Nath, R, Baishya, S, Nath, D, Nahar, L, Sarker, SD, Choudhury, MD and Talukdar, AD

Identifying druggable targets from active constituents of Azadirachta indica A. Juss. for non-small cell lung cancer using network pharmacology and validation through molecular docking.

Research Article

Identifying druggable targets from active constituents of *Azadirachta indica* A. Juss. for non-small cell lung cancer using network pharmacology and validation through molecular docking

Rajat Nath[1], Somorita Baishya[1], Deepa Nath[2], Lutfun Nahar[3*], Satyajit D. Sarker[4], Manabendra Dutta Choudhury[1], Anupam Das Talukdar[1*]

[1]*Department of Life Science and Bioinformatics, Assam University, Silchar, Cachar, Assam, India, 788011*

[2]*Department of Botany, Guru Charan College, Silchar, Cachar, Assam, India, 788004*

[3]*Laboratory of Growth Regulators, Palacký University and Institute of Experimental Botany, The Czech Academy of Sciences, Šlechtitelů 27, 78371 Olomouc, Czech Republic*

[4]*Centre for Natural Products Discovery (CNPD), School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, James Parsons Building, Byrom Street, Liverpool L3 3AF, United Kingdom*

***Corresponding authors:**

Anupam Das Talukdar

Department of Life Science and Bioinformatics, Assam University
Silchar, Assam, India, 788011
Phone: +91 9401416452

Lutfun Nahar
Laboratory of Growth Regulators, Palacký University and Institute of Experimental Botany, The Czech Academy of Sciences
Šlechtitelů 27, 78371 Olomouc, Czech Republic
Phone: +420 585634850

***Correspondence e-mails:** anupam@bioinfoaus.ac.in (A.D.T.); nahar@ueb.cas.cz (L.N.)

Abstract:

Introduction: *Azadirachta indica* A. Juss. is a well-known medicinal plant that has been used traditionally to cure various ailments in every corner of the globe. There are many *in vitro* and *in vivo* experimental evidences in connection with the bioactivity of this plant extracts. Lung cancer is the deadliest form of cancer and contributes to the most cancer related deaths. The mode of action of anticancer components of this plant is still to establish explicitly.

Objective: The objective of this study is to identify druggable targets of active constituents of *Azadirachta indica* A. Juss. for non-small cell lung cancer (NSCLC) using network pharmacology and validation of activity through molecular docking analysis.

Methodology: Targets of all the active phytochemicals from *A. indica* were predicted and genes related to NSCLC were retrieved. PPI network of the overlapping genes were prepared. Various databases and server were employed to analyze the disease pathway enrichment analysis of the clustered genes. Validation of the gene/protein activity was achieved by performing molecular docking, and ADMET profiling of selected phytocompounds was performed.

Result: Gene networking revealed three key target genes as EGFR,BRAF and PIK3CA against NSCLC by the active components of *A. indica*. Molecular docking and ADMET analysis further validated that desacetylnimbin, nimbandiol, nimbin, nimbinene, nimbolide, salannin and vepinin are the best suited anti NSCLC among all the phytocompounds present in this plants.

Conclusion: The present study has provided a better understanding of pharmacological effects of active components from *A. indica* and its potential therapeutic effect on NSCLC.

Keywords: *Azadirachta indica, Network pharmacology, Molecular docking, Non-small cell lung cancer, phytocompounds*

1. Introduction

Azadirachta indica (Meliaceae), commonly known as “Neem”, is an evergreen tree, native to tropical and subtropical regions such as Bangladesh, Burma, India, Nepal and Pakistan.[1,2] Many ailments, including infections, metabolic disorders and cancer, have been claimed to be treated with phytochemicals extracted from various part of this tree, and this plant is included in the Ayurveda, Unani, and homoeopathy systems of medicine.[3] In recent years, studies have demonstrated that *A. indica* constituents possess several biological and pharmacological properties including anti-inflammatory, antibacterial, analgesic, antipyretic, anti-arrhythmic, antihistamine, anti-arthritic, antiprotozoal, antifungal, diuretic, anti-ulcer, antitubercular, antimalarial, spermicidal, insect repellent, antifeedant, and antihormonal activities.[4-6] This *in silico* study was designed to understand the mechanism of anti-non-small cell lung cancer (NSCLC) activity of the active components present in this plant.

Cancer is a deadly disease, which killed 10 million people just in the year 2020 according to the WHO report (WHO factsheets, 3 February 2022; https://www.who.int/news-room/fact-sheets). The number is increasing rapidly day by day. Among all cancers, lung cancer causes the highest number of cancer related deaths and is the most predominantly diagnosed cancer amid patients.[7] Lung cancer causes massive haemorrhage, systemic air embolism, bronchopleural fistula leading to intractable pneumothorax, pneumonitis, and pulmonary artery pseudoaneurysm. It also damages nearby lung tissue, and leads to empyema, needle tract seeding, skin burns, and lung abscesses. Surgical removal of the tumour is both the first and best treatment option for lung cancer. However, this procedure is not suitable for all types of patients, like those with multiple health problems, advanced stages of cancer, or lungs that compromise its function. Chemotherapy, radiotherapy, or a combination of the two are other treatments for these patients, but the success rate is low.[8] So, to find new alternative treatments for NSCLC is a major goal for cancer researchers. In this study, *A. indica* was selected to understand its underlying molecular mechanism against NSCLC by the network pharmacology, which is a new drug discovery approach developed by Hopkins in 2007 that combines systematic medicine and information science. It highlights the concept of "network target, multicomponent

treatments," which shifts the paradigm from one gene, one target, and one disease. Network pharmacology is a powerful method for studying traditional medicine's synergistic actions and underlying mechanisms[9].

2. Methodology

2.1. Screening of active components and targets of *Azadirachta indica*

Active components present in *Azadirachta indica* were retrieved from the Dr. Duke's Phytochemical and Ethnobotanical Databases (https://phytochem.nal.usda.gov/phytochem/search) and literature mining. In the databases, the search keyword was used as "*Azadirachta indica*".

For literature mining, PubMed (https://pubmed.ncbi.nlm.nih.gov/), Google Scholar (https://scholar.google.com) and Science Direct (https://www.sciencedirect.com/) databases were used. In all the cases, keyword used as "phytocompounds of *Azadirachta indica*", "metabolites and *Azadirachta indica*", "secondary metabolites from *Azadirachta indica*", "bioactive compounds from *Azadirachta indica*".

The chemical structures of identified compounds were retrieved for target prediction. PubChem (https://pubchem.ncbi.nlm.nih.gov/) and ChemSpider (https://www.chemspider.com/) were employed to get the structure in canonical smile format. Swiss Target Prediction (https://www.swisstargetprediction.ch/) and search tool for interactions of chemicals (STICH) (https://stitch.embl.de/) databases were used to predict the probable target of the compounds. During the search, the selected species was *Homo sapiens*. The target was predicted using structural similarity and a reverse pharmacophore matching approach, and a target with a high probability was chosen for further experiments.

2.2. Collection of target genes for non-small cell lung cancer

Several database were employed for collection of target genes for non-small cell carcinoma, they were merged and any duplicates were removed. In all the platforms, keyword used was "non-small cell lung cancer" and species was selected as *Homo sapience*. Gene Cards (https://www.genecards.org/) is a robust

searchable database that compiles human gene data from a variety of sources, Drug Bank (https://go.drugbank.com/) a knowledge base for drug interactions, pharmacology, chemical structures, targets, metabolism, & more., NCBI Genes (https://www.ncbi.nlm.nih.gov/gene/) which is a powerful database by The National Center for Biotechnology Information (NCBI), part of the United States National Library of Medicine (NLM) provide gene integrates information from a wide range of species, The Online Mendelian Inheritance in Man (OMIM) (https://omim.org/search/advanced/) database is regularly updated with new information about the correlation between genetic and phenotypic features in humans and DiGeNET (https://www.disgenet.org/) a discovery platform containing one of the largest publicly available collections of genes and variants associated to human diseases were used for retrievals of data.

2.3. Acquisition of overlapping targets of the and NSCLC

The targets predicted from the active compounds of *A. indica* and targets related to NSCLC were put as two separate data set to Veny v.2.1, an online mapping tool for creation of venn diagram.[10] The overlapping targets were identified and proceeds for the further analysis.

2.4. Protein-protein interaction (PPI) network construction

Compounds and NSCLC overlap targets were deemed hub genes, and their PPI was obtained by online STRING v.11.5 (https://string-db.org/) analysis, with the species set to "*Homo sapiens*" and a confidence score of >0.990.[11] The output TSV file from STRING database was generated following evidences of various sources and was imported into Cytoscape v3.9.1 for further study. The network was built and visualized using Cytoscape v3.9.1, a programme frequently used in network pharmacology studies. For more advanced network analysis, it also provides a standard set of characters for data integration, analysis, and display. In a network, the most significant nodes are defined by the degree and betweenness centralities of their corresponding edges; larger numerical values for these topological metrics indicate greater significance.[12]

2.5. Gene clustering and network analysis

The Molecular Complex Detection (MCODE) tool in cytoscape v3.9.1 was used to extract the highly interaction regions (Clusters) from the massive gene network. The MCODE score, which takes into account both the number of edges and the distance between them, is used to rank the clusters.[13] For the cluster analysis with MCODE, the whole network was selected, where as other parameters was set as default; degree cutoff set at 2, node score cut of 0.2, k-score 2 and max. depth was set at 100.

