AUROC-Based Model Performance Across Multi-Omics Datasets

(A) Performance of four machine learning models (SVM, Random Forest, MLP, and 1D CNN) trained on microbial species data. (B) Performance of the same models trained on KO gene data. (C) Performance on metabolite data. (D) Performance on the combined multi-omics dataset (species + KO genes + metabolites).

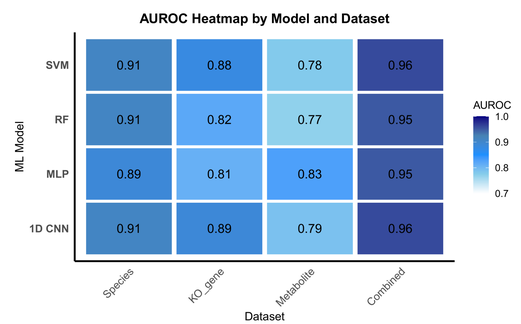

Fig 4. Comparative Heatmap of Model AUROC Values Across Single and Combined Omics Datasets

References

1. Ning, L., Zhou, Y. L., Sun, H., Zhang, Y., Shen, C., Wang, Z., ... & Hong, J. (2023). Microbiome and metabolome features in inflammatory bowel disease via multi-omics integration analyses across cohorts. *Nature Communications*, *14*(1), 7135.
2. Gao, W., Gao, X., Zhu, L., Gao, S., Sun, R., Feng, Z., ... & Jiao, N. (2023). Multimodal metagenomic analysis reveals microbial single nucleotide variants as superior biomarkers for early detection of colorectal cancer. *Gut Microbes*, *15*(2), 2245562.