Exploring the hub genes and mechanisms of Daphne altaica. treating esophageal squamous cell carcinoma based on Network Pharmacology and Bioinformatics Analysis

Sendaer Hailati
Xinjiang Medical University
Ziruo Talihati
Xinjiang Medical University
Kayisaier Abudurousuli
Xinjiang Medical University
Meng Han
Xinjiang Medical University
Muhadaisi Nuer
Xinjiang Medical University
Nawaz Khan
Xinjiang Medical University
Nulibiya Maihemuti
Xinjiang Medical University
Jimilihan Simayi
Xinjiang Medical University
Dilihuma Dilimulati
Xinjiang Medical University
Nuerbiye nueraihemaiti
Xinjiang Medical University
Wenting Zhou (zwt@xjmu.edu.cn)
Xinjiang Medical University

Research Article

Keywords: Esophageal Squamous Cell Carcinoma, Daphne altaica., Network Pharmacology, Bioinformatics Analysis, hub genes

Posted Date: April 18th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2812974/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Journal of Cancer Research and Clinical Oncology on April 23rd, 2023. See the published version at https://doi.org/10.1007/s00432-023-04797-w.

Abstract

Purpose

Esophageal squamous cell carcinoma (ESCC), is a frequent digestive tract malignant carcinoma with a high fatality rate. *Daphne altaica.* (*D. altaica*), a medicinal plant that is frequently employed in Kazakh traditional medicine, and which has traditionally been used to cure cancer and respiratory conditions, but research on the mechanism is lacking. Therefore, we examined and verified the hub genes and mechanism of *D. altaica* treating ESCC.

Methods

Active compounds and targets of *D. altaica* were screened by databases such as TCMSP, and ESCC targets were screened by databases such as GeneCards and constructed the compound-target network and PPI network. Meantime, datasets between tissues and adjacent non-cancerous tissues from GEO database (GSE100942, GPL570) were analyzed to obtain DEGs using the limma package in R. Hub genes were validated using data from the Kaplan-Meier plotter database, TIMER2.0 and GEPIA2 databases. Finally, AutoDock software was used to predict the binding sites through molecular docking.

Results

In total, 830 compound targets were obtained from TCMSP and other databases. And 17710 disease targets were acquired based on GeneCards and other databases. And we constructed the compound-target network and PPI network. Then, 127 DEGs were observed (82 up-regulated and 45 down-regulated genes). Hub genes were screened including TOP2A, NUF2, CDKN2A, BCHE, and NEK2, and had been validated with the help of several publicly available databases. Finally, molecular docking results showed more stable binding between five hub genes and active compounds .

Conclusions

In the present study, five hub genes were screened and validated, and potential mechanisms of action were predicted, which could provide a theoretical understanding of the treatment of ESCC with *D. altaica.*

Introduction

Esophageal carcinoma (EC) is the seventh most aggressive type of malignancy and because of lacking biomarkers that can facilitate early detection, its treatment remains a challenge. Identifications of EC are two major histological forms-namely Esophageal adenocarcinoma (EAC) and Esophageal squamous cell carcinoma (ESCC), and each of them shows differences in the incidence among populations that are geographically separated (Tungekar et al. 2018). ESCC tends to have a higher global proportion than EAC, and it accounts for 90% of all cases of Esophageal carcinoma worldwide. According to research, the ESCC incidence in the Kazakh population is much higher than in other minorities in Xinjiang, northwestern China (Zheng et al. 2010). There has been some progress in its possible etiological

mechanisms in several studies, including behavioral and environmental risk factors and genetic alterations (Zheng et al.2011). There are research reports, on the unique eating habits of the Kazakh population in northern Xinjiang, such as drinking hot milk tea, bacon, air-dried meat, fermented dairy products, and low intake of vegetables and fruits (Han et al. 2015). In addition, Esophageal carcinoma is also closely related to genetic factors, which makes the incidence of Esophageal carcinoma in the Xinjiang Kazakh population cluster in families (Zhang et al. 2015; Chen et al. 2006). But the most complex mechanisms by which the Kazakh population suffers abnormally remain largely unknown.

Daphne altaica. (*D. altaica*), it was first mentioned as spicy, heated, and lethal in the 15th-century Kazakh classic work "Chipagar Bayan" (Chinese: "Medicine"). The rare plant is primarily found in the Altai Mountains, namely in the Tacheng and Habahe region of Xinjiang's Junggar Basin, the Altai, Manrak, and Talbagatai Mountains in Kazakhstan, the Altai region of Russia, and northwest Mongolia (Kizaibek et al. 2011). The medicinal component of *D. altaica* is the stem bark, which the Kazakhs refer plant as "Wusuoyihe" or "Hasherjidek". It has been part of traditional Kazakh medicine for centuries and is used in the treatment of cancer and respiratory conditions (Kizaibek et al. 2020). It is mostly used to treat malignant tumors such as stomach cancer and esophageal cancer, as well as colds, bronchitis, pharyngeal swelling and pain, rheumatoid arthritis, toothaches, and snake bites which are poisonous (Murat et al. 2016). Studies have shown that ethanol extracts from different parts of *D. altaica* can inhibit human esophageal cancer cells (Eca-109), human gastric cancer cell lines (AGS), cervical cancer cells (Hela), liver cancer cells (SMMC-7721) four types of cancer cells proliferation in vitro (Kizaibek et al. 2011). The potential role of *D. altaica* on human esophageal cancer cells Eca-109 was investigated at the cellular level by Kizaibek M et al. (Kizaibek et al. 2011; Kizaibek et al. 2020). It was concluded that *D. altaica* may induce apoptosis, cause cell cycle arrest and upregulate intracellular PPARγ gene expression to accomplish these although the theoretical mechanism of action is still unclear.