2.6. Disease pathway analysis

To identify the most influential gene from all the generated clusters, disease pathway analysis was done using different server and algorithm. ClueGO v2.5.9, a cytoscape plugin, ShinyGO v0.76.3 (http://bioinformatics.sdstate.edu/go/) and Database for Annotation, Visualization and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/) were employed to perform this analysis. The disease was limited related to human in the search filter. ClueGo were used for human disease pathway analysis, ShinyGO predicted kyoto encyclopedia of genes and genomes (KEGG) and OMIM pathway and DAVID used for DisGeNET human disease pathway analysis.

2.7. Molecular docking

In computer aided drug discovery (CADD) approaches, molecular docking is one of the most important tool to understand the nature and strength of chemical bonding between the protein molecule and target ligands. The visual interface gives the information regarding the types and position of the bonds and the scoring system indicates the bond strength. The negative scores implies the spontaneous binding whereas positive score means external energy is required for formation of bond. Molegro virtual docker (MVD) v6.0 were used to perform the molecular docking analysis. Protein molecule were downloaded from protein data bank (https://www.rcsb.org/) in .pdb format and imported in to the workspace of MVD removing the water molecules and bonded ligands. In built protein preparation plugin of the software were used to correct the protonation state of the protein. The active site of the protein was detected using the detect cavity plugin. Details of the used target proteins are tubulated in the **Table 1**. All the ligands were then imported to the

workspace in mol format. Docking was performed replicating the run 10 times for each ligands and considering the moldock score, top 1 pose for each ligands were retrieved.[7]

Table 1: *Receptor information retrieved from the RCSB PDB database (www.rcsb.org/pdb).*

Sl No.	Receptor	PDB Id	Mutation	R-free	Resolution	Experimental methods
1	EGFR	1xkk	No mutation	0.255	2.40 Å	XRD
2	EGFR	6lub	L858R/T790M/C797S	2.31 Å	0.226	XRD
3	B-Raf	4r5y	V600E	3.50 Å	0.306	XRD
4	PIK3CA	7l1c	Mutant	1.96 Å	0.205	XRD

2.8. ADMET prediction and drug likeliness analysis

The ADME (Adsorption, Distribution, Metabolism, and Excretion) profiles of the screened compounds by molecular docking were calculated using SwissADME (http://www.swissadme.ch/), a web server from Swiss Institute of Bioinformatics. Canonical smile formats of the active screened compounds were imported to the server which generated the pharmacokinetics, drug likeliness and boiled egg models for the compounds.[14]

3. Result and Discussion

3.1. Active components and targets of *Azadirachta indica*

From the literature and phytochemical databases, a total of 79 phytocompounds were retrieved, whereas structure was available for 53 compounds. Target fishing enabled the prediction of targets effective against NSCLC related targets. Remaining 26 compounds were avoided in this study due to the lack of their structural elucidation. All the compounds and their predicted targets were given in **Supplementary Table: 1**. The efficacy of the compounds fished by such *in silico* methods can be considered to be complementary to experimental findings as these methods employ structure–activity relationship (SAR) data to predict bioactivity of the phytochemicals.

3.2. NSCLC related targets and identification of overlapping targets

From the aforementioned database, targets related to human NSCLC were retrieved, merged and removed duplicates. A total of 19970 target genes were found. The Venny v2.1.0 differentiated the overlapping targets of active components of *A. indica* and NSCLC. Generation of venn diagrams helped in easy dissection of the targets into identifiable groups relevant to our study. 724 common targets from both the lists were found. These targets were assumed to have roles to play in causing human NSCLC (**Supplementary file: 2**) **(Figure: 1**). These 724 common targets genes were used for further studies, which are somehow linked with the human NSCLC.

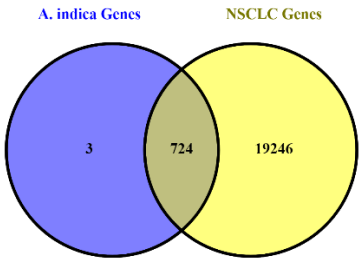

Figure 1: *Differentiation of overlapping target genes of active compounds of A. indica and targets related to human NSCLC.*

3.3. Construction of compound target network and clustering

The protein-protein interaction (PPI) of overlapping 724 genes were constructed by STRING database. The database generated interaction based on various types of evidences such as text mining, experimental results, annotated pathways etc. At the highest confidence score (>0.990), network was formed with 724 nodes, 3523 edges with average node degree of 9.73, indicating that these genes have highest likelihood of interaction. This NSCLC network belongs to the targets of active components from *A. indica.* PPI network generated from STRING was exported to Cytoscape for visualization and further analysis. MCODE app utilized for clustering showed 19 highly interactive clusters (gene regions) from the network. These 19 clusters are composed of total 235 genes. As genes interact in clusters to materialize a biological process, therefore, identification of the highly influential cluster is imperative in order to predict the most potent target causing this disease.

3.4. Disease enrichment pathway analysis of clustered genes

To compare and validate results, the disease enrichment analysis with 4 different methodology and server was carried out. First of all, clueGO enrichment app available at Cytoscape was used. Analysis with this software revealed that out of these 235 genes, 4 genes are highly linked with neoplasm of lung and lung cancer. These four genes are EGFR, BRAF, ERBB2 and PIK3CA, belongs to cluster 4, cluster 14, cluster 6 and cluster 2 respectively. Other than this, more 7 genes are also related to another 6 different types of human ailments (**Figure: 2**).

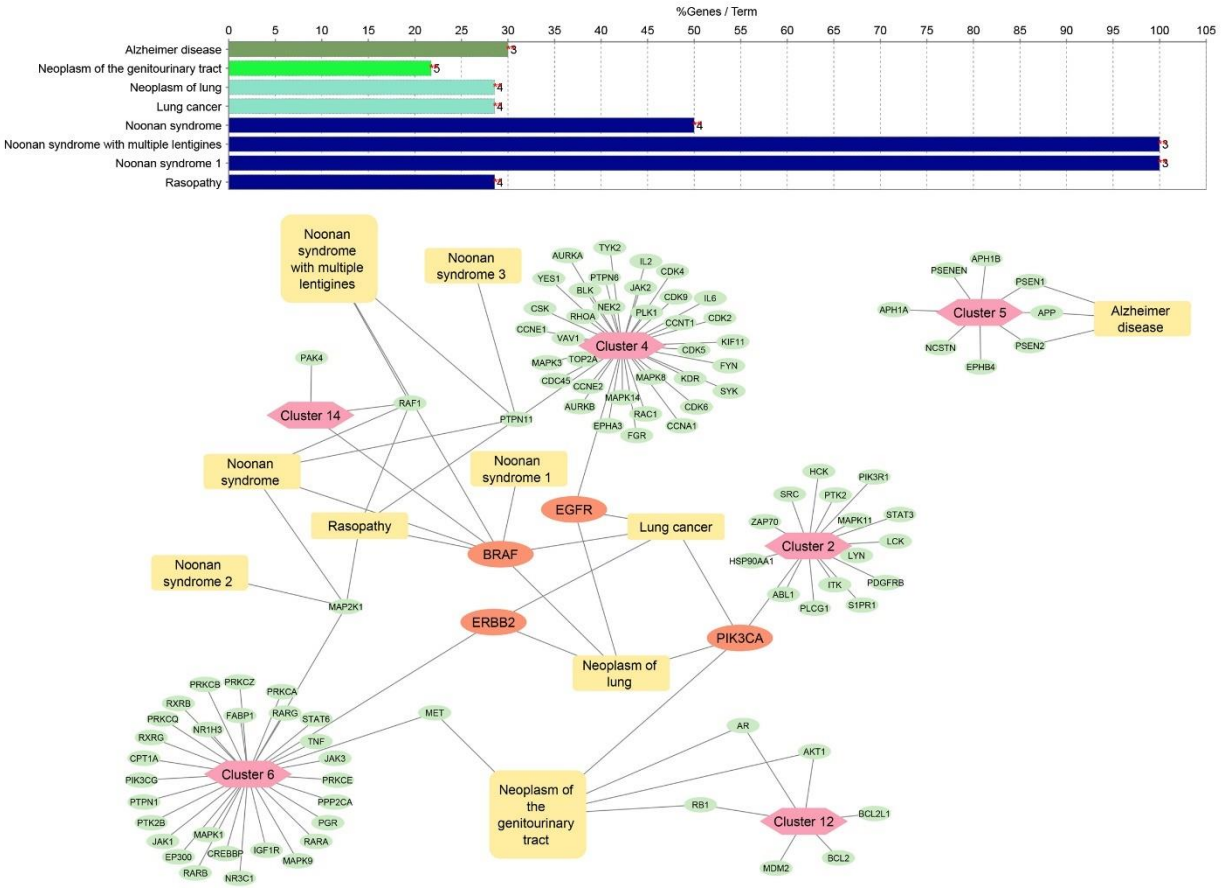

Figure 2: *Disease gene cluster pathway network by ClueGO enrichment analysis. 4 target genes were identified associated with human lung cancer and neoplasm of lung.*

OMIM and KEGG disease pathway were analyzed by using ShinyGO web server. According to OMIM analysis data, out of 235 genes, a subset is responsible for nine distinct human diseases, one of which is lung cancer. Three of the genes, EGFR, BRAF, and PIK3CA, were shown to be associated with lung cancer, and they were found to be located in clusters 4, 14, and 2 respectively (**Figure: 3**). KEGG shows 20 human disease pathway whereas non-small cell lung cancer stood for second position contributing by 21 responsible genes belonging to cluster 2, cluster 4, cluster 6, cluster 12 and cluster 14 (**Figure: 4**).

