Journal of Advanced Zoology

ISSN: 0253-7214

*Volume **44** Issue **S-3 Year 2023** Page **410**:423*

Construction and Analysis of Protein–Protein Interaction of Therapeutic Phytoconstituents of Garlic (*Allium sativum*) Targeting Non-small Cell Lung Cancer Protein

Dinesh[1], Khushboo Bhutani[2], Chanchal Malhotra[1*]

[1]*Department of Botany, Baba Mastnath University, Asthal Bohar, Rohtak-124021*
[2]*Department of Biotechnology, SRM University, Sonepat-131029, Haryana*

**Corresponding author's E-mail: Chanchalmalhotra2007@rediffmail.com*

Article History

Received: 06 June 2023
Revised: 05 Sept 2023
Accepted:05 Oct 2023

Abstract

Background: *Advances in genomics and taxonomic literature reveal that lung cancer is a polygenic disease and expose the complexity of cancer-related genes and molecular mechanisms. Several chemically synthesized drugs are used in cancer treatment, but this is still a challenging task because of their lower efficacy and side effects. In recent years, phytochemicals have played a vital role in novel drug discovery; garlic was used in this study to treat lung cancer. This study aimed to determine the specific target, pharmacological, and molecular mechanisms of garlic phytoconstituents in lung cancer treatment.* ***Methodology:*** *Protein–protein interaction network and KEGG signaling pathway analysis of no small cell lung cancer and garlic phytoconstituents revealed that out of 68 targets of non-small cell lung cancer, 14 were targeted by garlic phytoconstituents.* ***Results:*** *Only 14 phytoconstituents of Allium sativum L. have regulatory effects on 14 targets of non-small cell lung cancer. Garlic phytoconstituents have an inhibitory effect on lung cancer progression by regulating highly enriched pathways in lung cancer, such as the epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor resistance pathway.* ***Conclusion:*** *Some garlic phytoconstituents, such as allixin, ajoene, carvacrol, and some derivative of allicin have drug-like properties that target these targets and act by regulating different molecular pathways of cancer progression. The results of this study require further in vivo or experimental studies to confirm their value in lung cancer treatment and to identify the exact binding sites of the selected targets on which ligands can bind.*

Keywords: *PPI Network, Lung Cancer, Target Protein, Garlic Phytoconstituents*

1. Introduction

Lung cancer is a serious health issue in the modern world owing to environmental pollution, smoking, and genetic mutations. It is a polygenic disease and mutation in some genes like EGFR, KRAS, ALK, RET, MET, HRAS, BRAF, HER2/ERBB2, and ROT causes abnormalities in cell signaling pathways such as MAPK/ERK, JAK/STAT and PI3K/AKT. [1-2] It is essential to identify new targets and mechanisms or pathways at the early stages of tumors. Numerous serious human disorders, including are caused due to abnormalities in protein–protein interactions network. [3-4] We are blessed with a vast array of natural medicines, and more than half of all the medications used in modern treatments are made from natural materials and their derivatives. Plants contain secondary metabolites such as flavonoids, alkaloids, phenols, and terpenoids, which have therapeutic value in the homeopathic and Ayurvedic framework of medicine. Some biological active compounds of garlic also used in various disease treatment and have anti cancerous value.[5] Allicin is effective in almost all types of cancers, and the combined use of allicin and 5-FU is more effective against HCC. It also effective in

gastrointestinal cancer treatment by regulating caspase-3, p38 gene expression.[6] Other synthetic drugs used in cancer treatment as cisplatin and vinorelbine, gefitinib, Adriamycin, and campothecin targets KRAS, EGFR, and MYC gene family respectively; but have more side effects and drug resistance problem.[7] So, Natural product library can be used as starting points in drug development because they have good absorption and metabolism, wide pharmacological area and low toxicity as compared to synthetic chemicals.[8] Herein, *Allium sativum* L. (garlic) is used as a medicinal plant to analyze the interaction between the effective phytoconstituents of garlic and key targets of lung cancer. However, experimental approaches in PPIs network identification are expensive, time-consuming, laborious, and challenging. In response to this, many computational methods have been developed to predict the PPIs network of any disease.[9] Advances in the field of PPI targetable drugs accelerate the drug development from natural product library with negligible side effects.[10] Numerous efforts have been made to identify small compounds of garlic that can prevent protein dimerization or protein(peptide)-receptor interaction in lung cancer treatment.[11] PPIs interaction of seed proteins and GO analysis uncovered new pathways and roles for these targets in lung cancer development. Degree centrality, closeness centrality, betweenness centrality, and cluster coefficient are methods for hub gene identification, which play crucial roles in the network.[12-13] According to the literature, *EGFR* is the most common target in all types of cancer, including lung cancer, which is involved in most pathways and is highly regulated by neighboring proteins.[14] In this study, we identified MAPK1, MAPK3, STAT3, EGFR, GRB2, and JAK3 as potential targets of lung cancer that can be used in novel drug discovery from garlic products. The results of this study clarify the therapeutic value, multitargeting action, and ADME properties of garlic phytoconstituents, which can be used in the pharmaceutical industry for drug development against lung cancer with negligible side effects.

2. Materials And Methods

Selection of plant

In the current study, garlic (*Allium sativum* L.) was selected as the medicinal plant because of its anticancer activity. Garlic extract, its phytoconstituents, and formulations have been shown to deter different stages of cancer, such as initiation, promotion, and progression.

Withdrawing chemical components of garlic

All active ingredients of garlic (*Allium sativum* L.) were obtained manually from the IMPPAT database [15] and literature basis as shown in Table 1. To date, this is the largest database of phytochemicals from Indian medicinal plants.

Screening for drug-likeness

The 2D SDF files and canonical SMILES representations of each molecule were obtained from the PubChem database accessed in September 2021.[16] Subsequently, energy minimization was performed on these molecules using Open Babel to convert the SDF files into the PDB format. All phytochemical compounds were then evaluated in Drulito to assess their drug-likeness properties and bioactivity, based on Lipinski's Rule of Five. For an in-depth analysis of the physicochemical and Absorption, Distribution, Metabolism, and Excretion (ADME) properties of these phytoconstituents, the SWISS ADME database was used.[17]

Target gene prediction

Swiss Target Prediction was used to predict the potential targets of the screened phytoconstituents found in garlic.[18] To identify significant proteins associated with non-small cell lung cancer, data were sourced from the STRING v11.5 database using the KEGG pathway identifier map05223.[18]

Construction and analysis of common PPI network of predicted targets

The STRING v11.5 database was used to create a protein–protein interaction (PPI) network of the predicted targets.[18] Prior to this, duplicate and false-positive targets were meticulously eliminated from the list of phytoconstituent targets. The resulting refined list was then compared to potential targets associated with non-small cell lung cancer (NSCLC). Common targets between the two lists were utilized as input data for constructing the integrated PPI network using a medium confidence score threshold (>0.70). Subsequently, the network was analyzed for the number of nodes and edges.