Within the current study, we had examined and verified the hub genes and mechanism of the Kazakh traditional medicine *D. altaica* treating ESCC using network pharmacology in combination with bioinformatics Analysis and molecular docking, with the hope that this study will make a contribution to research on the therapies of ESCC.

Methods

Screening for *D. altaica* and ESCC targets

Through literature review learning about *D. altaica.* active compounds, we used the TCMSP Platform (Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform) (https://old.tcmsp-e.com/tcmsp.php), BATMAN-TCM Database (Bioinformatics Analysis Tool of Molecular Mechanism of Traditional Chinese Medicine) (http://bionet.ncpsb.org.cn/batman-tcm/) PubChem Database (https://pubchem.ncbi.nlm.nih.gov/), and SwissTargetPrediction Database (http://www.swisstargetprediction.ch/) to collect and predict targets of *D. altaica* active compounds. In clinical treatment, TCMs are often administered orally. ADME-related models such as oral bioavailability

(OB) (Xu et al. 2012) and drug-likeness (DL) (Walters and Murcko 2002) mostly affect the absorption of drugs through the gastrointestinal tract. The further characterization of bioactive components was performed following standards such as OB≥20% and DL≥0.18, and the related targets for each component were determined. And BATMAN-TCM under the standards of Score≥20. Our next step was to search for targets of disease using the search term "Esophageal squamous cell carcinoma (ESCC)" by The Comparative Toxicogenomics Database (CTD) (http://ctdbase.org/), Therapeutic Target Database (TTD) (http://db.idrblab.net/ttd/), and GeneCards Database (https://www.genecards.org/) under the standards of Score≥20.

Network Construction

The compound-target network was constructed by Cytoscape 3.9.1 to import the active compounds and targets of the drug. And we got the interaction of proteins using STRING (https://cn.string-db.org/). The protein-protein interaction (PPI) network was then generated based on Cytoscape 3.9.1

Differentially Expressed Genes Screening

From the GEO database, we obtained data on clinical samples of ESCC (https://www.ncbi.nlm.nih.gov/geo/) (series: GSE100942, GPL570). Clinical samples were collected from ESCC patients undergoing surgical tumor resection. For RNA extraction and hybridization on Affymetrix microarrays, paired adjacent non-tumor tissues were also collected from the proximal resection borders (>5 cm from the ESCC sample). Differential analysis was performed using R Studio, and the cut-off for identifying DEG was set at |log2 (Fold Change)|>2 and p.Value<0.05.

Functional Enrichment Analysis

Gene ontology analysis comprises three parts: Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) (The Gene Ontology Consortium 2017), which were used for the description of gene functions. The bioinformation and system function of the hub genes were identified using the Kyoto Encyclopedia of Genes and Gnomes enrichment analysis, which resulted in enrichment of biological signaling pathways (Ogata et al. 1999). We inputted the results of the analysis into R and used the R package Bioconductor ClusterProfiler to perform both of the above enrichment analysis by gene cluster (Yu et al. 2012).

Screening hub genes by prognostic analysis

For prognostic analysis, the KM Plotter database (http://kmplot.com/analysis/) was used to perform survival analysis and for the screening of hub genes, p<0.05 was considered statistically relevant (Nagy et al. 2021).

Validation analysis of hub genes

The TIMER2.0 database (http://timer.comp-genomics.org/) was used to perform Pan-cancer analysis on hub genes. Using this database, the expression differences between tumour and surrounding normal tissue were investigated for hub gene from each TCGA tumor. The results were shown as box plots showing the levels of gene expression. The stars were annotations of the statistic significances calculated by the Wilcoxon signed rank test. And for each type of cancer, it can be used to determine which genes are upregulated or downregulated in tumors as compared to normal tissue.

GEPIA2 database (http://gepia2.cancer-pku.cn/) Box Plot was used to examine the expression levels of hub genes. The differential analysis was performed on the chosen data sets (TCGA tumors vs. TCGA normal + GTEx normal). The method was one-way ANOVA, with disease status ("Tumor" or "Normal") as the variables used to calculate expression differences: Gene expression ~ Disease status. Data were initially log2(TPM+1) converted for differential analysis, and log2FC was expressed as the median (Tumor) - median (Normal). Genes were considered to be differentially expressed genes if |log 2FC|>1 and p value<0.01.

Molecular docking

Molecular docking was then used to determine whether core drug compounds discovered by network pharmacology were associated with hub genes. Docking was performed as described in the following:

1. On the basis of the results of the previous screening from Network Pharmacology and Bioinformatics analysis, we got five hub genes (TOP2A, NUF2, CDKN2A, BCHE, NEK2). The 3D structures of the target proteins were obtained using the Protein Database (PDB) (https://www.rcsb.org/) (Berman et al. 2000) and were downloaded in PDB format. Next, the protein was configured in AutoDock4.2.6 to remove the water, and add hydrogen and the structure was saved as PDBQT format (Morris et al. 2009).

2. In the above screening outcomes, we got the two most active compounds of the drug (Apigenin and Luteolin). Using the TCMSP and PubChem databases, we obtained the protein structures of the compounds (Kim et al. 2019). In the same process as in the previous step, the protein was modified in AutoDock4.2.6 to remove the water and add hydrogen, and the ligand files were saved in PDBQT format.

3. To perform molecular docking, we imported the above two structure files into AutoDock4.2.6. The minimum binding energy was calculated using the PDBQT format. The PDBQT format was then converted to the PDB format using the OpenBabel software (O'Boyle et al. 2011).

4. Finally, PyMOL2.5 software was used to visualize the molecular docking maps.

Results

Introduction to the present study

The present study aimed to obtain the hub genes of the Kazakh drug *D. altaica* for the treatment of ESCC and to investigate their potential molecular mechanisms. The flow chart of this study was shown in Fig. 1. First, we obtained targets of *D. altaica* and ESCC using databases such as TCMSP and GeneCards. Using the limma package in R, we analyzed 127 differentially expressed genes on the basis of the gene expression profiles of the ESCC and healthy patients in the GEO database. Then, 12 overlapping genes were obtained by Cross-tabulation analysis as the core genes for treatment. Survival analysis were performed to screen for five hub genes, and hub gene expression was validated using publicly available databases. Finally, to further investigate the interactions of the hub genes with the compounds, molecular docking was carried out using AutoDock software.