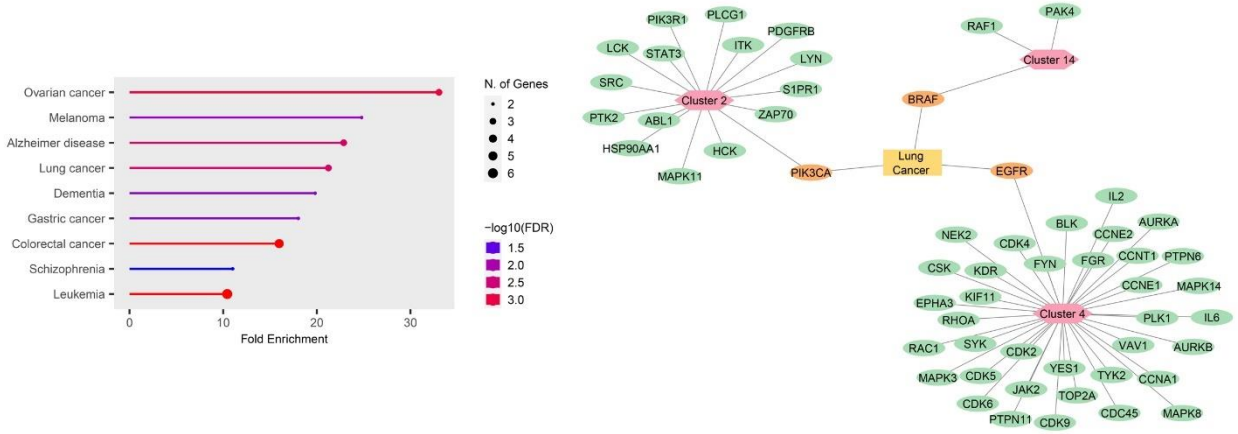

Figure 3: *Disease gene cluster pathway network by OMIM enrichment analysis. Three genes namely PIK3CA, BRAF and EGFR found to be associated with human lung carcinoma and these three genes are belongs to cluster 2, 14 and 4 respectively.*

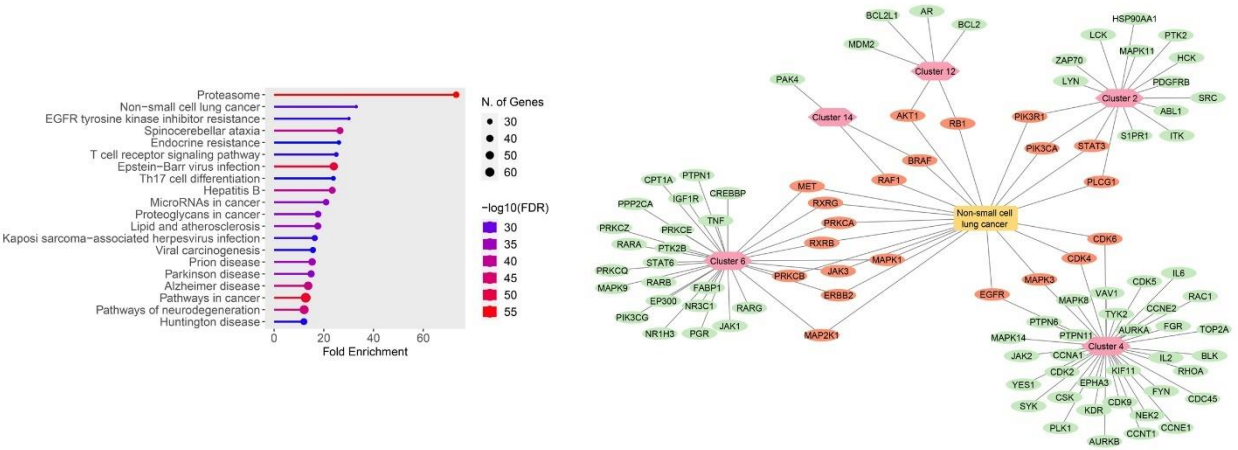

Figure 4: *Disease gene cluster pathway network by KEGG enrichment analysis. Twenty different human disease are found to be associated with the gene set whereas non-small cell lung carcinoma ranked 2 comprising 21 genes.*

Lastly, DAVID enrichment software was used, where DisGeNET identified 15 genes from cluster 2, cluster 4, cluster 6, cluster 12, cluster 13 and cluster 14 are responsible for non-small cell lung cancer (**Figure: 5**).

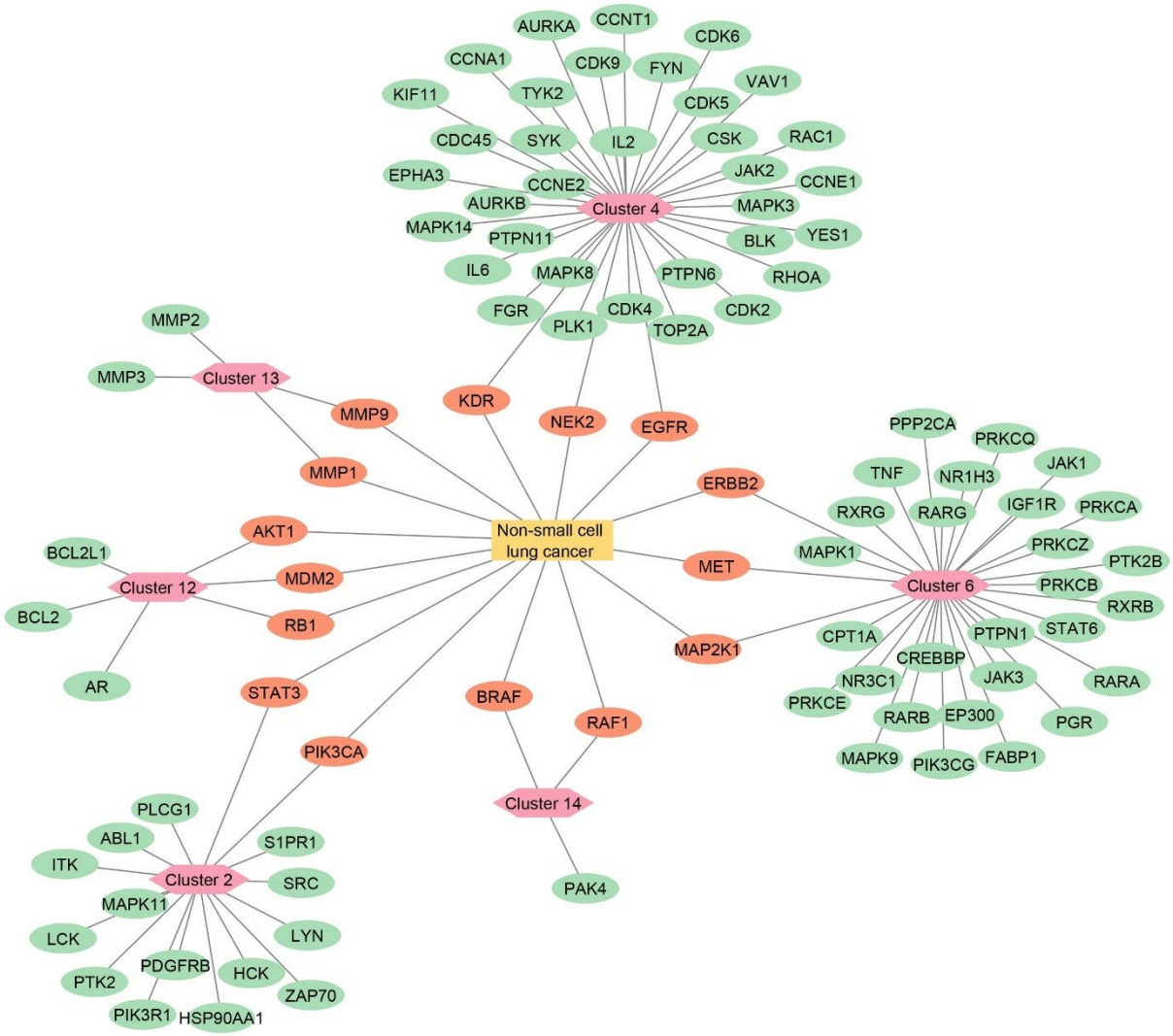

Figure 5: *Disease gene cluster pathway network by DisGeNET enrichment analysis. From the six different genes cluster, 15 genes are found associated with human non-small cell lung cancer.*

After all of these analyses, an extensive panoramic view of the interaction pattern was generated. Upon merging the result of aforesaid four methods, three genes, EGFR, BRAF and PIK3CA were found to be most common among all (**Figures: 6 and 7**).

