

Identification of Potential Genomic Biomarkers in Pancreatic Cancer, Colon Cancer, and Ulcerative Colitis Using Integrated Machine Learning and Bioinformatics Approaches.

Author

Md Ibrahim Sarker Raiyan

Sakib Sarker

Utsha Das

Emon Ahammed

Abstract:

Pancreatic cancer (PC) and colon cancer (CC) are aggressive gastrointestinal malignancies often diagnosed late, contributing to high mortality. Ulcerative colitis (UC), a chronic inflammatory bowel disease, also reduces quality of life through persistent colon inflammation. These challenges highlight the need for genomic biomarkers to aid early diagnosis and prognosis. In this study, we applied an integrated machine learning and bioinformatics approach to identify shared genomic biomarkers across PC, CC, and UC. Differential expression analysis was performed on three microarray datasets, followed by LASSO regression as a feature selection method to refine disease-specific differentially expressed genes (DEGs). Common genes across all three diseases were used to construct a protein-protein interaction network (PPINet) via the STRING database. The network was analyzed in Cytoscape, and ten hub genes were identified using MCC, MNC, and Degree centrality metrics. Three genes-COL11A2, COL5A2, and COL11A1-emerged as common hub genes. Their diagnostic potential was validated using ROC analysis in independent test datasets, where all achieved high AUC scores. Additionally, mRNA expression analysis in PC samples showed significant upregulation of COL11A2 and COL11A1 in tumor tissue. These genes may serve as valuable biomarkers for early detection, prognosis, and therapeutic targeting in PC, CC, and UC.

2025 IEEE International Conference on Quantum Photonics, Artificial Intelligence, and Networking, QPAIN 2025

Conference Paper2025

EID: 2-s2.0-105019049263

DOI: 10.1109/QPAIN66474.2025.11171969

ISBN:979-8-3315-9695-8