D. altaica and ESCC targets collection

We obtained ethanol and chloroform extraction and analytically identified 21 compounds from the studies of Shi Lei, Cong Jing et al. (Shi 2015; Cong 2015), of which 13 were active compounds. Then we obtained 165, 153, and 512 compound targets using the TCMSP, BATMAN-TCM, and SwissTargetPrediction three databases respectively (Table 1). And we used the CTD, TTD, and GeneCards databases to obtain 15912, 3, and 1795 disease targets respectively.

Table 1 The active components of *D. altaica*

No.	Component	TCMSP	BATMAN-TCM	SwissTargetPrediction	Total
1	Daphnoretin	12	0	99	111
2	β-sitosterol	3	34	44	81
3	Daphnetin	0	3	36	39
4	Daphnogitin	0	0	24	24
5	Daucosterol	17	0	2	19
6	Luteolin	56	0	100	156
7	Apigenin	77	3	100	180
8	Daphnodorin H	0	3	0	3
9	Daphnodorin G	0	3	0	3
10	4-Hydroxybenzoic acid	0	75	12	87
11	Umbelliferone	0	32	64	96
12	Dotriacontane	0	0	1	1
13	2-Hydroxydiplopterol	0	0	30	30

Compound-target network and PPI network construction

Based on 13 active drug compounds and 830 targets, we constructed a compound-target network using Cytoscape 3.9.1 (Fig. 2a). This network consisted of one drug, 13 compounds, and 830 compound targets, and the compounds were sorted by Degree value. The higher the value of the degree, the more biological functions it is involved in, and the stronger its biological importance (Li et al. 2009) (Table 2). Meantime, we got 409 overlapping genes of Drug-Disease by the Venn diagram (Fig. 2b). And then, the STRING database was used to obtain the PPI information. After that, we put it into Cytoscape 3.9.1 to obtain the PPI network diagram (Fig. 2c) and set the protein size according to the Degree value (Table 3).

Table 2 The information of 13 active compounds

NO.	Name	Betweenness	Degree
1	Apigenin	0.011562166	181
2	Luteolin	0.008207579	157
3	Daphnoretin	0.009092787	112
4	Umbelliferone	0.007661993	97
5	4-Hydroxybenzoic acid	0.008251402	88
6	β-sitosterol	0.007447578	82
7	Daphnetin	0.001622046	40
8	2-Hydroxydiplopterol	0.001725654	31
9	Daphnogitin	0.001382276	25
10	Daucosterol	0.001636758	20
11	Daphnodorin G	1.97E-04	4
12	Daphnodorin H	1.97E-04	4
13	Dotriacontane	4.38E-05	2

Table 3 The information of the top 20 genes of PPI

NO.	Gene name	Degree	NO.	Gene name	Degree
1	AKT1	199	11	CASP3	144
2	TP53	194	12	JUN	142
3	TNF	172	13	HIF1A	137
4	IL6	165	14	IL1B	135
5	INS	159	15	ESR1	134
6	EGFR	155	16	ERBB2	122
7	SRC	153	17	CCND1	121
8	HSP90AA1	150	18	MTOR	121
9	VEGFA	148	19	PTGS2	116
10	STAT3	145	20	PPARG	115

Acquisition of Differentially Expressed Genes

We obtained RNA extraction and hybridization data from ESCC patients' cancer and paracancerous tissues through the GEO database and obtained DEGs of micro RNA data using R Studio. We obtained 127 DEGs, including 82 ups and 45 downs (Supplementary Table1). A volcano plot (Fig. 3a), a heatmap (Fig. 3b), and a PCA plot (Fig. 3c-d) were generated to show the distribution of DEGs. A GSEA plot (Fig. 3e-f) showed that DEGs were closely related to DNA replication, Fanconi anemia pathway, Glycosaminoglycan biosynthesis-keratan sulfate, Homologous recombination, Mismatch repair, Cell cycle, Nucleocytoplasmic transport, Proteasome and Ribosome biogenesis in eukaryotes.

After the above analyses, we screened the genes for *D. altaica* treatment of ESCC and obtained 12 overlapping genes as the core genes for treatment, including BCHE, CDC20, CDKN2A, CXCR2, IL6, MMP1, MMP3, MMP12, NEK2, NUF2, PTGS2, and TOP2A (Fig. 4a).

Functional Enrichment Analysis

GO analysis and KEGG pathway enrichment analysis were carried out with the R package Bioconductor ClusterProfiler to clarify the signaling pathways of the 12 overlapping genes (Supplementary Table2). GO analysis results showed that the 12 overlapping genes in BP terms were mainly enriched in the positive regulation of cell death, cytokine-mediated signaling pathway, extracellular matrix disassembly, collagen metabolic process, and inflammatory cell apoptotic process (Fig. 4b). The CC terms included chromosome, endoplasmic reticulum lumen, extracellular matrix, kinetochore, spindle pole, centromeric region, and senescence-associated heterochromatin focus (Fig. 4c). In MF terms, they were enriched in enzyme binding, metallopeptidase activity, endopeptidase activity, metalloendopeptidase activity, histone deacetylase binding, serine-type endopeptidase activity, DNA topoisomerase activity (Fig. 4d). And the result of KEGG analysis indicated that mainly enriched in the IL-17 signaling pathway, Pathways in cancer, Platinum drug resistance, Cell cycle, Viral carcinogenesis, MicroRNAs in cancer, and p53 signaling pathway (Fig. 4e, Table 4).