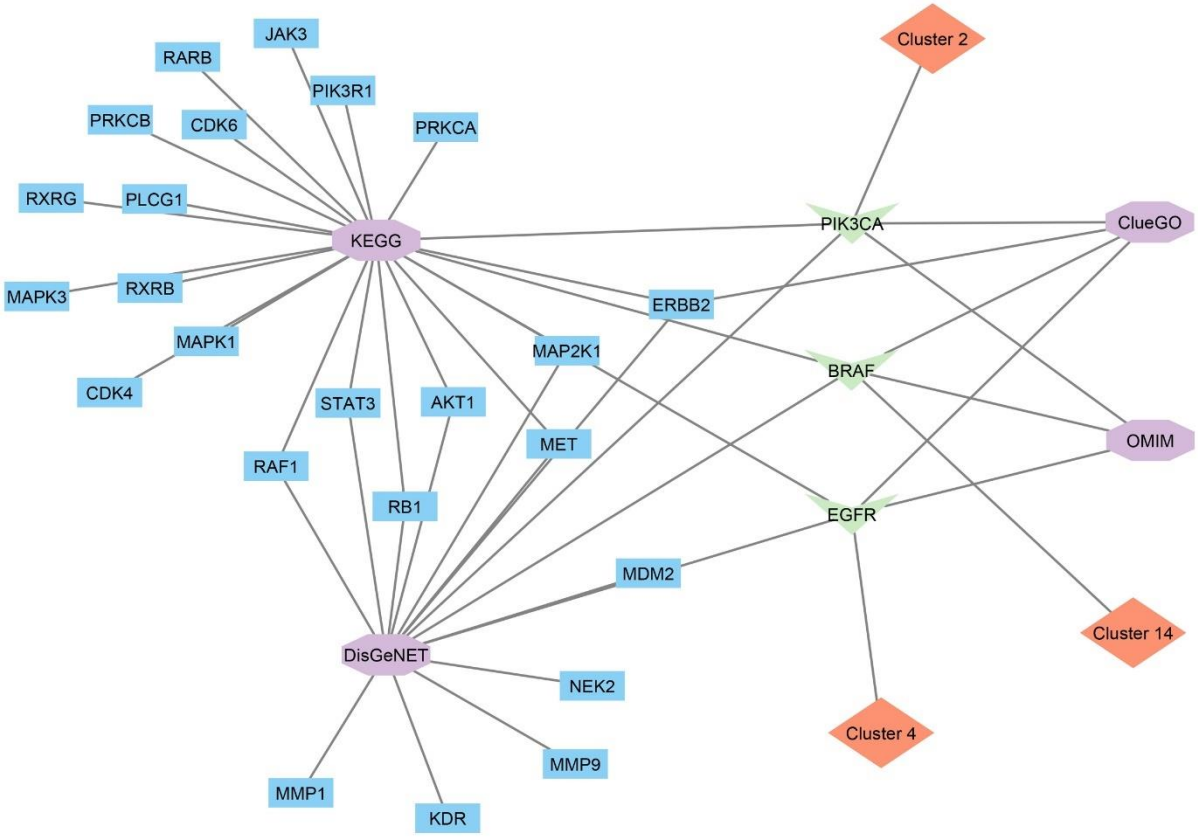

Figure 6*: Common Target genes from different disease pathway and network clusters. This analysis revealed PIK3CA, BRAF and EGFR are the most important target genes for the therapeutic use of non-small cell lung cancer.*

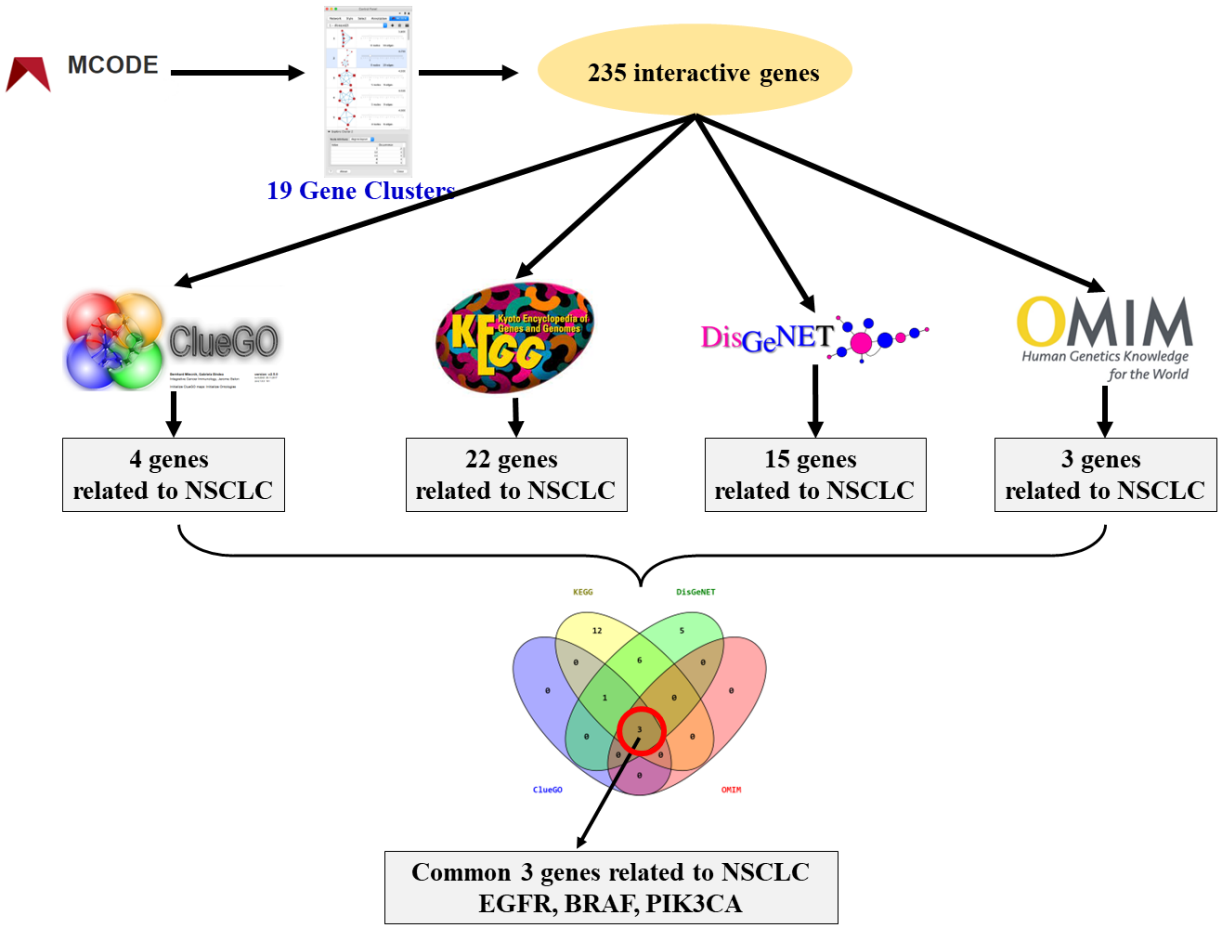

Figure 7: *The overall disease pathway analysis of four different database and finding of most influence genes associated with human non-small cell lung carcinoma.*

The epidermal growth factor receptor (EGFR) belongs to the family of proteins called ErbB receptors. The tyrosine kinase, which is located inside the cell, is responsible for signal transduction, whereas ligand binding occurs on the extracellular region of the receptor.[15] The extracellular receptor is activated by dimerization following the binding of a ligand, such as epidermal growth factor (EGF) or transforming growth factor-a (TGF-a). The signal travels through a number of downstream routes. The STAT3/STAT5 pathway, the PI3K/AKT pathway, and the RAS/RAF/MEK/MAPK route appear to be the most significant.[16] These pathways modify gene transcription and the cell cycle, leading in enhanced cell proliferation and angiogenesis, suppression of apoptosis, and changes in their ability to migrate, adhere, and invade. Changes in these factors and properties, as expected, have an effect on the development of the

cancer. The EGFR gene has been found to contain several mutations. Some of these mutations improve tyrosine kinase activity and so sensitivity to targeted therapy, whereas others increase cell resistance to treatment with tyrosine kinase inhibitors (TKIs).[15]

Mutant epidermal growth factor receptor (EGFR) tumours constitute a distinct subtype of NSCLC.[17] This EGFR mutation leads to cancer cell proliferation.[18] Adenocarcinomas, particularly those in younger women and girls who have never smoked, are characterized by mutations in the EGFR gene.[19,20]. Gefitinib is used as a first line drug for treatment of naive patients with activating EGFR mutations, but did not improve overall survival in T790M-mutant patients.[21]

BRAF is a serine/threonine protein kinase that operates as a downstream molecule of KRAS and is activated upon phosphorylation in a GTP dependent way, controlling important cell processes such as survival and proliferation by boosting the MEK/MAPK cascade.[22,23] Mutations that turn on the BRAF gene trigger constitutive MEK/ERK signaling, which in turn facilitates uncontrolled cell proliferation. About 30% of human malignancies are caused by BRAF mutations, which were first discovered in malignant melanomas. Around 50% of lung cancers have BRAF mutations, and these alterations might be V600E or non-V600E.[24,25]

In recent years, the PIK3CA gene has been considered as a potential driver gene in lung squamous cell cancer.[26-29] Wang et. al.[30], reported smoking may be associated with PIK3CA high expression in NSCLC patients, and PIK3CA mutation may alter lymph node metastasis and serve as a promising prognostic factor. Human cancer patients have been reported to carry somatic mutations in the PIK3CA (phosphatidylinositol 3-kinase catalytic subunit) gene. The PIK3CA gene has been shown to be mutated in roughly 4% of lung tumours, according to some studies.[31] The catalytic p110 subunit of PI3K is encoded by three genes (PIK3CA, PIK3CB, and PIK3CD), while the regulatory p85 subunit is encoded by only one gene (PIK3R1). The PIK3CA mutation is the most common in malignancies.[32] Some research has indicated that PIK3CA mutation occurs in a wide range of human cancers, with rates as high as 2-7% in NSCLC.[33,34] Moreover,

the PIK3CA gene may have a role in the expansion and development of NSCLC tumour cells and has been identified as a possible driver gene of lung squamous cell carcinoma.[26-29]

