ISSN 0974 - 9373

Vol. 26 No. 1 (2022) **J. Int. Acad. Phys. Sci.** pp. 65-78

Identification of Biomarkers and Enriched Pathways Involved in Lung Cancer using Statistical Techniques

Nikita Singh
Department of Bioinformatics, M.M.V., Banaras Hindu University, Varanasi, India
Email: singhnikita119@gmail.com

Mukesh Kumar*
Department of Statistics, M.M.V., Banaras Hindu University, Varanasi, India
*Corresponding Author Email: mukesh.mmv@bhu.ac.in

Atanu Bhattacharjee
Section of Biostatistics, Centre for Cancer Epidemiology, TMC Mumbai, India
Email: atanustat@gmail.com

Prashant Kumar Sonker and Agni Saroj
Department of Statistics, Banaras Hindu University, Varanasi, India
Email: prashteek12@gmail.com, agni.saroj4@bhu.ac.in

(Received February 23, 2022)

Abstract: The aim of the study is to find key genes and enriched pathways associated with lung cancer using bioinformatics and statistical techniques. A total of 54674 differentially expressed genes (DEGs) data genes based on clinical information of lung cancer was taken from 66 patients of African American (AAs) origin. The data was retrieved from https://www.ncbi.nlm.nih.gov/geo/. The accession number of data is GSE102287. The Protein-protein interaction (PPI) network, gene ontology (GO) and KEGG pathways were used to find association among DEGs. Total 33 common DEGs were found from stage, tumor and status of lung cancer patients. 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. 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 closed association with lung cancer. 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.

Keywords: Lung cancer, biomarker, gene ontology, survival analysis.

1. Introduction

Worldwide mortality from lung cancer growth expanded from 3.5 million in 1990 to 4.2 million in 2015[1] and it is assessed that there will be 2.1 million new lung cancer incidents and 1.8 million deaths in 2018, representing (18.4%) incidents of cancer-related mortality[2]. Lung cancer is a heterogeneous disease and various factors including hereditary transformations, ecological components and individual habits can add to cancer incident, evolution and metastasis[3]. According to histological disparity, lung cancer can be partitioned into non-small cell lung cancer (NSCLC) and small-cell lung cancer (SCLC), of which NSCLC represents roughly 85%, and 30% of SCLC cases can be named as lung squamous cell carcinoma[4]. It is reported that a number of genes and biological, cellular and molecular pathways take part in these processes. Hence, it is crucial to understand the important mechanisms that lead to the onset and development of lung cancer in order to produce diagnostic and therapeutic strategies. A past researches on gene expression profiling in cancer used microarray tools for examining oncology[5] however some of these studies have been directed on lung cancer with comparative analysis of the DEGs[6], and a very authentic biomarker profile refining cancerous tissues from normal ones remains to be discovered.

At present, there are different type of technology and software available for analysis of DEGs with the molecular behavior. Gene sequencing analysis is also an approach to find biological functioning of DEGs microarray data. The aim of this study is to find DEGs and related pathways for development of lung cancer and also identify possible genes biomarkers for identification and prospects of lung cancer. However, the DEGs are dependent on cancer stages, tumor grade, and status of patient. Statistical and bioinformatics techniques have been considered for analysis purpose. A data set (GSE102287) has been downloaded from GEO (https://www.ncbi.nlm.nih.gov/geo/), and was consequently analyzed to find the involved DEGs in lung cancer.

2. Materials and methods

2.1. Gene expression data: DEGs of mRNAs and miRNAs have been taken from 66 patients of AAs origin was used for analysis. A total of 54674 genes were screened, on the basis of stage (I or II), tumor (present or absent) and status (dead or alive). The DEGs data was based on stage, tumor and status of lung cancer patients.

Table I: Summary of cancer patients involved in study

Variable	Number (%)
Total number of patients	66
Age, median (range) (years)	60 (32-76)
Male sex	38(57.5%)
Female sex	28(42.4%)
Smoking history	41.47(62%)
Cancer stages	
Stage I	36(54.5%)
Stage II and III	30(45.4%)
Tumor	
Present	32(48.4%)
Absent	34(51.1%)
Status	
Dead	26(39.3%)
Alive	40(60.6%)

2.2. Statistical analysis*:* Student's t-test for difference of means assuming unequal variances was applied to test the datasets and two-tailed ($p < 0.05$) was considered statistically significant. We have categorized the genes on the basis of stage, tumor and status. This procedure is adopted to screen the DGEs data in Table II. Out of these, 2392 DEGs from stage of lung cancer, 13502 DEGs from tumor and 2979 DEGs from status were obtained. 33 common DEGs from stage, tumor and status of the lung cancer were found.

