Identification of Novel Biomarkers and enriched pathways involved in lung cancer using Statistical and Bioinformatics techniques

Nikita Singh[1], Mukesh Kumar[1], Atanu Bhattacharjee[2], Prashant Sonker[1], and Agni Saroj[1]

[1]Banaras Hindu University
[2]Tata Memorial Hospital

July 31, 2020

Abstract

Introduction: Lung cancer is the most important cause of cancer related deaths across the world. The aim of the study is to find key genes and enriched pathways associated with lung cancer using bioinformatics and statistical techniques, hence providing potential targets for the identification and treatment of the cancer. Methods: Differentially expressed genes (DEGs) data of 54674 genes based on stage, tumor and status of the lung cancer was taken from 66 patients of African American (AAs) origin. 2392 DEGs were found based on stage, 13502 DEGs were found based on tumor, 2927 DEGs were found based on status having p value ($p<0.05$). Results: Total 33 common DEGs were found from stage, tumor and status of lung cancer patients. Gene ontology (GO) and KEGG pathway enrichment analysis is performed and 49 significant pathways were obtained, out of which 10 pathways were found that were exclusively involved in lung cancer development. Protein-protein interaction (PPI) network analysis found 69 nodes and 324 edges and identified 10 hub genes based on their highest degrees. Additionally, module analysis of PPI found that 'Viral carcinogenesis', 'pathways in cancer', 'notch signaling pathway', 'AMPK signaling pathways' had close association with lung cancer. Conclusion: it is seen that these identified DEGs do not directly participate in growth of lung cancer but regulate other genes which play important role in growth of lung cancer. The key genes and enriched pathways identified can thus help in better identification and prediction of lung cancer.

Nikita Singh[1], Mukesh Kumar[2], Atanu Bhattacharjee[3], Agni Saroj[4], Prashant Kr. Sonker[4]

[1]Department of Bioinformatics, M.M.V., Banaras Hindu University, Varanasi, India

[2]Department of Statistics, M.M.V., Banaras Hindu University, Varanasi, India

[3]The Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Mumbai, India

[4]Department of Statistics, I.Sc., Banaras Hindu University, Varanasi, India

Corresponding Author:

Mukesh Kumar, PhD

Department of Statistics, MMV

Banaras Hindu University

Varanasi-221005

Email: mukesh.mmv@bhu.ac.in

Data availability statement

The data that support the finding of the study are available from the corresponding author upon reasonable request from author.

Conflicts of interest

The authors declare no conflict of interest.

Funding

The authors have not received any funding for research work.

Ethical approval statement

Not applicable.

Patients consent statement

Not applicable.

Hosted file

CSO Manuscript.docx available at https://authorea.com/users/347666/articles/473284-identification-of-novel-biomarkers-and-enriched-pathways-involved-in-lung-cancer-using-statistical-and-bioinformatics-techniques

Hosted file

CSO Figures.docx available at https://authorea.com/users/347666/articles/473284-identification-of-novel-biomarkers-and-enriched-pathways-involved-in-lung-cancer-using-statistical-and-bioinformatics-techniques

Hosted file

CSO Tables.docx available at https://authorea.com/users/347666/articles/473284-identification-of-novel-biomarkers-and-enriched-pathways-involved-in-lung-cancer-using-statistical-and-bioinformatics-techniques