The detailed pathway of these three genes are illustrated in **Figure 8.**

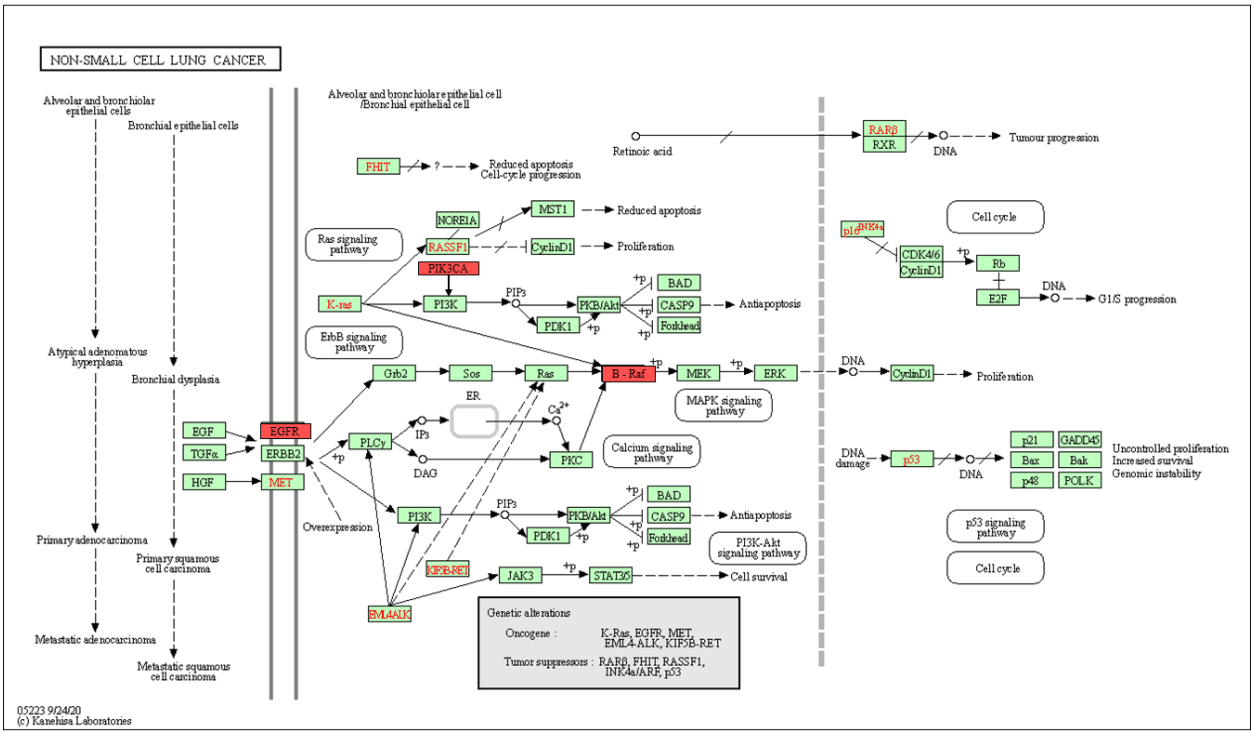

Figure 8: *The detailed pathway for Non-small cell lung cancer and the associated genes. Red boxes are the target genes identified by the network pharmacological approach using disease pathway analysis.*

3.5. Molecular Docking:

Molecular docking of active compounds from *A. indica* with the identified 3 targets viz. EGFR, B-Raf (encoded by BRAF gene) and PIK3CA. As the mutated form of all the proteins are responsible for NSCLC, therefore all the protein structure were downloaded in their mutated form. Additionally normal EGFR protein also selected for this docking study as the normal EGFR protein associated with the said type of cancer. This docking analysis revealed that ten compound from this plants namely salannolide, deacetylsalannin, isonimbinolide, nimbolide, nimbin, salannin, nimbinene, nimbocinolide, nimbandiol and vepinin exhibited better inhibition in all the four cases than their marketed drugs (**Figure 9**). Apart from

this three more compounds viz. 7-deacetyl-7-benzoylepoxyazadiradione, desacetylnimbin and 17-epiazadiradione showed activity against EGFR and PIK3CA. Normal and mutated EGFR were also inhibited by two more active compounds, including rutin and azadiradione (**Table 2**). If the top one compound is considered against each protein, salannolide showed the best inhibition against normal EGFR and B-Raf targets with a moldock score of -177.16 and -169.35, respectively, whereas positive control (marketed drug) gefitinib displayed moldock score of -127.63 and dabrafenib showed -137.31. Similarly, against PIK3CA and mutated EGFR, best inhibition was shown by salannin with a score of -178.13 and -161.54, respectively, while the positive control alpelisib exhibited -134.59 and gefitinib showed -107.35. The entire moldock score of the active compounds against these targets are tabulated in **Table 2.** Docking pose of best ligands with target receptors are shown in **Figure 10** and interaction of positive controls with the target molecules are shown in **Figure 11**.

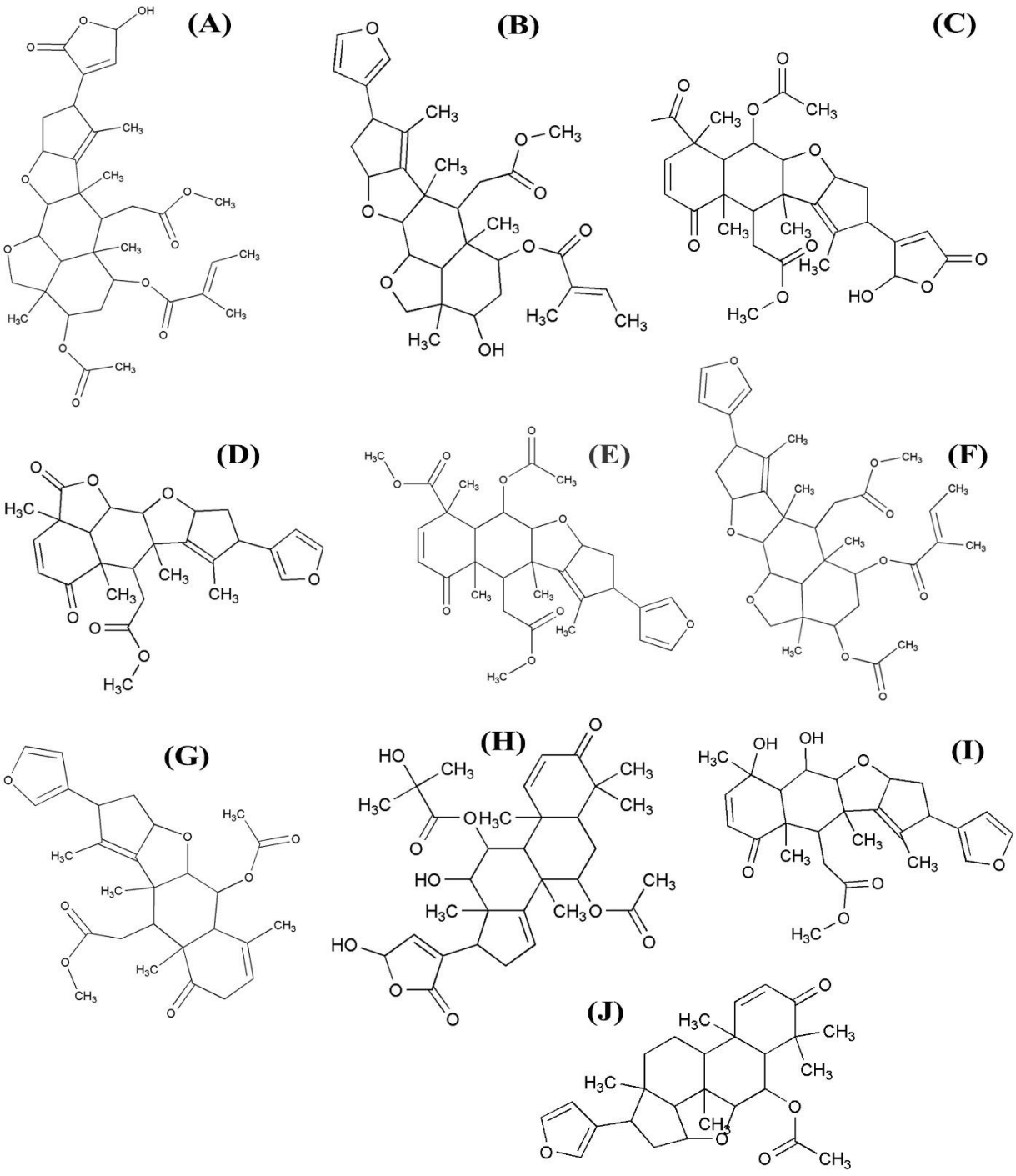

Figure 9: *Structure (A) salannolide, (B) deacetylsalannin, (C) isonimbinolide, (D) nimbolide, (E) nimbin, (F) salannin, (G) nimbinene, (H) nimbocinolide, (I) nimbandiol and (J) vepinin .*