2.3. Heat map: Heat map is used to represent the level of expression genes with comparable samples. By using R software we have created heat maps to show gene expressions level for DEGs obtained based on stage 1 and stage 2. Thereafter the patients were classified as tumor present, tumor absent and dead and alive status. Now the gene expression levels are shown by yellow, orange and red colors with gene affy ID and patients ID along with x-axis and y-axis respectively.

2.4 GO term enrichment analysis: GO of these 205 DEGs were done using DAVID Database that is available at https://david.ncifcrf.gov/. GO is a major bioinformatics activity to combine the demonstration of gene and gene product attributes with all variety. The aim is to: (1) maintain and expand its restricted vocabulary of gene and gene product attributes; (2) interpret genes and gene products data; and (3) provide tools for easy access to all aspects of the data[7].

2.5. Establishment of PPI Network: Search Tool for Retrieval of the Interacting genes (STRING) online database is used for representation of PPI networks and available at https://string-db.org/.

Table II: Common DEGs based on Status, tumor and status of lung cancer

Gene Symbol	Mean	Standard Deviation	P-Value	Gene Description
NKTR	5.43	0.54	0.01	Natural killer cell triggering receptor
C12orf80	1.93	0.04	0.02	Chromosome 12 open reading frame 80
LOC101927406	2.57	0.06	0.04	Uncharacterized LOC101927406
ACAA2	3.63	0.38	0.04	Acetyl-CoA acyltransferase
GPX2	3.84	3.66	0.01	Glutathione peroxidase
PARN	9.39	0.35	0.02	Poly(A)-specific ribonuclease
PPP1R3C	8.42	1.77	0.01	Protein phosphatase 1 regulatory subunit 3C
S100P	9.99	2.88	0.03	S100 calcium binding protein P
HOXC5	2.04	0.05	0.03	Homeobox C5
P2RY1	4.20	0.88	0.00	Purinergic receptor P2Y1
KLF12	2.14	0.03	0.04	Kruppel like factor
TRANK1	7.65	0.64	0.03	Tetratricopeptide repeat and ankyrin repeat containing 1
FXYD5	10.92	0.56	0.01	FXYD domain containing ion transport regulator 5
TRIM5	2.88	0.40	0.02	Tripartite motif containing 5
CISH	8.38	0.95	0.04	Cytokine inducible SH2 containing protein
PCSK9	5.45	0.93	0.01	Proproteinconvertasesubtilisin/ kexin type 9
SLIT1-AS1	2.18	0.04	0.01	SLIT1 antisense RNA 1
TFB1M	4.24	0.41	0.00	Transcriptional factor B1, mitochondrial
SLX1A-SULT1A3	7.12	0.48	0.03	SLX1A-SULT1A3 readthrough (NMD Candidate)
CC2D2A	4.43	0.61	0.02	Coiled-coil and C2 containing domain 2A
LOC100507277	1.87	0.04	0.04	Uncharacterized LOC100507277
FAM118B	3.78	0.38	0.00	Family with sequence similarity 118 member B
KMT2A	6.03	0.67	0.02	Lysine methyltransferase 2A
LOC100130502	1.78	0.06	0.03	Uncharacterized LOC100130502

A frame work comprehension of cell function requires information of all practical relations between the expressed proteins. The STRING database is used to collect and combine this information and predicted PPI for a large number of organisms[8]. Investigating the predicted interaction networks can recommend new directions for future computational research and provide cross-species expectations to efficient associated mapping[9]. STRING database gave the list of most significantly enriched pathways by KEGG pathway analysis. In the PPI network, the nodes involved in pathways exclusively involved in lung cancer with various colors were highlighted. These pathways showed genes that were involved in the NSCLC with their false discovery rates.

Table III: Gene Ontology analysis of DEGs

S.No.	Category	Term	Gene Count	P-Value
1.	GO_BP_DIRECT	Regulation of receptor activity	2	8.5E-3
2.	GO_BP_DIRECT	Anterior/ posterior pattern specification	2	7.4E-2
3.	GO_MF_DIRECT	Poly(A) RNA binding	4	7.4E-2
4.	GO_MF_DIRECT	Protein binding	12	8.0E-2

BP-biological process; FAT-functional annotation tool; MF-molecular function

Table IV: List of over represented pathway involved in Lung cancer using KEGG pathway analysis in STRING