Table 4 The information of top 10 significant KEGG enrichment analysis

Pathway ID	Pathway	Count	P.Value
hsa04657	IL-17 signaling pathway	4	1.21E-06
hsa04668	TNF signaling pathway	4	0.000146747
hsa05166	Human T-cell leukemia virus 1 infection	3	0.001056123
hsa05200	Pathways in cancer	4	0.001120432
hsa01524	Platinum drug resistance	2	0.002267077
hsa04110	Cell cycle	2	0.00641068
hsa05203	Viral carcinogenesis	2	0.016248416
hsa05206	MicroRNAs in cancer	3	0.036625941
hsa04370	VEGF signaling pathway	1	0.058132961
hsa04115	p53 signaling pathway	1	0.070536757

Screening hub gene by prognostic analysis

We performed survival analysis of 12 overlapping genes (BCHE, CDC20, CDKN2A, CXCR2, IL6, MMP1, MMP3, MMP12, NEK2, NUF2, PTGS2, and TOP2A) through the Kaplan-Meier plotter database, and obtained five significant genes ($p<0.05$) and used these as hub genes, including TOP2A, NUF2, CDKN2A, BCHE, and NEK2 (Fig. 5a).

Validation analysis of hub genes

The next five hub genes were validated. First, we explored hub genes' expression in multiple tumor types for pan-cancer analysis using the TIMER2.0 database. The pan-cancer analysis results revealed that the hub genes NEK2, NUF2, TOP2A, and CDKN2A were up-regulated in most tumors with high significance (*p*.value <0.001), while BCHE was down-regulated in tumors including ESCC (Fig. 5b). Similarly, Box Plot results in the GEPIA2 database revealed that the mRNA levels of NEK2, NUF2, TOP2A and CDKN2A were up-regulated considerably in ESCC samples compared to the normal group (*p*.value <0.01), while the mRNA levels of BCHE were not significantly down-regulated (Fig. 5c).

Molecular docking

We used the five hub genes NEK2 (PDB:6SGH), CDKN2A (PDB:1DC2), BCHE (PDB:4AQD), TOP2A (PDB:4R1F), and NUF2 (PDB:2VE7) with the two most active compounds of D. altaica for molecular docking using AutoDock 4.2.6. The docking scores were shown in Table 5. The more stable the ligand-receptor binding, the lower the binding energy will be (Hsin et al. 2013). It was found that NEK2-Apigenin

(PubChem CID:5280443), CDKN2A-Apigenin, BCHE-Apigenin, TOP2A-Apigenin, NUF2-Apigenin, NEK2-Luteolin (PubChem CID:5280445), CDKN2A-Luteolin, BCHE-Luteolin, TOP2A-Luteolin, and NUF2-Luteolin all have good binding affinities. Finally, PyMOL2.5 software was used to visualize the molecular docking maps, the visualization results were shown in Fig. 6.

Table 5 Docking scores of hub genes with compounds (kcal·mol-1)

Gene name	PDB ID	Compounds	
		Apigenin	Luteolin
NEK2	6SGH	-4.22	-3.64
CDKN2A	1DC2	-4.04	-3.38
BCHE	4AQD	-3.90	-2.54
TOP2A	4R1F	-3.11	-2.53
NUF2	2VE7	-2.79	-2.08

Discussion

In 2020, Esophageal cancer ranked as the seventh most common cancer, accounting for one in eighteen all-cancer deaths worldwide (Sung et al. 2021). ESCC is the most commonly occurring histological sub-type of EC, a common digestive system cancer, which has been a worldwide concern in developing countries such as China or India due to its high morbidity and mortality (Alsop and Sharma 2016). Chemotherapeutic agents, including docetaxel and paclitaxel, for the treatment of ESCC, have limited therapeutic potential and numerous side effects in clinical practice (Nakajima and Kato 2013). Although several treatment options have been developed, improvements to these existing treatment modalities are required because of poor prognosis and low survival rates related to late diagnosis (Tustumi et al. 2016). Consequently, it is needed to identify new potential targets of ESCC by exploring the molecular mechanisms that lead to ESCC in order to improve clinical outcomes.

TCM is also widely used as a medical treatment for ESCC. *Daphne altaica.* (*D. altaica*) is a medicinal herb used to treat many ailments, and is historically used to treat cancer and diseases of the respiratory system in traditional Kazakh medicine (Kizaibek et al. 2020). However, the complexity of the ingredients and targets limits the research into the mechanisms of Traditional Chinese Medicines (Liu et al. 2022).

In the present study, we identified 13 active compounds in *D. altaica* with 830 targets. Based on the data analyzed, it was known that the compounds Apigenin and Luteolin had higher degree values. By analyzing the PPI network diagram we obtained that AKT1 and TP53 were the two proteins with high relevance to the treatment. The 127 DEGs of the ESCC were obtained by analyzing GEO data, and 12 overlapped genes were identified by target screening analysis with drugs, diseases, and DEGs as the core genes for treatment, including BCHE, CDC20, CDKN2A, CXCR2, IL6, MMP1, MMP3, MMP12, NEK2, NUF2,

PTGS2, and TOP2A. Gene functional enrichment analysis revealed that they were mainly enriched in the IL-17 signaling pathway, Pathways in cancer, Cell cycle, MicroRNAs in cancer, and p53 signaling pathway. After survival analysis, we found that the levels of BCHE, CDKN2A, NEK2, NUF2, and TOP2A had a significant relationship with the prognosis, and these five genes were considered to be the hub genes in the D. altaica treatment of ESCC. Besides, molecular docking revealed that BCHE, CDKN2A, NEK2, NUF2, and TOP2A had good binding affinities with Apigenin and Luteolin.