Topological and functional enrichment analysis of string network

The constructed network was imported into the Cytoscape software to identify seed proteins based on topological parameters, including degree rank, bottleneck score, stress, clustering coefficient, and betweenness centrality. Cytoscape, a Java-based program equipped with various plugins, such as Cytohubba, String, and Cluster Maker, was employed for network analysis. In this study, two key parameters, degree rank and bottleneck score, were used for hub gene selection.[19] Functional enrichment analysis and examination of KEGG signaling pathways associated with the identified targets were performed using ShinyGO 0.80. [20].

3. Results and Discussion

Selection of active phytoconstituents

A total of 54 compounds from garlic (*Allium sativum L.*) were obtained from the IMPPAT database and literature, as detailed in Table 1. These compounds were screened process to identify active compounds with drug-like properties, favorable ADME characteristics, and good bioavailability.

Table 1. List of *Allium sativum* phytoconstituents obtained from IMPPAT database.

S.no.	Phytochemical Name	Canonical Smile
1.	Flavylium Perchlorate	c1ccc(cc1)c1ccc2c([o+]1)cccc2
2.	Gibberellin A7	OC(=O)C1C2[C@]3(C4[C@]51CC(=C)[C@@H](C5)CC4)C=C[C@@H](C2(C)C(=O)O3)O
3.	Gibberellins	OC(=O)[C@H]1[C@H]2[C@]3(C4[C@]51CC(=C)[C@](C5)(O)CC4)C=C[C@@H](C2(C)C(=O)O3)O
4.	Phytosterols	CC[C@H](C(C)C)CC[C@H]([C@H]1CCC2[C@]1(C)CC[C@H]1[C@H]2CC=C2[C@]1(C)CC[C@@H](C2)O)C
5.	(E)-1-Propenyl 2-Propenyl Disulfide	C=CCSS/C=C/C
6.	1-Hexanol	CCCCCCO
7.	1,2-Cyclopentanedithiol	SC1CCCC1S
8.	1,3-Dithiane	C1CCSCS1
9.	2-Coumarinate	OC(=O)/C=Cc1ccccc1O
10.	2,4-Dimethylsulfolane	CC1CC(S(=O)(=O)C1)C
11.	2,5-Dimethylpyrazine	Cc1ncc(nc1)C
12.	2,5-Dimethylthiophene	Cc1ccc(s1)C
13.	2,6-Dimethylpyrazine	Cc1cncc(n1)C
14.	3(Allylsulfinyl) Alanin	[O-]C(=O)[C@@H]([NH3+])CS(=O)CC=C
15.	3-Vinyl-4H-1,2-Dithiin	C=CC1CC=CSS1
16.	4-Aminospiroperidol	Nc1ccc(cc1)N1CNC(=O)C21CCN(CC2)CCCC(=O)c1ccc(cc1)F

17.	Ac1l96yb	OC(=O)[C@H]1OC(O[C@H]2[C@H](O[C@@H]([C@@H]([C@H]2O)O)O)C(=O)O)[C@@H]([C@H]([C@H]1OC1O[C@H](C(=O)O)[C@@H]([C@@H]([C@H]1O)O)O)O)O
18.	Adenosine	OC[C@H]1O[C@H]([C@@H]([C@@H]1O)O)n1cnc2c1ncnc2N
19.	Agapanthagenin	C[C@@H]1CC[C@@]2(OC1)O[C@@H]1[C@H]([C@@H]2C)[C@@]2([C@@H](C1)[C@@H]1CC[C@@]3([C@]([C@H]1CC2)(C)C[C@H]([C@@H](C3)O)O)O)C
20.	Ajoene	C=CCSS/C=C/CS(=O)CC=C
21.	Allicin	C=CCSS(=O)CC=C
22.	Allixin	CCCCCC1=C(C(=O)C(=C(O1)C)OC)O
23.	Allyl Alcohol	OCC=C
24.	Allyl Cis-1-Propenyl Disulfide	C=CCSS/C=CC
25.	Allyl Methyl Disulfide	CSSCC=C
26.	Allyl Methyl Sulfide	CSCC=C
27.	Allyl Methyl Trisulfide	CSSSCC=C
28.	Allyl Propyl Disulfide	CCCSSCC=C
29.	Beta-Carotene	C/C(=CC=CC=C(C=CC=C(C=CC1=C(C)CCCC1(C)C)/C)/C)/C=C/C=C(/C=C/C1=C(C)CCCC1(C)C)C
30.	Carvacrol	CC(c1ccc(c(c1)O)C)C
31.	Coniferyl Alcohol	OC/C=C/c1ccc(c(c1)OC)O
32.	Diallyl Disulfide	C=CCSSCC=C
33.	Diallyl Sulfide	C=CCSCC=C
34.	Diallyl Tetrasulfide	C=CCSSSSCC=C
35.	Diallyl Trisulfide	C=CCSSSCC=C
36.	Dimethyl Trisulfide	CSSSC
37.	Dipicrylamine	[O][N+](=O)c1cc(cc(c1Nc1c(cc(cc1[N+](=O)[O-])[N+](=O)[O-])[N+](=O)[O-])[N+](=O)[O-])[N+](=O)[O-]
38.	Disulfide, Methyl 1-Propenyl	CSS/C=C/C
39.	Divema	O=C1C=CC(=O)O1
40.	Dl-Alanine-15n	CC(C(=O)O)[15NH2]
41.	Dl-Tryptophan	OC(=O)C(Cc1c[nH]c2c1cccc2)N
42.	Eugenol	C=CCc1ccc(c(c1)OC)O
43.	Gamma-Glu-Met	CSCC[C@@H](C(=O)O)NC(=O)CC[C@@H](C(=O)O)N
44.	Isoeruboside B	OCC1OC(OC2CCC3(C(C2)C(O)CC2C3CCC3(C2CC2C3C(C)C3(O_2)CCC(CO3)C)C)C)C(C(C1OC1OC(CO)C(C(C1OC1OC(CO)C(C(C1O)O)O)O)OC1OC(CO)C(C(C1O)O)O)O)O)O
45.	Levan N	OC[C@H]1O[C@@]([C@H]([C@@H]1O)O)(CO)OC[C@H]1O[C@@]([C@H]([C@@H]1O)O)(CO)OC[C@H]1O[C@]([C@H]([C@@H]1O)O)(O)CO
46.	Metyrapone	O=C(C(c1cccnc1)(C)C)c1cccnc1