Table 2: *MolDock score of the ligands for their respective target proteins.*

Normal EGFR		Mutant EGFR		Mutant PIK3CA		Mutated B-Raf	
Ligand	**Moldock Score**	**Ligand**	**Moldock Score**	**Ligand**	**Moldock Score**	**Ligand**	**Moldock Score**
Gefitinib (Positive Control)	-127.63	Gefitinib (Positive Control)	-107.35	Alpelisib (Positive Control)	-123.51	Dabrafenib (Positive Control)	-137.31
Salannolide	-177.16	Salannin	-161.54	Salannin	-178.65	Salannolide	-169.35
Salannin	-168.95	Isonimbinolide	-145.61	Deacetylsalannin	-161.94	Deacetylsalannin	-163.84
Deacetylsalannin	-150.43	Salannolide	-141.77	Salannolide	-160.76	Isonimbinolide	-161.56
Isonimbinolide	-149.19	Deacetylsalannin	-135.77	Rutin	-155.26	Nimbolide	-155.21
Nimbinene	-147.41	Nimbolide	-133.51	Nimbinene	-147.68	Nimbin	-153.79
Rutin	-144.44	Nimbinene	-130.20	Nimbocinolide	-146.98	Salannin	-153.52
Nimbocinolide	-143.26	Nimbocinolide	-128.36	Nimbolide	-144.86	Nimbinene	-151.53
Vepinin	-142.50	Vepinin	-125.88	Isonimbinolide	-143.49	Nimbocinolide	-143.71
Nimbaflavone	-141.85	Nimolicinol	-124.30	17-Epiazadiradione	-140.66	Nimbandiol	-142.50
Nimbandiol	-137.51	Rutin	-123.53	Nimbin	-140.09	Vepinin	-141.03
Nimbosterol	-136.46	Nimbandiol	-120.98	Nimbandiol	-139.68	Nimocinol	-138.05
Nimbolide	-135.84	7-Deacetyl-7-benzoylepoxyazadiradione	-119.44	7-Deacetyl-7-benzoylepoxyazadiradione	-137.40		
7-Deacetyl-7-benzoylepoxyazadiradione	-135.72	Desacetylnimbin	-115.15	Azadiradione	-130.22		
Azadiradione	-135.29	Nimbaflavone	-114.91	Vepinin	-129.51		
Nimocinol	-131.90	Nimbin	-114.79	Linoleic acid	-126.75		
Nimbin	-131.78	Nimbosterol	-114.78	Nimolicinol	-126.58		
Desacetylnimbin	-131.095	Linoleic-Acid	-109.98	Nimocinol	-125.44		
Epoxyazadiradione	-130.056	17-Epiazadiradione	-109.72	Meldenin	-123.58		
Hyperoside	-129.818	Oleic-Acid	-108.44				
Azadirone	-129.709						
17-Epiazadiradione	-129.342						
Nimbinin	-127.79						

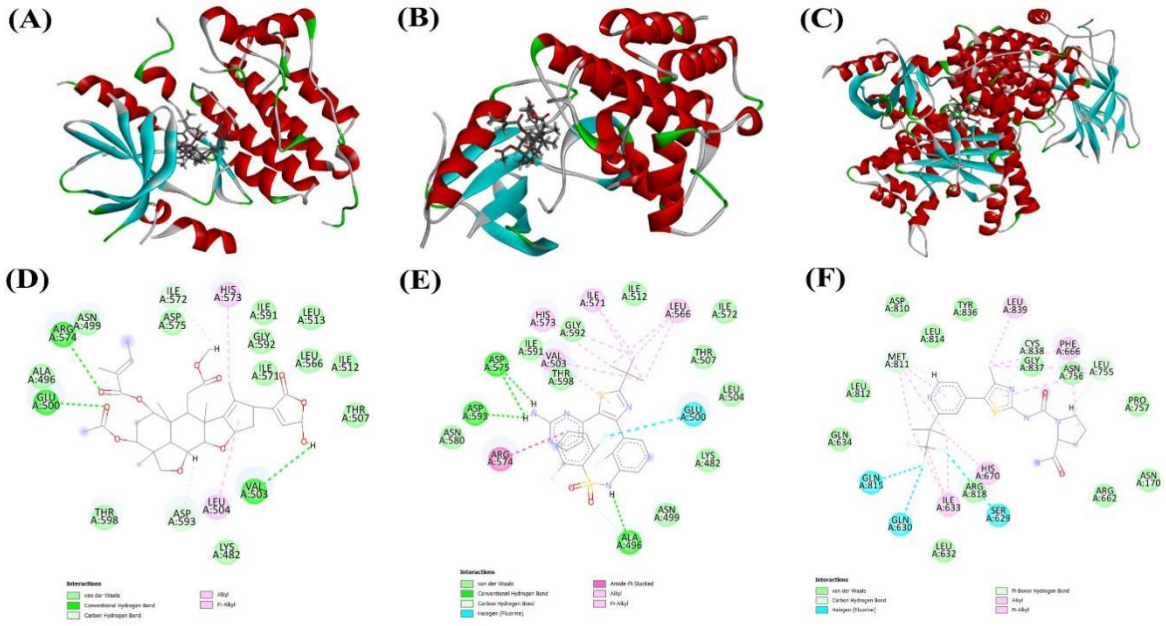

Figure 10: *Docking pose of ligand with the target receptors. 3D pose of (A) EGFR with Salannolide; (B) B-Raf with Salannolide; (C) PIK3CA with Salannin; 2D pose of (D) EGFR with Salannolide (E) B-Raf with Salannolide; (F) PIK3CA with Salannin;*

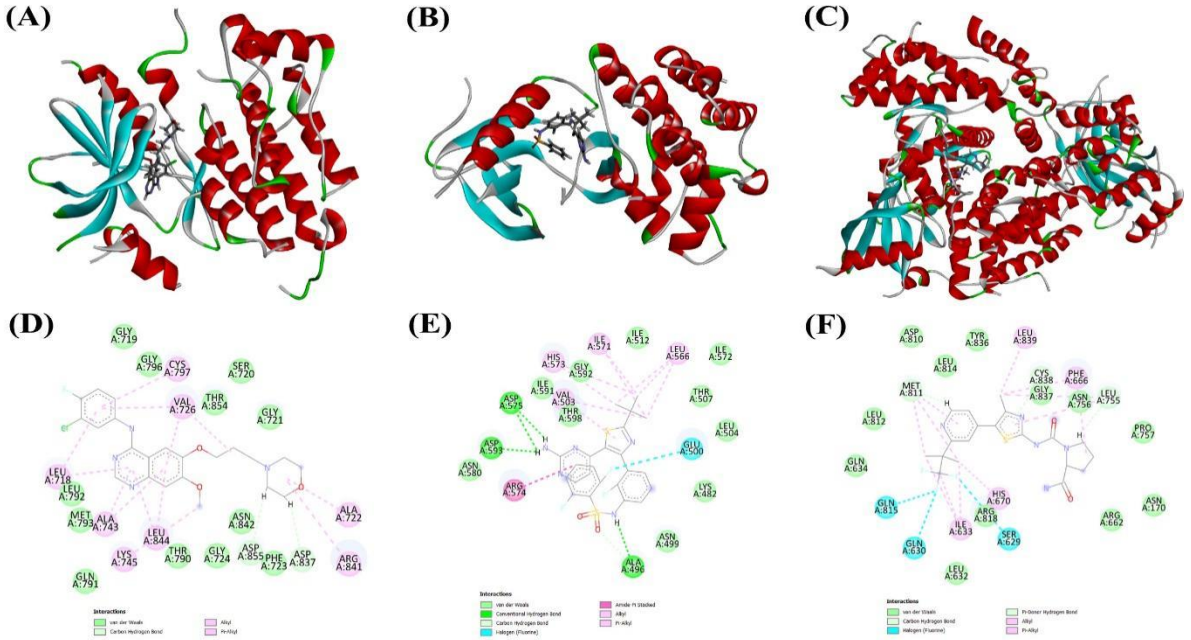

Figure 11: *Docking pose of positive control with the target receptors. 3D pose of (A) EGFR with Gefitinib; (B) B-Raf with Dabrafenib; (C) PIK3CA with Alpelisib; 2D pose of (D) EGFR with Gefitinib acid (E) B-Raf with Dabrafenib; (F) PIK3CA with Alpelisib;*

Protein-ligand molecular docking is an effective and crucial method in drug development, especially for characterizing the drug-likeness and mechanism of action of new compounds.[7] Inhibition of protein EGFR and B-Raf is associated with the inhibition of uncontrolled cell proliferations. Studies have revealed that, the L858R/T790M/C797S mutation in EGFR are associated with NSCLC. 50% of lung cancer cases are associated with T790M mutations. L858R is responsible for upregulation of the CXCR4 chemokine receptor and malignant pleural effusion (MPE) formation. C797S mutation is linked with tyrosine kinase inhibitors, it is found that almost 40% of the 3rd generation cancer drugs treated patients often get C797S mutations.[35-37] Similar studies has done by the Saini et. al.,[38] and shows that targeting EGFR can cure the NSCLC.

The common type of BRAF mutation is BRAF V600E which is responsible for NSCLC.[39] Targeted therapy of B-Raf protein is a key therapeutic strategy for NSCLC.[40,41] Somatic mutations in the PIK3CA gene have been found in high frequency in numerous cancer types, including colon, brain, stomach, breast, and ovary. More than 75% of these mutations are found in the PIK3CA gene's helical (exon 9) and kinase domains (exon 20). Mutations in PIK3CA's three mutation hotspots (E542K, E545K, and H1047R) have been demonstrated to increase its lipid kinase activity and activate the downstream Akt-signaling pathway. PIK3CA mutations were found 11.4% of lung cancer cases.[42-44] PIK3CA mutations are also commonly coexist with EGFR/KRAS mutations in NSCLC cases.[45] Studies has found that the treatment with the inhibitors to mutated PIK3CA in murine model shows the NSCLC protections .

3.6. ADME profile and drug likeliness analysis:

Considering the 10 compounds screened by the molecular docking analysis, uploaded for ADME and drug likeliness analysis by SwissADME server. The pharmacokinetics parameter used for the studies were selected Gastro Intestinal (GI) absorption, Blood Brain Barrier (BBB) permeability and P-Glycoprotein (P-GP) substrate. GI absorption is an important parameter to understand the drug delivery concept; a drug having high GI absorption is considered that it can be consumed orally.[46] BBB permeability is also an

crucial parameter for a drug candidates. If a drug crosses the BBB, may interact with the Central Nervous System (CNS) which may occurs ill effect in the brain. Additionally, P-GP substate is responsible for efflux out the drug from the BBB, if somehow permeable the barrier.[47] Other than the isonimbinolide, nimbocinolide, salannin and salannolide, all the rest seven compounds showed high GI absorption. All the eleven compounds were non-permeable to BBB, and excluding desacetylnimbin, isonimbinolide, nimbin and nimbinene, all displayed P-GP substrate activities (**Table 3**). The same result also obtained by the boiled egg analysis by the same server. The yolk part of the egg represents the BBB and white part represents the Gastro Intestinal absorption (**Figure 12**).