AKT1, phospho-form specific substrates of protein kinase B (McKenna et al. 2021), protein kinase B (AKT) is one of the AGC family of serine-threonine kinases (Manning and Toker 2017). AKT1 is an important regulator of the phosphoinositide 3-kinase (PI3K)/AKT signal transduction pathway, which regulates the proliferation and survivability of cells (Burgering and Coffer 1999). The study by Alwhaibi A et al. showed that AKT1 has a critical role in modulating tumor angiogenesis and cancer metastasis (Alwhaibi et al. 2019). Ou R et al. demonstrated that gain-of-function and loss-of-function tests indicated that circ-AKT1 promotes CC cell growth and migration (Ou et al. 2020). TP53, the p53 tumor suppressor protein, is an important transcription factor that acts as a failsafe in the cells' defence against tumors by inhibiting the proliferation or viability of cells under a variety of stresses (Kastenhuber and Lowe 2017; Vousden and Prives 2009; Macheret and Halazonetis 2015). From a clinical point of view, TP53 mutation has been associated with a worse prognosis in several types of cancer. However, this remains controversial (Robles and Harris 2010). The TP53 family of transcription factors, which includes TP53, TP63, and TP73, is involved in biologically and pathologically relevant development of cancer and neurons (Agostini et al. 2018). The study by Yao L et al. found that TP53 is up-regulated in ESCC tumors and has a critical function in ESCC cell growth, survival and migratory processes. It could be an activator of the mTOR pathway and an inhibitor of TP53 dependent autophagy. As a result, it could shed new light on the possible value of TP53 as a biomarker for diagnosis and as well as a useful therapeutic target in ESCC (Yao et al. 2021).

Butyrylcholinesterase (BChE, EC 3.1.1.8), also described as plasma cholinesterase or pseudocholinesterase, is a serine hydrolase that is found in most of the tissue of mammals, the most abundant being in the plasma and liver (Johnson and Moore 2012; Lockridge 2015). Clinical studies have shown that its levels in the serum are decreased in conditions including liver damage, inflammation, infection, and malignant tumor (Lampón et al. 2012). In the study of Lampón N. et al., BCHE as a biomarker for EC may be a predictor of its prognosis. It may also have significant immunotherapy implications (Liu et al. 2022).

The CDKN2A gene also known as p16INK4 is thought to help suppress tumors (Foulkes et al. 1997). The CDKN2A protein is critical for inhibiting cell cycle progression, and the down-regulation of p16 could lead to increased cell proliferation and may contribute to the pathogenesis of several cancers (Conceição et al. 2015). In a study by Chen Z et al, they analyzed CDKN2A levels in 33 tumors and found that CDKN2A may be a promising prognostic biomarker with potential molecular mechanisms to influence the survival outcomes of cancer patients (Chen et al. 2021)

Never in mitosis-related kinase 2 (NEK2) is a serine/threonine kinase that facilitates the division of centrosomes and guarantees the accurate segregation of chromosomes in the G2/M phase of the cell cycle (Xia et al. 2015). The prognostic significance of increased NEK2 expression and its role in ESCC in proliferation, migration, and EMT was demonstrated in a study by Su W et al. They concluded that NEK2 interacts with YAP1 and phosphorylates it at Thr-143, thereby attenuating its ubiquitylation in vitro and promoting its stability, thereby mediating the progression of ESCC (Su et al. 2022). Previous studies have indicated that NEK2 inhibitors efficiently inhibit cancer cell proliferation by causing arrest of the cellular cycle and apoptotic cell death, and markedly inhibition of tumor growth in-vivo (Xi et al. 2017; Fang et al. 2016).

NUF2, also named CDCA1, is a critical component of the Ndc80/NUF2 composite and is necessary for the establishment of a stable kinetochore-microtubule attachment and chromosome alignment throughout mitosis (Xie et al. 2021). In the study by Zheng L et al, NUF2 was identified as one of the six novel markers of ESCC and could also be used as one of the potential putative markers for ESCC treatment (Zheng et al. 2021).

TOP2A is an isoform of Topoisomerase II (TOP2), a core protein necessary for DNA replication and cell division (Wang et al. 2022). TOP2A has been identified in numerous studies as a biomarker for prognosis and a promising target for therapy in bladder cancer, including bladder urothelial carcinoma (BLCA), lung adenocarcinoma, prostate cancer, colorectal cancer, and breast cancer (Zeng et al. 2019; Du et al. 2020; de Resende et al. 2013; Zhang et al. 2018; Badawy and Loay 2019). Yang W et al. identified TOP2A as 1 of the top 10 candidate genes predicted to be implicated in the treatment of ESCC, and it was also found to be the top three hub genes targeted by the majority of miRNAs (Yang et al. 2019).

Among all polyphenols, the five most common plant polyphenols are kaempferol, quercetin, prunetin, apigenin and lignan (Abid et al. 2022). A study showed that a significant reduction in the risk and potential of ovarian cancer was observed due to regular intake of apigenin. After a course of medical trials, it was postulated that apigenin together with lignocaine is an effective chemotherapeutic agent (Wang et al. 2019). In a study by Rahmani AH et al, apigenin was demonstrated to play an essential role in cancer inhibition through a multitude of molecular interactions and processes that modulate apoptotic mechanisms, abnormal cell signaling, and oncogenic protein networks to inhibit carcinogenesis (Rahmani et al. 2022). Apigenin causes apoptosis in esophageal cancer cells through disruption of membrane structure, according to Zhu H et al. (Zhu et al. 2016) Chen P et al. found Luteolin could cause G2/M arrest of the cellular cycle and mitochondrially mediated apoptotic cell death in EC cells, while in vivo studies also demonstrated that lignocaine could markedly inhibit tumor progression in a mouse model of ESCC (Chen et al. 2017). According to the above expressions, Apigenin and Luteolin have some therapeutic effects on ESCC as well as other cancers.

The treatment targets and drugs revealed by this study are both promising and innovative for personalized treatment, although this study was merely a prediction that needed to be validated by further experiments in-vivo and in-vitro.