Pathway	Description	False Discovery Rate	Color Representation
hsa05203	Viral carcinogenesis	0.00071	Dark purple
hsa04152	AMPK signaling pathway	0.00096	Pale yellow
hsa03320	PPAR signaling pathway	0.0017	Cyan
hsa05200	Pathways in cancer	0.0042	Dark green
hsa4330	Notch signaling pathway	0.0059	Pink
hsa04151	PI3K-Akt signaling pathway	0.0099	Yellow
hsa05206	microRNAs in cancer	0.0113	Green
hsa04066	HIF-1 signaling pathway	0.0228	Dark blue
hsa00280	Valine, leucine and isoleucine degradation	0.0427	Purple
hsa04310	Wnt signaling pathway	0.0423	Red

2.6. Cytoscape: This is online open software platform for **representation** molecular communication networks and genetic pathways and **combine** these networks with annotations, gene expression profiles and other state of data and can be downloaded from https://cytoscape.org/. Cytoscape is used to provide a basic set of features for data integration, analysis, and representation. The string file was saved in *.tsv* format and was imported in Cytoscape software. Using the MCODE (molecular complex detection) plug-in of Cytoscape, top 3 modules of protein-protein interactions were visualized that are seen to be involved in the lung cancer. By using CYTOHUBBA plug-in of Cytoscape, we found top 10 hub genes which are highly involved in lung cancer.

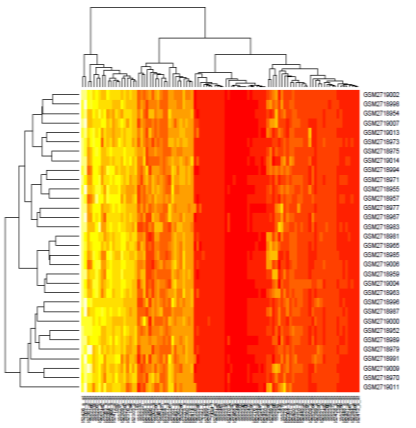
2.7. Kaplan Meier (KM) plot: Survival analysis is used to analysis of life time until one or more event happens. The KM curve is used to estimate the survival of patients from time dependent data. In medical sciences, it is often used to find the fraction of patients living for a certain time after treatment. Here, we have plotted the KM curve using R software for the stage-wise survival of lung cancer patients[23-26].

Table V: Top 10 Hub genes with their ranks and scores respectively

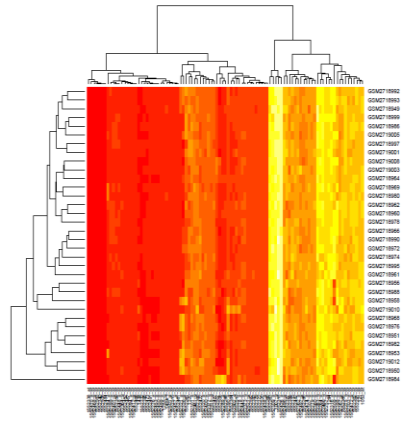
Rank	Name	Score
1	EP300 (E1A Binding Protein P300)	29
2	TP53 (Tumor protein 53)	25
3	KAT2B(K(*lysine) acetyltransferase 2B)*	24
4	HDAC1 (Histone Deacetylase 1)	22
5	SIRT1 (sirtuin 1)	17
6	KMT2A(*Lysine Methyltransferase 2A*)	16
6	ASH2L (histone lysine methyltransferasecomplex subunit)	16
6	SETD1B (SET Domain Containing 1B)	16
6	SETD1A (SET Domain Containing 1A)	16
10	KMT2C (*Lysine Methyltransferase 2C*)	15

3. Results

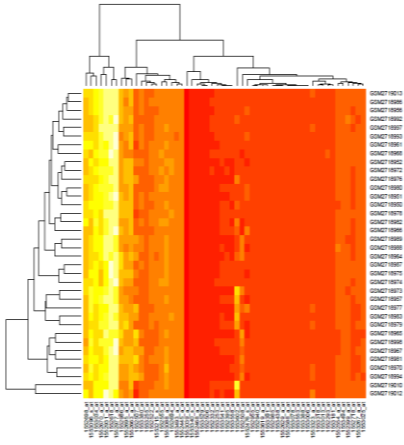
After applying student's t-test for unequal variances on 54647 genes with their gene expression values, we obtained 33 common DEGs. The selected genes had ($p<0.05$) in (Table II). The description of cancer patients is shown in (Table I). A total of 6 heat maps were plotted to show gene expressions level for DEGs obtained based on stage 1 and stage 2, tumor present and absent, and dead status and alive status.



