

Title: Identification of Biomarkers and Molecular Pathways Implicated in Smoking and COVID-19 Associated Lung Cancer Using Bioinformatics and Machine Learning Approaches

Author: Md Ali Hossain, Mohammad Zahidur Rahman, Touhid Bhuiyan, Mohammad Ali Moni

Abstract: Lung cancer (LC) is a significant global health issue, with smoking as the most common cause. Recent epidemiological studies have suggested that individuals who smoke are more susceptible to COVID-19. In this study, we aimed to investigate the influence of smoking and COVID-19 on LC using bioinformatics and machine learning approaches. We compared the differentially expressed genes (DEGs) between LC, smoking, and COVID-19 datasets and identified 26 down-regulated and 37 up-regulated genes shared between LC and smoking, and 7 down-regulated and 6 up-regulated genes shared between LC and COVID-19. Integration of these datasets resulted in the identification of ten hub genes (SLC22A18, CHAC1, ROBO4, TEK, NOTCH4, CD24, CD34, SOX2, PITX2, and GMDS) from protein-protein interaction network analysis. The WGCNA R package was used to construct correlation network analyses for these shared genes, aiming to investigate the relationships among them. Furthermore, we also examined the correlation of these genes with patient outcomes through survival curve analyses. The gene ontology and pathway analyses were performed to find out the potential therapeutic targets for LC in smoking and COVID-19 patients. Moreover, machine learning algorithms were applied to the TCGA RNAseq data of LC to assess the performance of these common genes and ten hub genes, demonstrating high performances. The identified hub genes and molecular pathways can be utilized for the development of potential therapeutic targets for smoking and COVID-19-associated LC.

Keywords: COVID-19; ROC curve; WGCNA; comorbidity; lung cancer; pathway analysis; protein-protein interaction; smoking; survival analysis.

DOI: <https://doi.org/10.3390/ijerph21111392>

<https://www.sciencedirect.com/science/article/pii/S2590291124003139>