Conclusion

To conclude, we identified five hub genes and two active compounds for the *D. altaica* treatment of ESCC by network pharmacology and validated their effects by bioinformatics analysis and molecular docking. The five hub genes may act as potential novel targets for detecting, prognosticating, and targeting ESCC. These findings predicted new therapeutic targets for the Kazakh traditional drug *D. altaica* treatment of ESCC.

Declarations

Supplementary Information The Supplementary Material for this article can be found online.

Acknowledgments We are grateful to all the researchers who provided the data we used.

Author Contributions Sendaer Hailati, Ziruo Talihati, and Wenting Zhou designed the study. Sendaer Hailati, Kayisaier Abudurousuli, Meng yuan Han performed bioinformatic analysis. Jimilihan Simayi, Dilihuma Dilimulati, and Nuerbiye nueraihemaiti contributed to the conception of the study and drafted the manuscript. Muhadaisi Nuer, Nawaz Khan, and Nulibiya Maihemuti contributed to the writing of the manuscript. All authors contributed to the article and approved the submitted version.

Funding The authors appreciate financial support from the funding of Tianshan Talents-Youth Science and Technology Innovation Talents Training Program of Xinjiang Autonomous Region (2022TSYCCX0035); National Natural Science Foundation of China (81660696); "Fourteenth Five-Year Plan" Key Discipline Construction Project of Xinjiang Autonomous Region (2021); The key Laboratory of Xinjiang Autonomous Region (XJDX1713).

Data Availability Statement The datasets (GSE100942, GPL570) for this study can be found in the [The Gene Expression Omnibus (GEO) database] [https://www.ncbi.nlm.nih.gov/geo/]. All the data in this paper support the results of this study, Other datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflict of Interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Alsop BR, Sharma P (2016) Esophageal cancer. Gastroenterol Clin North Am 45:399–412. https://doi:10.1016/j.gtc.2016.04.001
2. Agostini M, Melino G, Bernassola F (2018) The p53 Family in Brain Disease. Antioxid Redox Signal 29(1):1–14. https://doi:10.1089/ars.2017.7302
3. Alwhaibi A, Verma., Adil MS, Somanath PR (2019) The unconventional role of Akt1 in the advanced cancers and in diabetes-promoted carcinogenesis. Pharmacol Res 145:104270.

https://doi:10.1016/j.phrs.2019.104270

4. Abid R, Ghazanfar S, Farid A, Sulaman SM, Idrees M, Amen RA, Muzammal M, Shahzad MK, Mohamed MO, Khaled AA, Safir W, Ghori I, Elasbali AM, Alharbi B (2022) Pharmacological Properties of 4', 5, 7-Trihydroxyflavone (Apigenin) and Its Impact on Cell Signaling Pathways. Molecules 27(13):4304. https://doi:10.3390/molecules27134304
5. Burgering BM, Coffer PJ (1995) Protein kinase B (c-Akt) in phosphatidylinositol-3-OH kinase signal transduction. Nature 376, 599–602. https://doi:10.1038/376599a0
6. Berman HM, Westbrook J, Feng Z et al (2000) The protein data bank. Nucleic Acids Research 28(1):235–242. https://doi:10.1093/nar/28.1.235
7. Badawy OM, Loay I (2019) FISH Analysis of TOP2A and HER-2 Aberrations in Female Breast Carcinoma on Archived Material: Egyptian NCI Experience Appl Immunohistochem Mol Morphol. https://doi:10.1097/PAI.0000000000000574
8. Cong J (2015) Research on the Chemical Constituents of the Chloroform Extraction Layer of Daphne Altaica Pall. Northeast Normal University https://kns.cnki.net/KCMS/detail/detail
9. Chen XC, Pang LJ, Li F (2006) Progress in the study of esophageal cancer in Kazakhs in Xinjiang. Journal of Nongken Medicine 28(5): 384–387.https://kns.cnki.net/kcms/detail/detail
10. Conceição AL, Babeto E, Candido NM, Franco FC, de Campos Zuccari DA, Bonilha JL, Cordeiro JA, Calmon MF, Rahal P (2015) Differential Expression of ADAM23, CDKN2A (P16), MMP14 and VIM Associated with Giant Cell Tumor of Bone. J Cancer 6(7):593–603. https://doi:10.7150/jca.11238
11. Chen P, Zhang JY, Sha BB, Ma YE, Hu T, Ma YC, Sun H, Shi JX, Dong ZM, Li P (2017) Luteolin inhibits cell proliferation and induces cell apoptosis via down-regulation of mitochondrial membrane potential in esophageal carcinoma cells EC1 and KYSE450. Oncotarget 8(16):27471–27480. https://doi:10.18632/oncotarget.15832
12. Chen Z, Guo Y, Zhao D, Zou Q, Yu F, Zhang L, Xu L (2021) Comprehensive Analysis Revealed that CDKN2A is a Biomarker for Immune Infiltrates in Multiple Cancers. Front Cell Dev Biol 9:808208. https://doi:10.3389/fcell.2021.808208
13. de Resende MF, Vieira S, Chinen LT et al (2013) Prognostication of prostate cancer based on TOP2A protein and gene assessment: TOP2A in prostate cancer. J Transl Med 11:36. https://doi:10.1186/1479-5876-11-36
14. Du X, Xue Z, Lv J, Wang H (2020) Expression of the topoisomerase II alpha (TOP2A) gene in lung adenocarcinoma cells and the association with patient outcomes. Med Sci Monit 26:e929120. https://doi:10.12659/MSM.929120
15. Foulkes WD, Flanders TY, Pollock PM, Hayward NK (1997) The CDKN2A (p16) gene and human cancer. Mol Med 3:5–20.
16. Fang Y, Kong Y, Xi J, Zhu M, Zhu T, Jiang T, Hu W, Ma M, Zhang X (2016) Preclinical activity of MBM-5 in gastrointestinal cancer by inhibiting NEK2 kinase activity. Oncotarget 7:79327–79341.
17. Hsin KY, Ghosh S, Kitano H (2013) Combining machine learning systems and multiple docking simulation packages to improve docking prediction reliability for network pharmacology. PloS one

8:e83922.