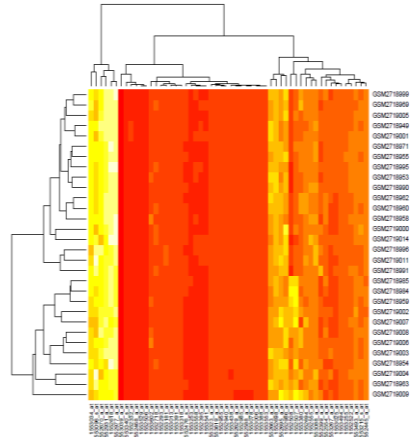
a) tumor present



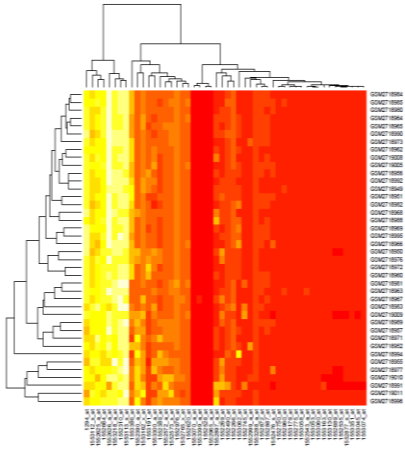
b) tumor absent



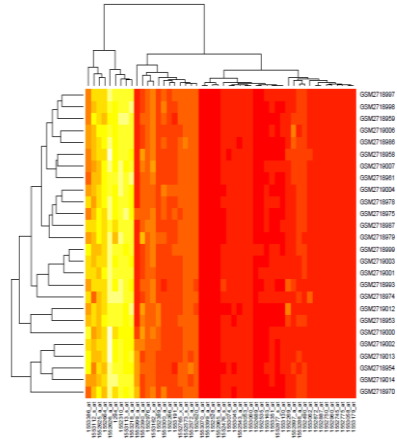
c) stage 1



d) stage 2



e) dead status



f) alive status

Figure 1: Heat maps showing the gene expression levels

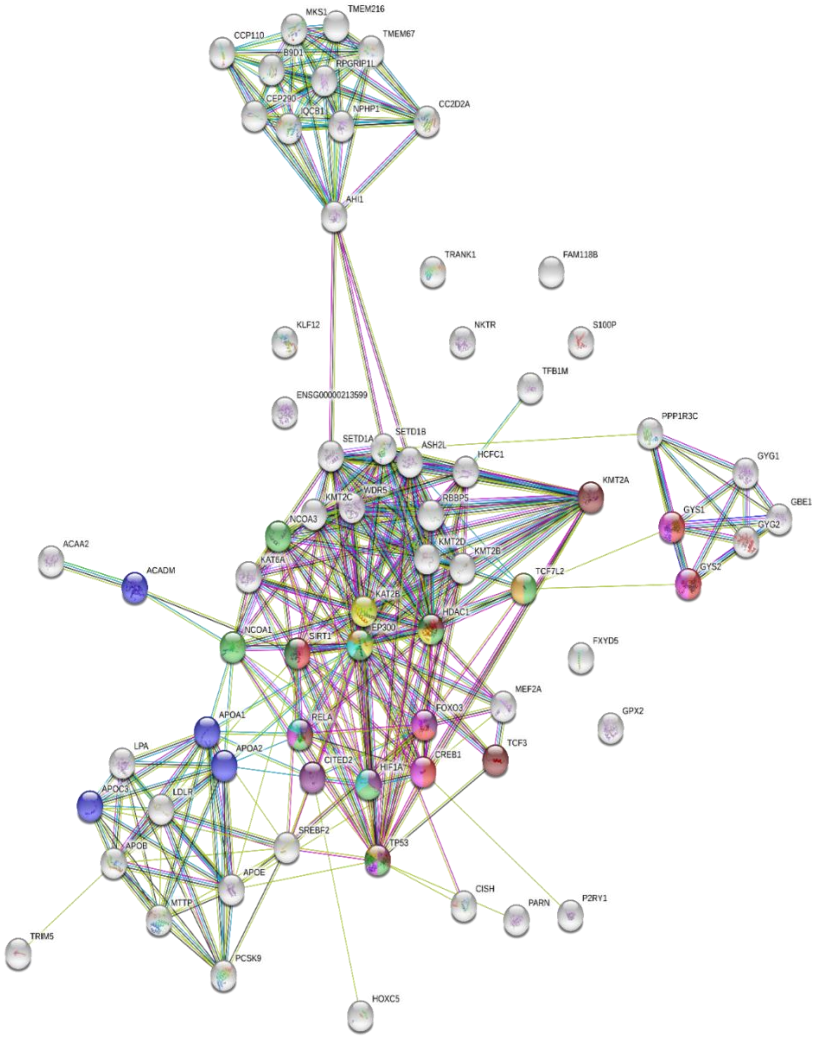
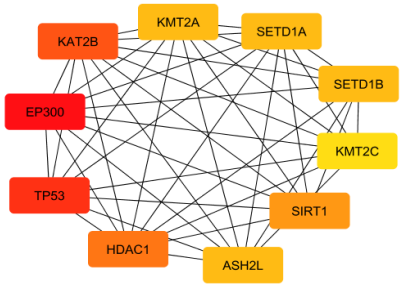
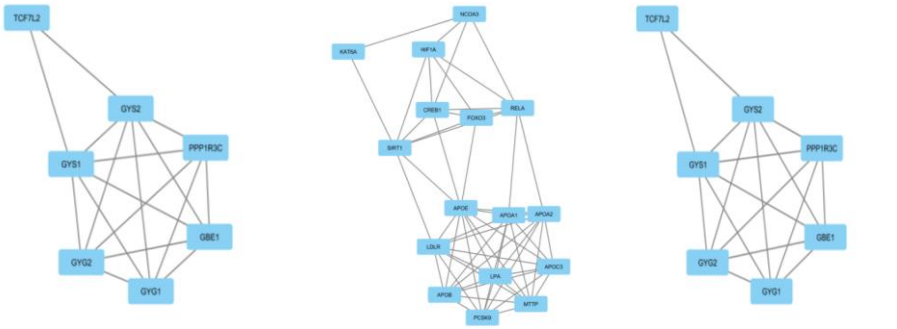