47.	Moupinamide	COc1cc(/C=C/C(=O)NCCc2ccc(cc2)O)ccc1O
48.	Ninhydrin	O=C1c2ccccc2C(=O)C1(O)O
49.	Nonasaccharide Glc4xyl3gal2	OCC1OC(O)C(C(C1OC1OC(COC2OCC(C(C2OC2OC(CO)C(C(C2O)O)O)O)O)C(C(C1O)O)OC1OC(COC2OCC(C(C2OC2OC(CO)C(C(C2O)O)O)O)O)C(C(C1O)O)OC1OC(COC2OCC(C(C2O)O)O)C(C(C1O)O)O)O)O
50.	Oxycellulose	COCC1OC(OC2C(COCC(O)C)OC(C(C2OCC(O)C)OCC(O)C)OC)C(C(C1OC1OC(CO)C(C(C1O)O)OC)OC)OC
51.	Pcso	NC(C(=O)O)CS(=O)CC=C
52.	alliin	C=CCS(=O)CC(C(=O)O)N
53	s-allyl cysteine	C=CCSCC(C(=O)O)N
54	Z-ajoene	C=CCSSC=CCS(=O)CC=C

Among the 52 compounds assessed using the Drulito tool, only 21 met the criteria outlined in Lipinski's Rule of Five, indicating their potential suitability as oral drug candidates. These criteria encompass a molecular weight range of 130–725 Da, a log P (lipophilicity) value of less than 5, fewer than five rotatable bonds, and 0–6 hydrogen bond donors and 0–20 hydrogen bond acceptors. Their physiochemical and ADME properties (high gastrointestinal absorption and blood–brain barrier permeation) are shown in Table 2.[21]

Table 2. Physiochemical and ADME properties of active compounds of garlic.

Compound Name	*MW*	*ROTB*	*HBA*	*HBD*	*AROM*	*MLOGP*	*ALERT*	*GIA*	*BBBP*
(E)-1-Propenyl 2-Propenyl Disulfide	146.27	4	0	0	0	2.35	0	High	Yes
1-Hexanol	102.17	4	1	1	0	1.53	0	High	Yes
1,3-Dithiane	120.24	0	0	0	0	1.42	0	High	Yes
2-Coumarinate	164.16	2	3	2	6	1.28	0	High	Yes
2,5-Dimethylpyrazine	108.14	0	2	0	6	−0.14	0	High	Yes
2,6-Dimethylpyrazine	108.14	0	2	0	6	−0.14	0	High	Yes
3-Vinyl-4H-1,2-Dithiin	144.26	1	0	0	0	1.96	0	High	Yes
Ajoene	234.4	8	1	0	0	2.1	0	High	No
Allicin	162.27	5	1	0	0	1.18	0	High	Yes
Allixin	226.27	5	4	1	6	0.46	0	High	Yes
Allyl Methyl Trisulfide	152.3	4	0	0	0	1.67	0	High	Yes
Allyl Propyl Disulfide	148.29	5	0	0	0	2.46	0	High	Yes
Carvacrol	150.22	1	1	1	6	2.76	0	High	Yes
Diallyl Disulfide	146.27	5	0	0	0	2.35	0	High	Yes
Diallyl Sulfide	114.21	4	0	0	0	2.35	0	High	Yes
Diallyl Trisulfide	178.34	6	0	0	0	2.35	0	High	Yes
Dimethyl Trisulfide	126.26	2	0	0	0	0.84	0	High	Yes
Disulfide, Methyl 1-Propenyl	120.24	2	0	0	0	1.67	0	High	Yes
Alliin	177.22	5	4	2	0	−2.88	0	High	No
s-allyl cysteine	161.22	5	3	2	0	−1.93	0	High	No
Z-ajoene	234.4	8	1	0	0	2.1	0	High	No

* MW-Molecular weight, ROTB-Rotatable bonds, HBD-H-bond donor, HBA-H-bond acceptor, MR - Molecular Refractivity, AROM-Aromatic heavy atoms, GIS- gastrointestinal absorption, BBBP-Blood–brain barrier permeability.

Prediction of target genes and identification of potential targets

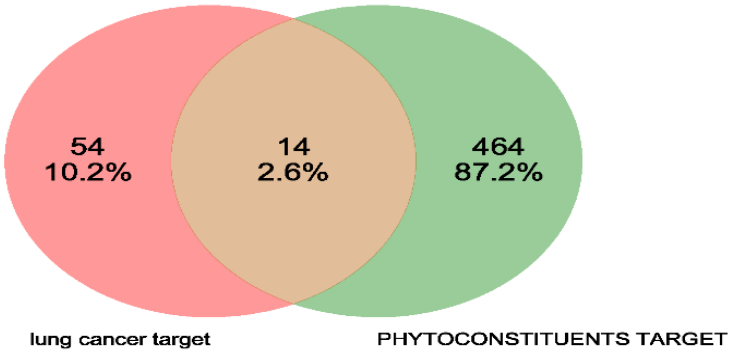


Fig 1. Venn diagram of phytoconstituents and NSCLC targets

A total of 1,150 targets associated with 21 phytoconstituents were initially obtained through Swiss target prediction. However, after meticulously removing duplicates and false-positive targets, only 478 unique targets for the phytoconstituents were retained for comparison with 68 targets related to no small cell lung cancer (NSCLC), as detailed in Table 3. Among these, only 14 targets were found to be common to both datasets, as shown in Figure 1, and were selected for protein–protein interaction (PPI) network analysis.

Table 3. List of phytoconstituents and NSCLC targets.