Table 3: *Calculated pharmacokinetics and drug-likeness parameters of the selected phytocompounds from A. indica.*

Compounds Name	**Pharmacokinetics**			**Drug likeness**					
	GI absorption	BBB permeant	P-GP substrate	Lipinski (Pfizer)	Ghose (Amgen)	Veber (GSK)	Egan (Pharmacia)	Muegge (Bayer)	Bioavailability Score
Desacetylnimbin	High	No	No	Yes	No (1)	Yes	yes	Yes	0.55
Isonimbinolide	Low	No	No	No (2)	No (3)	No (1)	No (1)	No (2)	0.17
Nimbandiol	High	No	Yes	Yes	Yes	Yes	Yes	Yes	0.55
Nimbin	High	No	No	Yes	No (3)	Yes	yes	Yes	0.55
Nimbinene	High	No	No	Yes	No (1)	Yes	yes	Yes	0.55
Nimbocinolide	Low	No	Yes	Yes	No (3)	No (1)	No (1)	No (1)	0.55
Nimbolide	High	No	Yes	Yes	Yes	Yes	yes	Yes	0.55
Salannin	Low	No	Yes	Yes	No (3)	Yes	yes	Yes	0.55
Salannolide	Low	No	Yes	No (2)	No (3)	No (1)	No (1)	No (2)	0.17
Vepinin	High	No	Yes	Yes	Yes	Yes	yes	No (1)	0.55

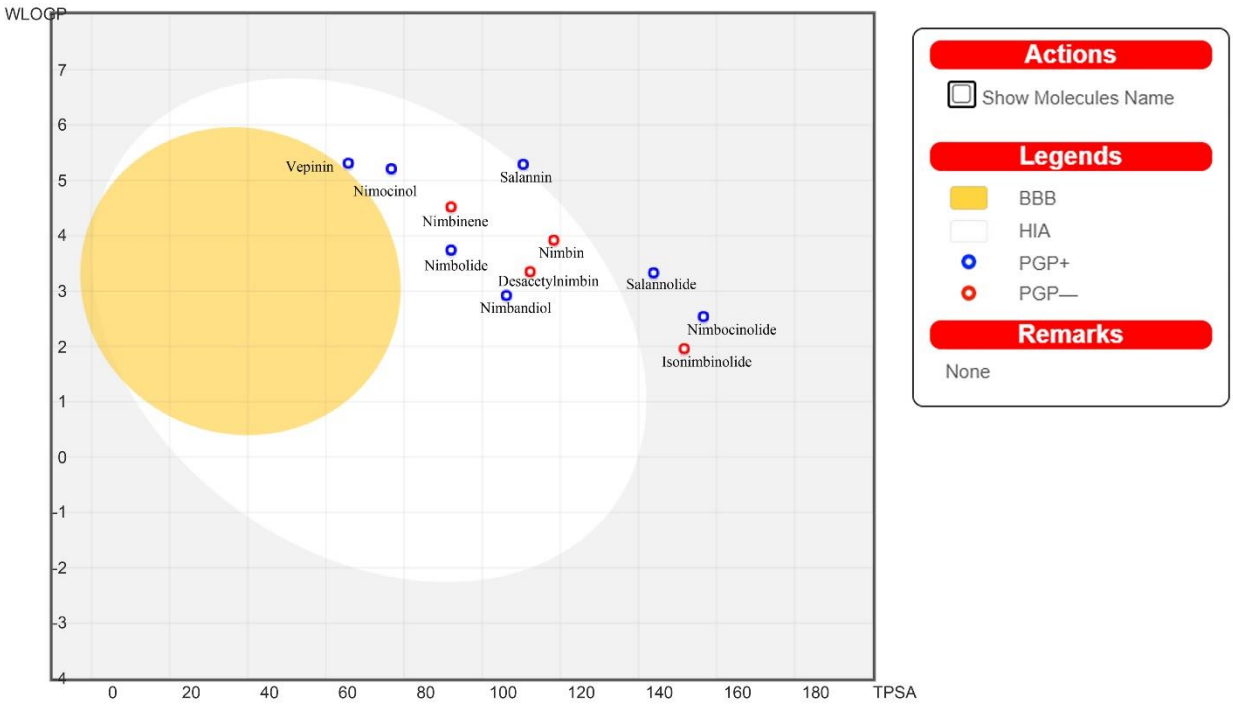

Figure 12: *BOILED-EGG MODEL for Absorption in the Gastrointestinal tract and Penetration into the Brain: Molecules in the yolk of the BOILED-Egg are thought to passively pass through the Blood-brain barrier (BBB). Molecules in the white of a Boiled Egg are thought to be absorbed passively by the digestive tract. The p-glycoproteins are thought to get rid of the blue-dotted molecules from the Central Nervous System (CNS)*

Apart from this pharmacokinetics data, another important parameter is drug likeliness. An orally bioavailable drug candidate is one that displays drug-like properties. Pharmaceutical giants routinely utilise filters like the Lipinski (Pfizer), Ghose (Amgen), Veber (GSK), Egan (Pharmacia), and Muegge (Bayer) filters to weed out candidates with undesirable pharmacokinetic profiles.[47] The most promising result was shown by nimbandiol and nimbolide without violating a single drug likeliness methods. Desacetylnimbin, nimbin, nimbinene, salannin and vepinin violated single out of five methods and all had a good bioavailability score. The most unsatisfactory result was observed with isonimbinolide, nimbocinolide and salannolide violating all the drug likeliness filters and had low bioavailability score (**Table 3**). Similar type of ADMET and drug likeliness studies were reported earlier by Laskar et. al.[47]

Acknowledgement

Authors are grateful to Shuvasish Choudhury, Assistant Professor of the Department of Life Science and Bioinformatics, Assam University, Silchar, Assam, India and DBT- Bioinformatics Infrastructure Facility (BIF) of Assam University, India for software and infrastructure support, and DelCON's e-Journal Access Facility for e-journal access facility. Lutfun Nahar gratefully acknowledges the financial support of the European Regional Development Grants - Projects 23-05474S and 23-05389S.

References

1. Alzohairy, M.A. Therapeutics role of Azadirachta indica (Neem) and their active constituents in diseases prevention and treatment. *Evidence-Based Complementary and Alternative Medicine* 2016;2016*:* 7382506.

2. Tiwari, R.; Verma, A.K.; Chakraborty, S.; Dhama, K.; Singh, S.V. Neem (Azadirachta indica) and its potential for safeguarding health of animals and humans: A review. *Journal of Biological Sciences* 2014; 14: 110-123.

3. Brahmachari, G. Neem—an omnipotent plant: a retrospection. *Chembiochem* 2004; 5: 408-421.

4. Santhosh, V.; Navartnam, V. Neem prehistory to contemporary medicinal uses to humankind, Asian Pac J Trop. *Biomed* 2013; 3: 505-519.

5. Biswas, K.; Chattopadhyay, I.; Banerjee, R.K.; Bandyopadhyay, U. Biological activities and medicinal properties of neem (Azadirachta indica). *Current science* 2002; 82(11):1336-1345.

6. Patel, S.M.; Venkata, K.C.N.; Bhattacharyya, P.; Sethi, G.; Bishayee, A. Potential of neem (Azadirachta indica L.) for prevention and treatment of oncologic diseases. In Proceedings of the Seminars in cancer biology, 2016; 40(41):100-115.

7. Nath, R.; Singha, S.; Nath, D.; Das, G.; Patra, J.K.; Talukdar, A.D. Phytochemicals from Allium tuberosum Rottler ex Spreng Show Potent Inhibitory Activity against B-Raf, EGFR, K-Ras, and PI3K of Non-Small Cell Lung Cancer Targets. *Applied Sciences* 2022; 12: 11749.

8. Hiraki, T.; Gobara, H.; Fujiwara, H.; Ishii, H.; Tomita, K.; Uka, M.; Makimoto, S.; Kanazawa, S. Lung cancer ablation: complications. *Semin Intervent Radiol* 2013; 30: 169-175.

9. Zhang, X.; Shen, T.; Zhou, X.; Tang, X.; Gao, R.; Xu, L.; Wang, L.; Zhou, Z.; Lin, J.; Hu, Y. Network pharmacology based virtual screening of active constituents of Prunella vulgaris L. and the molecular mechanism against breast cancer. *Scientific Reports* 2020; 10: 1-12.

10. Oliveros, J.C. VENNY. An interactive tool for comparing lists with Venn Diagrams. *http://bioinfogp. cnb. csic. es/tools/venny/index. html* 2007.

11. Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K.P. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Research* 2015; 43: D447-D452.

12. Huang, S.; Zhang, Z.; Li, W.; Kong, F.; Yi, P.; Huang, J.; Mao, D.; Peng, W.; Zhang, S. Network pharmacology-based prediction and verification of the active ingredients and potential targets of zuojinwan for treating colorectal cancer. *Drug Design, Development and Therapy* 2020; 14: 2725.

13. Bader, G.D.; Hogue, C.W.J.B.B. An automated method for finding molecular complexes in large protein interaction networks. 2003; 4: 1-27.

14. Daina, A.; Michielin, O.; Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports* 2017; 7: 1-13.

15. Putora, P.; Schneider, T.; Rodriguez, R.; Früh, M. Targeted therapy in non-small cell lung cancer. *Breathe* 2012; 8: 206-215.

16. Normanno, N.; Bianco, C.; Strizzi, L.; Mancino, M.; Maiello, M.; Luca, A.; Caponigro, F.; Salomon, D. The ErbB receptors and their ligands in cancer: an overview. *Current Drug Targets* 2005; 6: 243-257.