Figure 2: PPI network of DEGs identified by STRING



Rank	Node
1	EP300
2	TP53
3	KAT2B
4	HDAC1
5	SIRT1
6	KMT2A
6	ASH2L
6	SETD1B
6	SETD1A
10	KMT2C

Figure 3: Top 10 Hub Genes with highest degree of interaction in Lung cancer as analyzed by CYTOHUBBA plugin of Cytoscape



(A) (B) (C)

Figure 4: Top 3 modules of PPI networks. Nodes and links show human proteins and their PPI. A) Enriched pathway of module A; B) Enriched pathway of module B; C) Enriched pathway of module C

The yellow color in the heat map indicates lower values of gene expression values while the orange ones are intermediate and red ones have the highest values which had high values for the gene expression (Figure 1).To find the role of the DEGs, GO term enrichment analysis was performed with online database DAVID. The genes were significantly enriched in biological process (BP), molecular function (MF) and cellular component (CC) (Table III). The genes were enriched significantly in BP, including 'directive of receptor activity', 'Anterior/posterior pattern specification'. The genes enriched in MF, including 'Poly (A) RNA binding 'and 'Protein binding'. The over represented pathways using KEGG pathway analysis in String related with lung cancer were 'AMPK signaling pathway', 'PPAR signaling pathway', 'pathways in cancer', 'PI3K-Akt signaling pathway', 'notch signaling pathway', 'viral carcinogenesis', 'microRNAs in cancer', 'HIF-1 signaling pathway', 'Valine, leucine and isoleucine degradation' and 'Wnt signaling pathway' (Figure 2 and Table IV). The PPI network is constructed to classify the mainly important proteins and genetic modules that may serve critical roles in the growth of lung cancer. A total of 69 nodes and 324 edges were screened from PPI network (Figure 2). The average node degree was 9.39, the average local coefficient clustering was 0.694 and the PPI enrichment ($p < 0.01$). Each gene was entrusted a degree that predicted number of adjacent nodes in the network and changes in proteins/genes. The top 10 hub genes with the highest degrees in lung cancer were EP300 (E1A Binding Protein P300), TP53 (Tumor protein 53), KAT2B (Lysine Acetyltransferase 2B), HDAC1 (Histone Deacetylase 1), SIRT1 (sirtuin1), KMT2A (Lysine Methyltransferase 2A), ASH2L (Histone Lysine Methyltransferase complex subunit), SETD1B (SET Domain Containing 1B), SETD1A (SET Domain

Containing 1A), KMT2C (Lysine Methyltransferase 2C) (Table V and Figure 3). EP300 has highest degree of 29. It is found that high degree of these hub genes which play an important role in maintaining the entire PPI. In addition, to find the significance DEGs, the top 3 significant modules were selected and functional interpretation of genes related with the modules were analyzed (Figure 4 and Table VI). The results described that these modules had pathways that were seen to play a critical role in lung cancer. Module 1 was associated with viral carcinogenesis, pathways in cancer, notch signaling pathway, microRNAs in cancer, wnt signaling pathway. Module 2 was associated with AMPK signaling pathway, PPAR signaling pathway, PI3K-Akt signaling pathway, HIF-1 signaling pathway. Module 3 was associated with AMPK signaling pathway, pathways in cancer, wnt signaling pathway. It was found that out of 10 hub genes, only 4 were exclusively involved in lung cancer EP300, TP53, KMT2A and KMT2C. The 4 genes underwent mutations largely. KM plotted for the stage-wise survival curves of lung cancer AAs patients. Stage 3 clearly depicts the lowest rate of survival among all the 3 stages (Figure 5).