NSCLC Targets	GADD45B, MAPK1, SOS2, PIK3R2, HGF, CCND1, POLK, PLCG1, RAF1, GADD45G, KRAS, DDB2, CDK4, PIK3R3, E2F3, MAP2K2, MAPK3, PRKCG, AKT3, PIK3CA, STAT3, EGF, CDK6, RB1, TP53, ERBB2, EGFR, BRAF, PIK3CB, ARAF, BAX, STAT5B, TGFA, MAP2K1, PRKCB, MET, EML4, CASP9, RARB, STAT5A, PDPK1, E2F1, RASSF1, RXRG, E2F2, NRAS, GADD45A, STK4, BAK1, RXRB, PIK3CD, ALK, AKT2, GRB2, BAD, CDKN1A, FOXO3, SOS1, JAK3, HRAS, PRKCA, FHIT, CDKN2A, RXRA, PIK3R1, AKT1, RASSF5, PLCG2
Phyto-constituents targets	PIN1, PPARA, CNR2, BCHE, MAOA, ACHE, PTGS1, SLC6A2, HTR2C, CYP1A2, MGLL, PTGS2, ESR2, KIF11, ADRA2A, ADRA2B, DRD1, DRD2, DRD4, OPRM1, DRD3, OPRK1, SLC6A3, CCR4, CXCR2, ADORA3 HTR6, MAPK3, CCR2, HTR1B, KDM4A, CA2, CA4, BBOX1, ALOX15, CXCR3, HTR2A, SHBG, RAPGEF4, CA1, CDC25A, CDC25B, NR1I3, AR, NR1H4, GPBAR1, SHH, TRPM8, MIF, DPP4, SIGMAR1, NPC1L1, CYP11B1, CYP11B2, ESR1, UGT2B7, SLC22A6, HMOX1, JAK1, JAK2, CDK1, FABP4, NR1H3, FABP3, FABP5, SPHK1, FFAR1, FABP2, JAK3, TYK2, GSR, POLA1, AOC3, HSD11B1, CHEK1, CDK5R1 CDK5, ACPP, PARP1, MAP3K14, HSD17B3, G6PD, GABBR1, CHRNA3 CHRNB4, IKBKB, NR3C2, PGR, HMGCR, PKM, CSNK2A1, BRD4, BRD2, MPO, TRPV3, F2RL3, SLC5A7, NR3C1, CES1, CES2, HDAC1, LRRK2, TYMS, ABCC9, EGFR, KCNJ11, ABCC9, PYGL, MAPK1, LTA4H, KDM1A, EPHA4, RPS6KA5, GABRA2 GABRB2, GABRG2, AKR1B10, POLB, STK3, CASP1, STK26, CA7, CA3, CA6, CA12, CA14, CA9, CA13, CA5B, ADA, CTSB, ALDH1A1, CTSK, CTSL, AKR1B1, CA5A, KDM4E, ALOX5, MMP9, MMP1, MMP2, PTPN1, HCAR2, KDM3A, KDM6B, FTO, KDM4C, TPMT, TLR4, KMO, NGFR, KDM2A, CPA1, DAO, TRPA1, GPR17, FBP1, F3, HSD17B2, HSD17B1, CYP2C9, CYP3A4, CYP2C19, PTGDR2, ALPL, ADORA1, ADORA2A, ADORA2B, FOLH1, PTGER4, PTGER2, PTGER3, CTBP2, SLC13A5, CAPN2, CAPN1, APEX1, IDO1, SLC16A1, GSK3B, TLR9, ERBB2, AKR1C4, SRD5A2, DYRK1A, DYRK1B, FYN,LCK, MAOB, PTPN2, SRC, EEF2K, GABRR1, ACE, ELANE, APP, EIF2AK2, LAP3, MMP8, NQO2,

	CHRNA7, DBF4 CDC7, AMPD3, F2, CCND1, CDK4, MME, ACLY, PRKCQ, RARB, PDE4A, PDE4B, PDE4D, PDE4C, CYP2A6, NISCH, CYP19A1, MCL1, GSTP1, CSF1R, MAPK8, MAPK10, FAAH, CTSS, GRM5, CHRM1, PTK2 HSP90AA1, METAP2, MAPK14, GHSR, GRK2, PRSS1, TDO2, CACNA1B, SCN10A, CHRM2, CCR3, NAT1, CHRM4, CHRM5, CHRM3, KCNJ5, KCNJ3, KCNJ6, HCRTR2, HCRTR1, HTR2B, KCNJ3, CMA1, PREP, NPY5R, PSEN2 PSENEN NCSTN APH1A PSEN1 APH1B, ICAM1, SELE, RAF1, PDGFRB, MAP2K2, JAK3 JAK1, CDK4,M AP2K1, AURKC, MAPK11, CAMK2A, HDAC6, RAPGEF3, SIRT2, IDH1, FAP, CASP3, CASP7, HIF1A, AURKA, P2RX7, MTNR1A, MTNR1B, CNR1, EPHX2, TOP2A, FNTA FNTB, GABRB3 GABRG2 GABRA5, PGGT1B FNTA, CTSH, GRM4, KDR, PABPC1, ADRA1D, RPS6KA2, NAAA, FABP1, MKNK1, TNKS2, GRM2, HRH3, HRH4, CTSG, ADH1B, ADH1C, TBXAS1, GCK, CTSC, CTSF, CTSV, ALDH2, PARP3, TYMP, PTPRC, BACE1, DNMT3, ACCNE2, CDK2, CCNE1, CCNB3 CDK1 CCNB1 CCNB2, SLC6A4, PDE7A, PDE7B, TSPO, MAP3K20, ABL2, TGFBR2, TGFBR1, ABAT, QPCT, EGLN1, VEGFA, PDE5A, IMPDH2, PTGES, TUBB1, CASP6, TNKS, HDAC2, HDAC8, ALOX12, KCNMA1, TLR7, SIRT1, PARP2, GABRA1 GABRB2 GABRG2, MMP13, STK33, NQO1, MDM2, RET, FGFR1, TTR, ATR, GABRA1, FADS1, MTOR, MYLK, PLA2G7, RIPK2, LDHA, PDE10A, SLC9A1, PNP, CASR, PLK4, CBR1, HDAC3, ERN1, GSTA1, SRD5A1, DHFR, TNNC1 TNNT2 TNNI3, AOC1, CYP17A1, TERT, CHRNA4 CHRNB2, ALB, PRKCA, TYR, FLT3, DRD5, ESRRG, NOS1, NOS3, ASAH1,MC4R, MAPT, MAP2, MAPK9, LTB4R, CLK1, HDAC11, HDAC10, GAPDH, DUSP3, GPR84, AKT1, ADRB2, ADRB1, CDK2, CCNA1 CCNA2, CDK2, PTK6, PANK3, LDHB, CYP26B1, CYP26A1, ADRA1B, PAK1, GABBR2 GABBR1, EPHX1, GPR55, GPR18, RNASEH1, CD38, HPRT1, GRIA1, GRIA4, GRIA2, GRIK5, KYNU, GFPT1, TH, PEPD, ENPEP, RNPEP, CPB2, ANPEP, ODC1, SLC6A1, OAT, SLC1A1, SLC1A2, CACNA2D1, ERAP2, GRIK1, GRIK2, GRM3, GABRA3 GABRB2 GABRG2, GRIK3, GRM6, SLC1A3, NOS2, SLC6A12, GRM8, SI, GRB2, BHMT, HTR1AHTR7, FUCA1, CPB1, MGAM, ITGB1, ITGA5, AHCY, CPN1, GSTK1, SLC15A1, GRM1, GBA, ARG1, FDPS, PLG, GLA, PTPRA, GBA2, AGL, GLB1, GANAB, GAA, GCLC, ECE1, MAPKAPK2, PLAU, SLC7A5, CPM,DPP7, PIM1, NEU3, PIM2, GRK6, KDM5C, KDM4B, KDM5B, DHODH, CCNC CDK8, PTK2B, GABRB3, GABRA3, GABRG2, GABRB3 GABRG2 GABRA1, GABRA2 GABRB3 GABRG2, STAT3, S1PR3, GABRA5, PTP4A3, HPGDS
Common targets	AKT1, CDK4, EGFR, ERBB2, GRB2, JAK3, MAP2K1, MAP2K2, MAPK1, MAPK3, PRKCA, STAT3, RAF1, RARB