18. Han RH (2015) China Kazakh and Kazakhstan Kazakh cultural comparison diet. Urumqi: Xinjiang Normal University. https://kns.cnki.net/kcms/detail/detail
19. Johnson G, Moore SW (2013) Why has butyrylcholinesterase been retained? Structural and functional diversification in a duplicated gene. Neurochem. Int 61:783–797. https://doi:10.1016/j.neuint.2012.06.016
20. Kizaibek M, Daniar M, Li L, Upur H (2011) Antiproliferative activity of different extracts from Daphne altaica Pall. On selected cancer cells. J Med Plants Res 5:3448–3452.
21. Kastenhuber ER, Lowe SW (2017) Putting p53 in Context. Cell 170(6):1062–1078. https://doi:10.1016/j.cell.2017.08.028
22. Kizaibek M, Wubuli A, Gu Z et al (2020) Effects of an ethyl acetate extract of Daphne altaica stem bark on the cell cycle, apoptosis and expression of PPARγ in Eca-109 human esophageal carcinoma cells. Mol Med Rep 22(2):1400–1408. https://doi:10.3892/mmr.2020.11187
23. Li YL, Min W, Zhang W (2009) Influence of network topology on critical coupling parameters in synchronization of small world neutral networks. Far East J Dyn Syst 11(1): 65–75.
24. Lockridge O (2015) Review of human butyrylcholinesterase structure, function, genetic variants, history of use in the clinic, and potential therapeutic uses. Pharmacol 148:34–46. https://doi:10.1016/j.pharmthera.2014.11.011
25. Lampón N, Hermida-Cadahia EF, Riveiro A, Tutor JC (2012) Association between butyrylcholinesterase activity and low-grade systemic inflammation. Annals of Hepatology 11(3):356–363. https://doi:10.1016/S1665-2681(19)30932-9
26. Liu Z, Sun Y, Zhen H, Nie C (2022) Network Pharmacology Integrated with Transcriptomics Deciphered the Potential Mechanism of Codonopsis pilosula against Hepatocellular Carcinoma. Evid Based Complement Alternat Med 2022:1340194. https://doi:10.1155/2022/1340194
27. Liu J, Tian T, Liu X, Cui Z (2022) BCHE as a Prognostic Biomarker in Endometrial Cancer and Its Correlation with Immunity. J Immunol Res 2022:6051092. https://doi:10.1155/2022/6051092
28. Morris GM, Huey R, Lindstrom W et al (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. Journal of Computational Chemistry 30(16):2785–2791. https://doi:10.1002/jcc.21256
29. Macheret M, Halazonetis TD (2015) DNA replication stress as a hallmark of cancer. Annu Rev Pathol 10:425 – 48. https://doi:10.1146/annurev-pathol-012414-040424
30. Manning BD, Toker A (2017) AKT/PKB signaling: navigating the network. Cell 169, 381–405. https://doi:10.1016/j.cell.2017.04.001
31. Murat K, Xamxinur Y, Su JC (2016) Progress in the traditional application and anticancer activity of Kazakh medicine Daphne altaica and its congeners. Hebei Medical Journal 38(19):3007–3010. https://kns.cnki.net/kcms/detail/detail

32. McKenna M, Balasuriya N, Zhong S, Li SS, O'Donoghue P (2021) Phospho-Form Specific Substrates of Protein Kinase B (AKT1). Front Bioeng Biotechnol 8:619252. https://doi:10.3389/fbioe.2020.619252
33. NM O'Boyle, M Banck, CA James, C Morley, T Vandermeersch, G R Hutchison (2011) Open Babel: an open chemical toolbox, J. Cheminf 3,2011, 33.
34. Nakajima M, Kato H (2013) Treatment options for esophageal squamous cell carcinoma. Expert Opin Pharmacother 14:1345–1354. https://doi:10.1517/14656566.2013.801454
35. Nagy A, Munkacsy G, Gyorffy B (2021) Pancancer survival analysis of cancer hallmark genes, Sci Rep. 11(1):6047. https://doi.org/10.1038/s41598-021-84787-5
36. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 27(1):29–34. https://doi:10.1093/nar/27.1.29
37. Ou R, Mo L, Tang H, Leng S, Zhu H, Zhao L, Ren Y, Xu Y (2020) circRNA-AKT1 Sequesters miR-942-5p to Upregulate AKT1 and Promote Cervical Cancer Progression. Mol Ther Nucleic Acids 20:308–322. https://doi:10.1016/j.omtn.2020.01.003
38. Robles AI, Harris CC (2010) Clinical outcomes and correlates of TP53 mutations and cancer. Cold Spring Harb Perspect Biol 2(3):a001016. https://doi:10.1101/cshperspect.a001016
39. Rahmani AH, Alsahli MA, Almatroudi A, Almogbel MA, Khan AA, Anwar S, Almatroodi SA (2022) The Potential Role of Apigenin in Cancer Prevention and Treatment. Molecules 27(18):6051. https://doi:10.3390/molecules27186051
40. Shi L. (2015) Research on the Chemical Constituents of Ethyl Acetate Extract from Daphne Altaica. Northeast Normal University https://kns.cnki.net/KCMS/detail/detail
41. S Kim, J Chen, T Cheng, A Gindulyte, J He, S He, Q Li, BA Shoemaker, P A Thiessen, B Yu, L Zaslavsky, J Zhang, EE Bolton (2019) PubChem 2019 update: improved access to chemical data, Nucleic Acids Res 47 D1102-d1109.
42. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71:209–249.
43. Su W, Hu H, Ding Q, Wang M, Zhu Y, Zhang Z, Geng Z, Lin S, Zhou P (2022) NEK2 promotes the migration and proliferation of ESCC via stabilization of YAP1 by phosphorylation at Thr-143. Cell Commun Signal 20(1):87. https://doi:10.1186/s12964-022-00898-0
44. Tustumi F, Kimura MS, Takeda FR, Uema RH, Salum RA, Ribeiro-Junior U, Cecconello I (2016) Prognostic factors and survival analysis in esophageal carcinoma. Arq Bras Cir Dig 29:138–141. https://doi:10.1590/0102-6720201600030003
45. Tungekar A, Mandarthi S, Mandaviya PR et al (2018) ESCC ATLAS: A population wide compendium of biomarkers for Esophageal Squamous Cell Carcinoma. Sci Rep 8(1):12715. https://doi:10.1038/s41598-018-30579-3
46. The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and resources (2017). Nucleic Acids Res 45(D1):D331-D338. https://doi:10.1093/nar/gkw1108