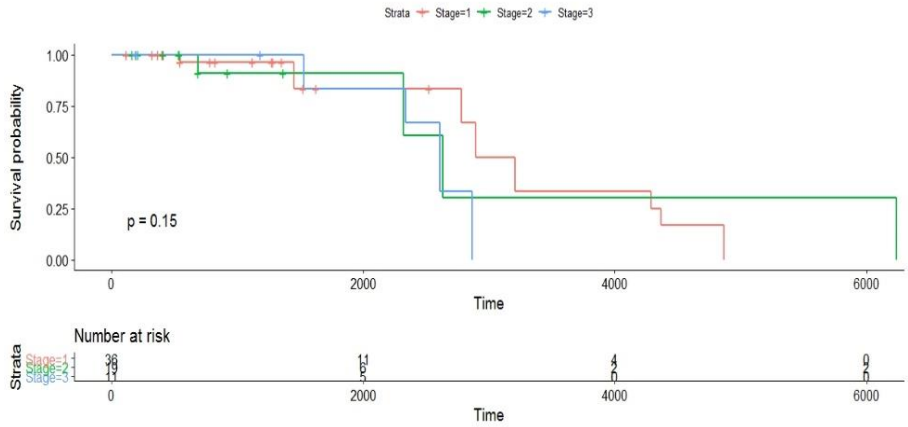


Figure 5: Stage wise survival of lung cancer patients among AAs.

4. Discussion

Cancer is basically a hereditary disease, and different hereditary changes collect during the multistep process of carcinogenesis, which finally leads to anomalous excessive cell development and malignant phenotype[10]. Lung cancer is basically essential pulmonary malignant tumor in terms of incidence and mortality[11]. Early identification and efficient treatment of lung cancer is need of the hour and it can be achieved by the identification of significant genes and understanding their molecular mechanisms which play an important role in causing lung cancer. DEGs data of various genes can be

used for further functional analysis and to screen biomarkers that can serve for early identification and remedial targets. Therefore, they may help in finding of lung cancer in the early stages and can be used for the development of targeted treatment. In present study statistical and bioinformatics methods are applied to identify new candidate genes that can serve critical roles in development of lung cancer. The data used here has gene expression values of 54674 genes for 66 patients, being categorized on the basis of stage, tumor and status of the lung cancer. A total of 33 common DEGs from stage, tumor and status were obtained based on their p-value score calculated by t-test for difference of means with unequal variances. Then, GO and KEGG pathway analyses are performed to find the associations of these significant genes. Finally, a PPI network was constructed that depicted that these identified DEGs directly do not play role in causing lung cancer, but they interact and regulate other neighboring genes that play a very important role in development of lung cancer (Figure 2).

GO analysis is helpful for annotating genes and gene products. GO analysis in the present study showed that these significant genes involved in biological process like 'Regulation of receptor activity', Anterior/posterior pattern specification', molecular functions like 'Poly (A) RNA binding' and 'Protein binding'. It is observed that defective functioning of biological processes and body system status are important causes of tumor growth and evolution. Hence, monitoring the expression of these genes may help in discovery of tumor mechanisms. The KEGG pathway database carries methodical analysis of gene functions, linking genomics and the functional information. Enrichment analysis is used to find important and most significant KEGG pathways which are related with lung cancer and its growth were 'AMPK signaling pathway', 'PPAR signaling pathway', 'pathways in cancer', 'PI3K-Akt signaling pathway', 'notch signaling pathway', 'viral carcinogenesis', 'microRNAs in cancer', 'HIF-1 signaling pathway', 'Valine, leucine and isoleucine degradation' and ' Wnt signaling pathway' (Figure 2 and Table IV). Taking pathways into consideration, AMPK plays a central role in the control of cell growth, prevalence and autotrophic through the rule of mTOR activity, which is consistently uncontrolled in cancer cells. Targeting of AMPK/mTOR is thus strategy in the growth of remedial elements against NSCLC[12]. The PI3K pathway is frequently uncontrolled in lung[13]. Cancer due to hereditary variation affecting its components resulting in increased PI3K signaling PPAR-γ factor bring development and promote changes related with separation as well as apoptosis in different lung carcinoma cell lines[14]. Thus, defects in PPAR signaling pathway can promote tumor growth. In case of notch signaling pathway and Dang et al. found that