Network analysis

The full string network of the 14 targets generated in the string consists of 14 nodes, 52 edges, and 0.798 average local clustering coefficient, 7.43 average node degree, and PPI enrichment p-value is <1.0e-16. All the interactions between proteins are evidence-based and appear by different color of edges like text-mining, curated database, co-expression, and gene-fusion (Fig. 2).[22]

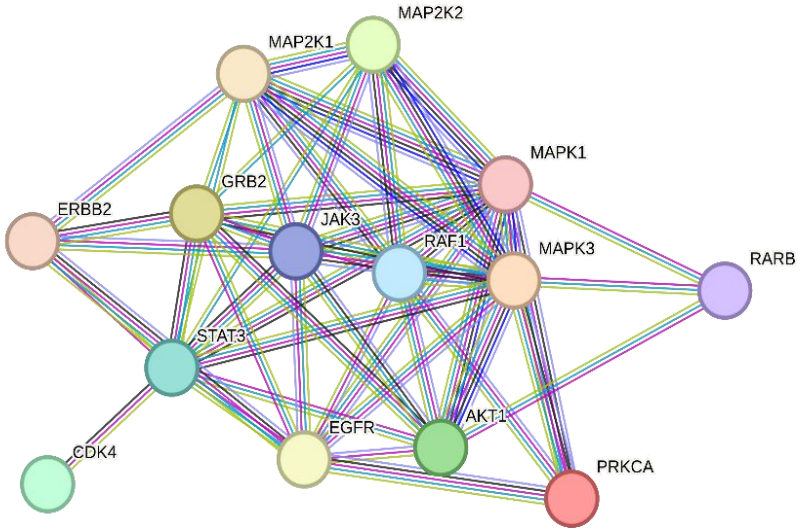


Fig. 2 Integrated PPI network of potential targets.

Topological and KEGG pathway analysis

Topological parameters such as node degree score, bottleneck, closeness, stress, clustering coefficient, and EC centricity represent the significance of nodes inside the network Table 4. The target with the highest degree score represents more interaction in the network, and the highest bottleneck score represent the number of the pathways that passed through them.

Table 4. Topological properties of all targets.

Targets	*D*	*BC*	*B*	*C*	*EC*	*S*	*CC*
CDK4	1	0	1	6.5	0.33333	0	0
RARB	3	0	1	7.66667	0.33333	0	1
PRKCA	4	0	1	8.33333	0.33333	0	1
ERBB2	5	0.28571	1	8.83333	0.33333	2	0.9
RAF1	8	3.61905	1	10.33333	0.33333	16	0.71429
MAP2K2	7	0.85714	1	10	0.5	6	0.90476
AKT1	8	5.42857	2	10.5	0.5	26	0.75
MAP2K1	8	3.10952	1	10.5	0.5	18	0.78571
EGFR	9	8.25238	2	11	0.5	32	0.69444
JAK3	9	3.51429	1	11	0.5	22	0.77778
STAT3	10	27.51429	3	11.5	0.5	66	0.62222
MAPK1	11	14.19048	2	12	0.5	56	0.63636
MAPK3	11	14.19048	2	12	0.5	56	0.63636
GRB2	10	5.0381	1	11.5	0.5	30	0.75556

*D-degree, BC-Betweenness Centricity, B-Bottleneck, C-Closeness, EC- Eccentricity, S- stress, CC-clustering coefficient.

The key proteins identified in this study played critical roles in various biological functions, molecular processes, and KEGG pathways. Table 5 presents the results of the KEGG pathway analysis, with pathways sorted based on their significance. In the bubble chart shown in Figure 3, pathways with the highest enriched FDR values and a greater number of associated genes were considered more significant. Figure 4 represents a KEGG graph, where all 14 genes are highlighted in red and are directly related to non-small cell lung cancer. Furthermore, the KEGG graph and topological analysis of the targeted proteins confirmed that EGFR, MAPK1, MAPK3, GRB2, AKT1, STAT3, and JAK3 exhibited the highest degree scores. These proteins are interconnected with each other and collectively contribute to the regulation of cancer-causing mechanisms and signaling pathways, including the PI3K/AKT, calcium, and MAPK signaling pathways, as described in reference.[24] Inhibition of EGFR expression

can be used to treat NSCLC by modulating several downstream signaling pathways, such as the ErbB signaling pathway, PI3K/AKT pathway, calcium signaling pathway, and MAPK signaling pathway. Out of the 21 garlic phytoconstituents that were examined, only 14 were found to have anticancer properties against the 14 NSCLC-associated targets. These 14 phytoconstituents include 2,5-Dimethylpyrazine, 2,6-Dimethylpyrazine, S-allylcysteine, alliin, allicin, ajoene, allixin, Z-ajoene, 1-hexanol, 2-coumarinate, carvacrol, Allyl Methyl Trisulfide, Diallyl Trisulfide, (E)-1-Propenyl 2-Propenyl Disulfide, and Diallyldisulfide.