17. Pao, W.; Chmielecki, J. Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. *Nature Reviews Cancer* 2010; 10: 760-774.

18. Park, S.J.; More, S.; Murtuza, A.; Woodward, B.D.; Husain, H. New targets in non–small cell lung cancer. *Hematology/Oncology Clinics* 2017; 31: 113-129.

19. Shigematsu, H.; Lin, L.; Takahashi, T.; Nomura, M.; Suzuki, M.; Wistuba, I.I.; Fong, K.M.; Lee, H.; Toyooka, S.; Shimizu, N. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *Journal of the National Cancer Institute* 2005; 97: 339-346.

20. Tsao, A.S.; Scagliotti, G.V.; Bunn Jr, P.A.; Carbone, D.P.; Warren, G.W.; Bai, C.; De Koning, H.J.; Yousaf-Khan, A.U.; McWilliams, A.; Tsao, M.S. Scientific advances in lung cancer 2015. *Journal of Thoracic Oncology* 2016; 11: 613-638.

21. Miller, V.A.; Hirsh, V.; Cadranel, J.; Chen, Y.-M.; Park, K.; Kim, S.-W.; Zhou, C.; Su, W.-C.; Wang, M.; Sun, Y. Afatinib versus placebo for patients with advanced, metastatic non-small-cell lung cancer after failure of erlotinib, gefitinib, or both, and one or two lines of chemotherapy (LUX-Lung 1): a phase 2b/3 randomised trial. *The lancet Oncology* 2012; 13: 528-538.

22. Daga, A.; Ansari, A.; Patel, S.; Mirza, S.; Rawal, R.; Umrania, V. Current drugs and drug targets in non-small cell lung cancer: limitations and opportunities. *Asian Pacific Journal of Cancer Prevention* 2015; 16: 4147-4156.

23. Cantwell-Dorris, E.R.; O'Leary, J.J.; Sheils, O.M. BRAFV600E: implications for carcinogenesis and molecular therapy. *Molecular Cancer Therapeutics* 2011; 10: 385-394.

24. Garnett, M.J.; Marais, R. Guilty as charged: B-RAF is a human oncogene. *Cancer Cell* 2004; 6: 313-319.

25. Davies, H.; Bignell, G.R.; Cox, C.; Stephens, P.; Edkins, S.; Clegg, S.; Teague, J.; Woffendin, H.; Garnett, M.J.; Bottomley, W. Mutations of the BRAF gene in human cancer. *Nature* 2002; 417:949-954.

26. Lu, H.-Y.; Qin, J.; Han, N.; Lei, L.; Xie, F.; Li, C. EGFR, KRAS, BRAF, PTEN, and PIK3CA mutation in plasma of small cell lung cancer patients. *OncoTargets and Therapy* 2018; 11: 2217.

27. Christensen, T.D.; Palshof, J.A.; Larsen, F.O.; Poulsen, T.S.; Høgdall, E.; Pfeiffer, P.; Jensen, B.V.; Yilmaz, M.K.; Nielsen, D. Associations between primary tumor RAS, BRAF and PIK3CA mutation status and metastatic site in patients with chemo-resistant metastatic colorectal cancer. *Acta Oncologica* 2018; 57: 1057-1062.

28. Sawa, K.; Koh, Y.; Kawaguchi, T.; Kambayashi, S.; Asai, K.; Mitsuoka, S.; Kimura, T.; Yoshimura, N.; Yoshimoto, N.; Kubo, A. PIK3CA mutation as a distinctive genetic feature of non-

small cell lung cancer with chronic obstructive pulmonary disease: a comprehensive mutational analysis from a multi-institutional cohort. *Lung Cancer* 2017; 112: 96-101.

29. Yu, J.; Bai, H.; Wang, Z.; Wei, Z.; Ding, X.; Duan, J.; Yang, L.; Wu, M.; Wang, Y.; Wang, J. PIK3CA mutation in Chinese patients with lung squamous cell carcinoma. *Chinese Journal of Cancer Research* 2013; 25: 416.

30. Wang, Y.; Wang, Y.; Li, J.; Li, J.; Che, G. Clinical significance of PIK3CA gene in non-small-cell lung cancer: a systematic review and meta-analysis. *BioMed Research International* 2020; 2020: 3608241.

31. Kawano, O.; Sasaki, H.; Endo, K.; Suzuki, E.; Haneda, H.; Yukiue, H.; Kobayashi, Y.; Yano, M.; Fujii, Y. PIK3CA mutation status in Japanese lung cancer patients. *Lung Cancer* 2006; 54: 209-215.

32. Beck, J.T.; Ismail, A.; Tolomeo, C. Targeting the phosphatidylinositol 3-kinase (PI3K)/AKT/mammalian target of rapamycin (mTOR) pathway: an emerging treatment strategy for squamous cell lung carcinoma. *Cancer Treatment Reviews* 2014; 40: 980-989.

33. Chen, J.-Y.; Cheng, Y.-N.; Han, L.; Wei, F.; Yu, W.-W.; Zhang, X.-W.; Cao, S.; Yu, J.-P. Predictive value of K-ras and PIK3CA in non-small cell lung cancer patients treated with EGFR-TKIs: a systemic review and meta-analysis. *Cancer Biology & Medicine* 2015; 12: 126-139.

34. Li, S.; Li, L.; Zhu, Y.; Huang, C.; Qin, Y.; Liu, H.; Ren-Heidenreich, L.; Shi, B.; Ren, H.; Chu, X. Coexistence of EGFR with KRAS, or BRAF, or PIK3CA somatic mutations in lung cancer: a comprehensive mutation profiling from 5125 Chinese cohorts. *British Journal of Cancer* 2014; 110: 2812-2820.

35. Suda, K.; Onozato, R.; Yatabe, Y.; Mitsudomi, T. EGFR T790M mutation: a double role in lung cancer cell survival? *Journal of Thoracic Oncology* 2009; 4: 1-4.

36. Tsai, M.-F.; Chang, T.-H.; Wu, S.-G.; Yang, H.-Y.; Hsu, Y.-C.; Yang, P.-C.; Shih, J.-Y. EGFR-L858R mutant enhances lung adenocarcinoma cell invasive ability and promotes malignant pleural effusion formation through activation of the CXCL12-CXCR4 pathway. *Scientific Reports* 2015; 5: 1-14.

37. Duggirala, K.B.; Lee, Y.; Lee, K. Chronicles of EGFR Tyrosine Kinase Inhibitors: Targeting EGFR C797S Containing Triple Mutations. *Biomolecules & Therapeutics* 2022; 30: 19-27.

38. Saini, N.; Grewal, A.S.; Lather, V.; Gahlawat, S.K. Natural alkaloids targeting EGFR in non-small cell lung cancer: Molecular docking and ADMET predictions. *Chemico-Biological Interactions* 2022; 358: 109901.

39. Alvarez, J.G.B.; Otterson, G.A. Agents to treat BRAF-mutant lung cancer. *Drugs in Context* 2019; 8: 212566

40. Jain, A.S.; Prasad, A.; Pradeep, S.; Dharmashekar, C.; Achar, R.R.; Ekaterina, S.; Victor, S.; Amachawadi, R.G.; Prasad, S.K.; Pruthvish, R.; et al. Everything Old Is New Again: Drug Repurposing Approach for Non-Small Cell Lung Cancer Targeting MAPK Signaling Pathway. *Front Oncol* 2021; 11: article no. 741326.

41. Thirunavukkarasu, M.K.; Suriya, U.; Rungrotmongkol, T.; Karuppasamy, R. In Silico Screening of Available Drugs Targeting Non-Small Cell Lung Cancer Targets: A Drug Repurposing Approach. *Pharmaceutics* 2021; 14*:* article no. 59.

42. Kawano, O.; Sasaki, H.; Endo, K.; Suzuki, E.; Haneda, H.; Yukiue, H.; Kobayashi, Y.; Yano, M.; Fujii, Y. PIK3CA mutation status in Japanese lung cancer patients. *Lung Cancer* 2006; 54: 209-215.

43. Chujo, M.; Noguchi, T.; Miura, T.; Arinaga, M.; Uchida, Y.; Tagawa, Y. Comparative genomic hybridization analysis detected frequent overrepresentation of chromosome 3q in squamous cell carcinoma of the lung. *Lung Cancer* 2002; 38: 23-29.

44. McGowan, M.; Hoven, A.S.; Lund-Iversen, M.; Solberg, S.; Helland, Å.; Hirsch, F.R.; Brustugun, O.T. PIK3CA mutations as prognostic factor in squamous cell lung carcinoma. *Lung Cancer* 2017; 103: 52-57.

45. Wang, L.; Hu, H.; Pan, Y.; Wang, R.; Li, Y.; Shen, L.; Yu, Y.; Li, H.; Cai, D.; Sun, Y.; et al. PIK3CA mutations frequently coexist with EGFR/KRAS mutations in non-small cell lung cancer and suggest poor prognosis in EGFR/KRAS wildtype subgroup. *PLoS One* 2014; 9: article no. e88291.

46. Gupta, R.; Rai, B. Effect of size and surface charge of gold nanoparticles on their skin permeability: a molecular dynamics study. *Scientific Reports* 2017; 7: 1-13.

47. Laskar, Y.B.; Mazumder, P.B.; Talukdar, A.D. *Hibiscus sabdariffa* anthocyanins are potential modulators of estrogen receptor alpha activity with favourable toxicology: a computational analysis using molecular docking, ADME/Tox prediction, 2D/3D QSAR and molecular dynamics simulation. *Journal of Biomolecular Structure and Dynamics* 2021; 39: 1-23.