the over expression of Notch3 was perceived in 40% of patients with NSCLC, and that this over expression was connected with a translocation including 19p[15]. In HIF-1 signaling pathway, Hypoxia-inducible factor-1α (HIF-1α) is over expressed in human lung diseases, especially in NSCLC, and is firmly related with a propelled tumor grade, expanded angiogenesis, and protection from chemotherapy and radiotherapy[16]. In case of Wnt signaling pathway, over expression of Wnt-1, -2, -3, and -5a and of Wnt-pathway components Frizzled-8, Disheveled, Porcupine, and TCF-4 is common in NSCLC and is associated with poor prognosis[17]. p53 is the most frequently mutated gene in lung cancer[18]. Most clinical studies suggest that NSCLC with TP53 alterations carries a worse prognosis and may be relatively more resistant to chemotherapy and radiation[19]. Inactivation of TP53 capacity or its orderly pathway is a typical component of human tumors that regularly relates with expanded danger, poor patient survival, and protection from treatment[20-22].

It is observed that many genes though not in our 33 common DEGs, comes into picture because it is regulated by genes present in our initial DEGs list such as PPP1R3C, ACAA2, TRIM5, PCSK9, P2RY1, CISH, PARN and KMT2A (Figure 2). Hence, it is clearly seen that the 33 DEGs do not directly participate in development of lung cancer but some of them influence and regulate other genes which play key role in development of lung cancer. PPP13RC is predicted functional partner of GYS1 AND GYS2. ACAA2 is neighbor of ACADM. TRIM5 and PCSK9 are in a cluster network of APOA1 and APOA2. P2Y1 is connected to CREB1. CISH is connected to two most crucial genes TP53 and EP300. PARN is found associated with TP53. KMT2A is the gene with high no. of degree among our 33 DEGs. It is connected to CREB1, EP300, TP53, HDAC1 and SIRT1.

The string file is imported in cytoscape software and using CYTOHUBBA plug-in, top 10 hub genes based on their degree was found. The gene with highest score was EP300, followed by TP53 and KAT2B (Figure 3 and Table V). These 10 hub genes played important role in growth of lung cancer. Using MCODE plug-in of cytoscape, top 3 modules of this network were seen which were again observed to take part in pathways that caused lung cancer (Figure 4 and Table VI). The oncoprint and cancer summary type study is done by cBioPortal of TCGA database shows that TP53 is most mutated gene among all the top 10 hub genes. Also among the 10 hub genes, only 4 genes are exclusively involved in lung cancer viz. EP300, TP53, KMT2C and KMT2A. Cancer type summary is depicted in Figure 5. The survival analysis was done and KM Plot was plotted which demonstrated that Stage 3 clearly has the lowest rate of survival among all the 3 stages.

4. Conclusions

This study made us to reach on a conclusion that DEGs may directly be involved in the pathways that lead to the development of cancer or may sometimes be indirectly involved like influencing and regulating other genes and their pathways that may play a crucial role in development of a tumor or a cancer.

Acknowledgement: Author would like to thank the chief editor and refereed reviewer for their valuable comments and suggestions to improve the article.

Conflict of interest: The authors declare that there is no conflict of interest.

Funding: The authors did not receive any funding for this work.

References

1. A. J. Cohen, M. Brauer, R. Burnett, H. R. Anderson, J. Frostad, K. Estep and V. Feigin; Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, *The Lancet*, 389(10082) (2017) 1907-1918.
2. F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre and A. Jemal; Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA: a cancer journal for clinicians*, 68(6) (2018) 394-424.
3. V. Aran, A. P. Victorino, L. C. Thuler and C. G. Ferreira; Colorectal cancer: epidemiology, disease mechanisms and interventions to reduce onset and mortality, *Clinical colorectal cancer*, 15(3) (2016) 195-203.
4. B. Piperdi, A. Merla and R. Perez-Soler; Targeting angiogenesis in squamous non-small cell lung cancer, *Drugs*, 74(4) (2014) 403-413.
5. F. Descotes, P. Dessen, P. P. Bringuier, M. Decaussin, P. M. Martin, M. Adams and M. Devonec; Microarray gene expression profiling and analysis of bladder cancer supports the sub-classification of T 1 tumours into T 1a and T 1b stages, *BJU international*, 113(2) (2014) 333-342.
6. H. Sakashita, H. Inoue, S. Akamine, T. Ishida, N. Inase, K. Shirao and K. Mimori; Identification of the NEDD4L gene as a prognostic marker by integrated microarray analysis of copy number and gene expression profiling in non-small cell lung cancer, *Annals of surgical oncology*, 20(3) (2013) 590-598.
7. M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry and M.A. Harris; Gene ontology: tool for the unification of biology, *Nature genetics*, 25(1) (2000) 25-29.
8. D. Szklarczyk, J. H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic and L. J. Jensen; The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible, *Nucleic acids research*, 45 (2016) D362-D368.
9. A.S. Schwartz, J. Yu, K. R. Gardenour, R. L. Finley Jr and T. Ideker; Cost-effective strategies for completing the interactome, Nature methods, 6(1) (2009) 55-61.