Table 5. KEGG signaling pathways of key proteins of string network (ShinyGO 0.80 http://bioinformatics.sdstate.edu/go80/).

Pathway	Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Genes
Nonsmall cell lung cancer	7.17E-28	14	72	113.5833333	CDK4 EGFR ERBB2 AKT1 GRB2 JAK3 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1 RARB STAT3
EGFR tyrosine kinase inhibitor resistance	9.62E-19	11	79	81.3363472	EGFR ERBB2 AKT1 GRB2 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1 STAT3
Glioma	1.19E-16	10	75	77.88571429	CDK4 EGFR AKT1 GRB2 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1
ErbB signaling pathway	2.97E-16	10	84	69.54081633	EGFR ERBB2 AKT1 GRB2 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1
Pathways in cancer	6.39E-16	14	530	15.43018868	CDK4 EGFR ERBB2 AKT1 GRB2 JAK3 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1 RARB STAT3
Endocrine resistance	7.21E-16	10	95	61.4887218	CDK4 EGFR ERBB2 AKT1 GRB2 MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1
Endometrial cancer	1.08E-15	9	58	90.64285714	EGFR ERBB2 AKT1 GRB2 MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1
Pancreatic cancer	1.11E-14	9	76	69.17481203	CDK4 EGFR ERBB2 AKT1 MAPK1 MAPK3 MAP2K1 RAF1 STAT3
Proteoglycans in cancer	1.11E-14	11	202	31.80975955	EGFR ERBB2 AKT1 GRB2 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1 STAT3

Human cytomegalovirus infection	2.82E-14	11	224	28.68558673	CDK4 EGFR AKT1 GRB2 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1 STAT3
Breast cancer	3.55E-14	10	147	39.73760933	CDK4 EGFR ERBB2 AKT1 GRB2 MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1
Gastric cancer	3.55E-14	10	148	39.46911197	EGFR ERBB2 AKT1 GRB2 MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1 RARB
PI3K-Akt signaling pathway	3.81E-14	12	354	19.80145278	CDK4 EGFR ERBB2 AKT1 GRB2 JAK3 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1
Hepatitis C	5.57E-14	10	157	37.20655141	CDK4 EGFR AKT1 GRB2 MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1 STAT3
Prostate cancer	6.05E-14	9	97	54.1988218	EGFR ERBB2 AKT1 GRB2 MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1
Choline metabolism in cancer	6.05E-14	9	98	53.64577259	EGFR AKT1 GRB2 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1
MicroRNAs in cancer	6.05E-14	10	161	36.28216504	EGFR ERBB2 GRB2 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1 STAT3
Hepatitis B	6.05E-14	10	162	36.05820106	AKT1 GRB2 JAK3 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1 STAT3
Hepatocellular carcinoma	7.71E-14	10	167	34.9786142	CDK4 EGFR AKT1 GRB2 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 RAF1
HIF-1 signaling pathway	1.47E-13	9	109	48.23197903	EGFR ERBB2 AKT1 PRKCA MAPK1 MAPK3 MAP2K1 MAP2K2 STAT3

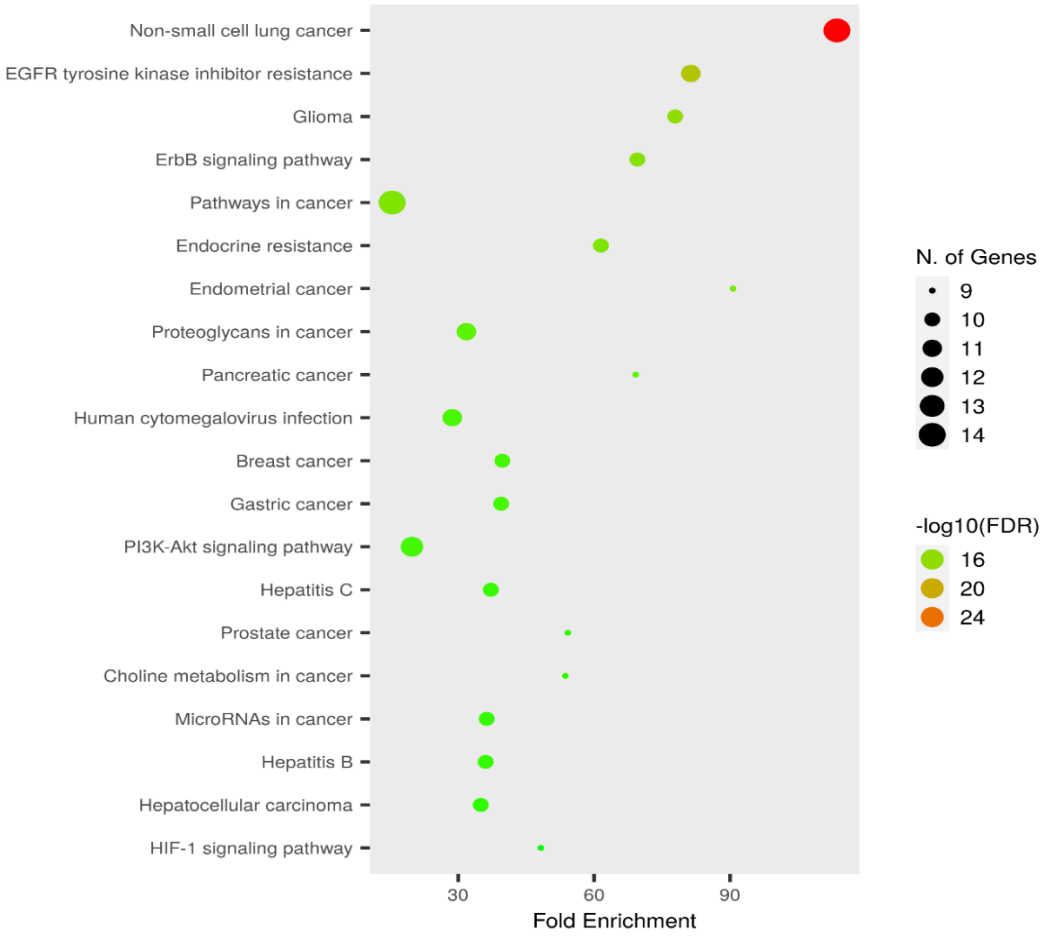
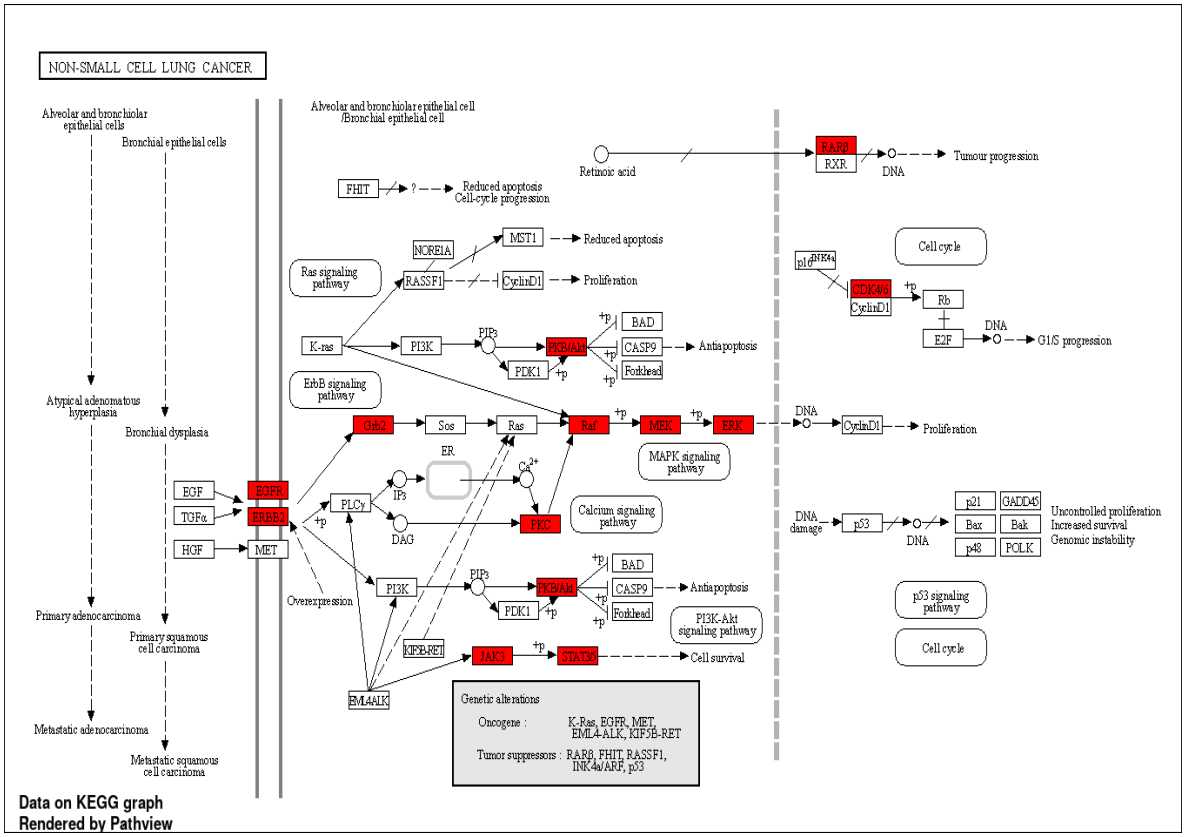