10. S. L. Carter, A. C. Eklund, I. S. Kohane, L. N. Harris amd Z. Szallasi; A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers, *Nature genetics*, 38(9) (2006) 1043-1048.
11. C. S. D. Cruz, L. T. Tanoue and R. A. Matthay; Lung cancer: epidemiology, etiology, and prevention, *Clinics in chest medicine*, 32(4) (2011) 605-644.
12. D. Han, S. J. Li, Y. T. Zhu, L. Liu and M. X. Li; LKB1/AMPK/mTOR signaling pathway in non-small-cell lung cancer, *Asian Pac J Cancer Prev*, 14(7) (2013) 4033-4039.
13. V. Duronio; The life of a cell: apoptosis regulation by the PI3K/PKB pathway, *Biochemical Journal*, 415(3) (2008) 333-344.
14. S. Han and J. Roman; Peroxisome proliferator-activated receptor γ: a novel target for cancer therapeutics, *Anti-cancer drugs*, 18(3) (2007) 237-244.
15. T.P. Dang, A.F. Gazdar, A.K. Virmani, T. Sepetavec, K.R. Hande, J.D. Minna D.P. Carbone; Chromosome 19 translocation, overexpression of Notch3, and human lung cancer, *Journal of the National Cancer Institute*, 92(16) (2000) 1355-1357.
16. Q. Zhang, X. Tang, Z. F. Zhang, R. Velikina, S. Shi and A. D. Le; Nicotine induces hypoxia-inducible factor-1α expression in human lung cancer cells via nicotinic acetylcholine receptor–mediated signaling pathways, Cl*inical Cancer Research*, 13(16) (2007) 4686-4694.
17. D. J. Stewart; Wnt signaling pathway in non–small cell lung cancer, *JNCI: Journal of the National Cancer Institute*, 106(1) (2014) 1-11.
18. M. Olivier, A. Petitjean, V. Marcel, A. Petre, M. Mounawar, A. Plymoth and P. Hainaut, Recent advances in p53 research: an interdisciplinary perspective, *Cancer gene therapy*, 16(1) (2009) 1-12.
19. A. Mogi and H. Kuwano; TP53 mutations in nonsmall cell lung cancer, *BioMed Research International*, 583929 (2011) 1-9.
20. A. J. Munro, S. Lain and D. P. Lane; P53 abnormalities and outcomes in colorectal cancer: a systematic review, *British journal of cancer*, 92(3) (2005) 434-444.
21. B. Vogelstein, D. Lane and A. J. Levine; Surfing the p53 network, Nature, 408(6810) (2000) 307-310.
22. K. H. Vousden and X. Lu; Live or let die: the cell's response to p53, *Nature Reviews Cancer*, 2(8) (2002) 594-604.
23. M. Kumar, P. K. Sonker, A. Saroj, A. Jain, A. Bhattacharjee and R. K. Saroj; Parametric survival analysis using R: Illustration with lung cancer data, *Cancer Reports*, 3(4) (2019) 1-6.
24. N. Tripathi, S. Keshari, P. Shahi, P. Maurya, A. Bhattacharjee, K. Gupta and M. Kumar; Human papillomavirus elevated genetic biomarker signature by statistical algorithm, *Journal of Cellular Physiology*, 135(9) (2020) 1-11.
25. R. K. Saroj, K. N. Murthy, M. Kumar, A. Bhattacharjee and K. K. Patel; Bayesian competing risk analysis: An application to nasopharyngeal carcinoma patients data, *Computational and Systems Oncology*, 1(1) (2021) 1-11.
26. M. Rajput, M. Kumar, M. Kumari, A. Bhattacharjee and A. A. Awasthi; Identification of key genes and construction of regulatory network for the progression of cervical cancer, *Gene Reports*, 21 (2020) 1-10.