Fig. 3 Bubble chart of KEGG signaling pathways analysis.



Fifig. 4 KEGG graph of selected target (red color) targeted by garlic phytoconstituents.

Topological analysis of PPI network of lung cancer and bioactive compounds of garlic supporting the idea that lung cancer is a polygenic disease and help to understand the mechanism of cancer development and molecular pathways and specific target of garlic.[25-26] Swiss ADME properties of garlic demonstrates that garlic extract have high gastrointestinal absorption and also interact with CYP isoform play important role in drug metabolism and also as xenobiotics defends our body from certain type of carcinogens.[27] 2,6-Dimethylpyrazine targeting CYP2A6 (CYP family1-4) involved in drug detoxification and metabolism of drugs.[28] Based on the result of network pharmacology it was found that most of the hub genes of lung cancer and targeted by garlic phytoconstituents which confirm their anti-cancerous value in lung cancer treatment.[29] *EGFR* is also involved in various types of cancers, such as bladder, prostate, endometrial, breast, and pancreatic cancers. Some synthetic drugs, such as gefitinib, erlotinib, and crizotinib, act on only a single *EGFR* target. [30] According to this study garlic products have multitargeting action or polypharmacological effect,[31] targeted all of 11 targets of *EGFR, ERBB2, AKT1, GRB2, PRKCA, MAPK1, MAPK3, MAP2K1, MAP2K2, RAF1* and *STAT3* related to EGFR tyrosine kinase inhibitor resistance pathways and regulate different stage and other pathways of lung cancer.[32] A meta-analysis by Ponvilawan et al. 2020, found the relationship between HCV and lung cancer [33] and this study also provide a link between hepatitis B, C viral infection, and lung cancer. These results reveal that garlic has anticancer properties and can be used in novel drug discovery because all 14 targets linked to lung cancer, *EGFR, ERBB2, AKT1, GRB2, PRKCA, MAPK1, MAPK3, MAP2K1, MAP2K2, RAF1*, and *STAT3*, are targeted by garlic phytoconstituents at various stages of development. It can be used as a novel strategy for drug discovery from natural products for lung cancer treatment because of its good ADME properties and high regulation of various pathways of cancer proliferation.[34]

4. Conclusion

This study found a link between lung cancer and other diseases such as hepatitis, cytomegalovirus, and other types of cancer. Garlic products have multitargeting action and polypharmacological effects on all 14 targets linked to lung cancer development stages, such as initiation, proliferation, and metastasis. It also regulates various signaling pathways like ErbB signaling pathway, VEGF signaling pathway, MAPK signaling pathway and PI3K-Akt signaling pathway and EGFR tyrosine kinase inhibitor

resistance pathways in lung cancer treatment. Carvacrol is only a single component targeting PRKCA in cancer cells. Some phytoconstituents, such as s-allylcysteine, alliin, allicin, allixin and carvacrol, ajoene, 2,5-dimethylpyrazine and 2,6-dimethylpyrazine selected as key components with anticancer properties. Z-ajoene, one of the seed components, has the highest degree and targets the top bottleneck proteins of the ppi network, as such MAPK1, MAPK3, STAT3, EGFR, GRB2, and JAK3, and controls the survival and proliferation of lung cancer cells. These results can be used in further *in silico* docking studies and *in vivo* experiments of novel drug discovery to confirm the exact binding, affinity, and action of these selected components on specific targets.

Acknowledgments
The authors are thankful to the Department of Botany, Faculty of Science, for providing the necessary facilities to carry out this research work.

Funding source
This research did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

References:

1. Chan BA, Hughes BG. Targeted therapy for non-small cell lung cancer: Current standards and the promise of the future. Transl Lung Cancer Res 2015; 4: 36-54.
2. Byers LA, Rudin CM. Small cell lung cancer: Where do we go from here? Cancer 2015; 121: 664-672.
3. Archakov AI, Govorun VM, Dubanov AV, Ivanov YD, Veselovsky AV, Lewi P, Janssen P. Protein-protein interactions as a target for drugs in proteomics. Proteomics 2003; 3: 380-391.
4. Silverbush D, Sharan R. A systematic approach to orient the human protein-protein interaction network. Nat Commun 2019; 10: 3015.
5. Huang M, Lu JJ, Ding J. Natural products in cancer therapy: Past, present and future. Nat Prod Bioprospect. 2021; 11: 5-13.
6. Zhang Y, Liu X, Ruan J, Zhuang X, Zhang X, Li Z. Phytochemicals of garlic: Promising candidates for cancer therapy. Biomed Pharmacother 2020; 123: 109730.
7. Cohen MH, Gootenberg J, Keegan P, Pazdur R. FDA drug approval summary: Bevacizumab (Avastin®) plus carboplatin and paclitaxel as first-line treatment of advanced/metastatic recurrent nonsquamous non-small cell lung cancer. *Oncologist* 2007; 12: 713-718.
8. Jin X, Lee K, Kim NH, Kim HS, Yook JI, Choi J. Natural products used as a chemical library for protein-protein interaction targeted drug discovery. J Mol Graph Model 2018; 79: 46-58.
9. Jiang M, Niu C, Cao J, Ni DA, Chu Z. In silico-prediction of protein–protein interactions network about MAPKs and PP2Cs reveals a novel docking site variant in *Brachypodium distachyon.* Sci Rep 2018; 8: 15083.
10. Huang CY, Ju DT, Chang CF, Muralidhar RP, Velmurugan BK. A review on the effects of current chemotherapy drugs and natural agents in treating non-small cell lung cancer. Biomedicine 2017; 7: 23.
11. Lu H, Zhou Q, He J, et al. Recent advances in the development of protein-protein interactions modulators: Mechanisms and clinical trials. Signal Transduct Target Ther 2020; 5: 213.
12. Lee SA, Chan CH, Chen TC, et al. POINeT: Protein interactome with sub-network analysis and hub prioritization. BMC Bioinformatics 2009; 10: 114.
13. Ran J, Li H, Fu J, et al. Construction and analysis of the protein-protein interaction network related to essential hypertension. BMC Syst Biol 2013; 7: 32.
14. Li Z, Ivanov AA, Su R, Gonzalez-Pecchi V, Qi Q, Liu S, et al. The OncoPPi network of cancer-focused protein-protein interactions to inform biological insights and therapeutic strategies. Nat Commun 2017; 8: 14356.
15. IMPPAT: Indian Medicinal Plants, Phytochemistry And Therapeutics [Homepage on the Internet]. Available from: https://cb.imsc.res.in/imppat/home.
16. PubChem.NCBI. [Homepage on the Internet]. Available from: https://pubchem.ncbi.nlm.nih.gov/.
17. Swisstargetprediction.org. [Homepage on the Internet]. Available from: http://www.swisstargetprediction.ch/.
18. String-db.org [Homepage on the Internet]. Available from: https://string-db.org/
19. Su W, Guan Y, Huang B, et al. Bioinformatic analysis reveals hub genes and pathways that promote melanoma metastasis. BMC Cancer 2020; 20: 863.
20. ShinyGo [Homepage on the Internet]. Available from: http://bioinformatics.sdstate.edu/go80/
21. Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep 2017; 7: 42717.

22. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019; 47: 607-613.
23. Rezaei-Tavirani M, Rezaei-Taviran S, Mansouri M, Rostami-Nejad M, Rezaei-Tavirani M. Protein-protein interaction network analysis for a biomarker panel related to human esophageal adenocarcinoma. Asian Pac J Cancer Prev 2017; 18: 3357-3363.
24. Karimizadeh E, Sharifi-Zarchi A, Nikaein H, et al. Analysis of gene expression profiles and protein-protein interaction networks in multiple tissues of systemic sclerosis. BMC Med Genomics 2019; 12: 199.
25. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 2014; 505: 495-501.
26. Wu I, Wang X. A novel approach to topological network analysis for the identification of metrics and signatures in non-small cell lung cancer. Sci Rep 2023; 13: 8223.
27. Jablonsky M, Haz A, Burcova Z, Kreps F, Jablonsky J. Pharmacokinetic properties of biomass-extracted substances isolated by green solvents. *BioRe*s 2019; 14: 6294-6303.
28. Fan Z, Wang Z, Chen W, Cao Z, Li Y. Association between the CYP11 family and six cancer types. Oncol Lett 2016; 12: 35-40.
29. Kodali RT, Eslick GD. Meta-analysis: Does garlic intake reduce risk of gastric cancer? Nutr Cancer 2015; 67: 1-11.
30. Ahmad U, Abdullah S, Chau M, et al. Analysis of PPI networks of transcriptomic expression identifies hub genes associated with Newcastle disease virus persistent infection in bladder cancer. Sci Rep 2023; 13: 7323.
31. Sahin K, Ali S, Sahin N, Orhan C, Kucuk O. Lycopene: Multitargeted applications in cancer therapy. Nat Prod Cancer Drug Dis 2017; 79.
32. Lee DH, Kwak HJ, Shin Y, Kim SJ, Lee GH, Park I-H, et al. Elucidation of phytochemicals affecting platelet responsiveness in Dangguisu-san: Active ingredient prediction and experimental research using network pharmacology. *Plants* 2023; 12: 1120.
33. Ponvilawan B, Charoenngam N, Rujirachun P, Wattanachayakul P, Tornsatitkul S, Rittiphairoj T, et al. Chronic hepatitis C virus infection is associated with an increased risk of lung cancer: A systematic review and meta-analysis. Lung 2020; 198: 705-714.
34. Bommhardt U, Schraven B, Simeoni L. Beyond TCR signaling: Emerging functions of Lck in cancer and immunotherapy. Int J Molecular Sci 2019; 20: 3500.