47. Vousden KH, Prives C (2009) Blinded by the Light: The Growing Complexity of p53. Cell 137(3):413–31. https://doi:10.1016/j.cell.2009.04.037
48. Walters WP, Murcko MA (2002) Prediction of 'drug-likeness'. Adv Drug Deliv Rev 54(3):255–71. https://doi:10.1016/s0169-409x(02)00003-0
49. Wang M, Firrman J, Liu L, Yam K (2019) A Review on Flavonoid Apigenin: Dietary Intake, ADME, Antimicrobial Effects, and Interactions with Human Gut Microbiota. Biomed Res Int 2019:7010467. https://doi:10.1155/2019/7010467
50. Wang T, Lu J, Wang R, Cao W, Xu J (2022) TOP2A promotes proliferation and metastasis of hepatocellular carcinoma regulated by miR-144-3p. J Cancer 13(2):589–601. https://doi:10.7150/jca.64017
51. Xu X, Zhang W, Huang C, Li Y, Yu H, Wang Y, Duan J, Ling Y (2012) A novel chemometric method for the prediction of human oral bioavailability. Int J Mol Sci 13(6):6964–82. https://doi:10.3390/ijms13066964
52. Xia J, Franqui-Machin R, Gu Z, Zhan F (2015) Role of NEK2A in human cancer and its therapeutic potentials. Biomed Res Int 2015, 12.
53. Xi JB, Fang YF, Frett B, Zhu ML, Zhu T, Kong YN, Guan FJ, Zhao Y, Zhang XW, Li HY, Ma ML, Hu W (2017) Structure-based design and synthesis of imidazo[1,2-a] pyridine derivatives as novel and potent Nek2 inhibitors with in vitro and in vivo antitumor activities. Eur J Med Chem 126:1083–1106.
54. Xie X, Jiang S, Li X (2021) Nuf2 Is a Prognostic-Related Biomarker and Correlated With Immune Infiltrates in Hepatocellular Carcinoma. Front Oncol 11:621373. https://doi:10.3389/fonc.2021.621373
55. Yu G, Wang LG, Han Y, He QY (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16(5):284–7. https://doi:10.1089/omi.2011.0118
56. Yang W, Zhao X, Han Y, Duan L, Lu X, Wang X, Zhang Y, Zhou W, Liu J, Zhang H, Zhao Q, Hong L, Fan D (2019) Identification of hub genes and therapeutic drugs in esophageal squamous cell carcinoma based on integrated bioinformatics strategy. Cancer Cell Int 19:142. https://doi:10.1186/s12935-019-0854-6
57. Yao L, Zhong X, Huang G, Ma Q, Xu L, Xiao H, Guo X (2021) Investigation on the Potential Correlation Between TP53 and Esophageal Cancer. Front Cell Dev Biol 9:730337. https://doi:10.3389/fcell.2021.730337
58. Zhang HX, Chen Y, Yin D et al (2009) Exploration of the risk factors of esophageal cancer in Xinjiang Kazakh nationality. Modern Preventive Medicine 36(10): 1804–1806. https://kns.cnki.net/kcms/detail/detail
59. Zheng S, Vuitton L, Sheyhidin I, Vuitton DA, Zhang Y, Lu X (2010) Northwestern China: a place to learn more on oesophageal cancer. Part one: behavioural and environmental risk factors. European Journal of Gastroenterology & Hepatology 22(8):917–925. https://doi:10.1097/MEG.0b013e3283313d8b

60. Zheng S, Vuitton L, Sheyhidin I, Vuitton DA, Zhang Y, Lu X (2011) Northwestern China: a place to learn more on oesophageal cancer. Part two: gene alterations and polymorphisms. European Journal of Gastroenterology and Hepatology 23(12):1087–1099. https://doi:10.1097/MEG.0b013e32834a14d9
61. Zhu H, Jin H, Pi J, Bai H, Yang F, Wu C, Jiang J, Cai J (2016) Apigenin induced apoptosis in esophageal carcinoma cells by destruction membrane structures. Scanning 38(4):322–8. https://doi:10.1002/sca.21273
62. Zhang R, Xu J, Zhao J, Bai JH (2018) Proliferation and invasion of colon cancer cells are suppressed by knockdown of TOP2A. J Cell Biochem 119(9):7256–7263. https://doi:10.1002/jcb.26916
63. Zeng S, Liu A, Dai L et al (2019) Prognostic value of TOP2A in bladder urothelial carcinoma and potential molecular mechanisms. BMC Cancer 19(1):604. https://doi:10.1186/s12885-019-5814-y
64. Zheng L, Li L, Xie J, Jin H, Zhu N (2021) Six Novel Biomarkers for Diagnosis and Prognosis of Esophageal squamous cell carcinoma: validated by scRNA-seq and qPCR. J Cancer 12(3):899–911. https://doi:10.7150/jca.50443

Figures

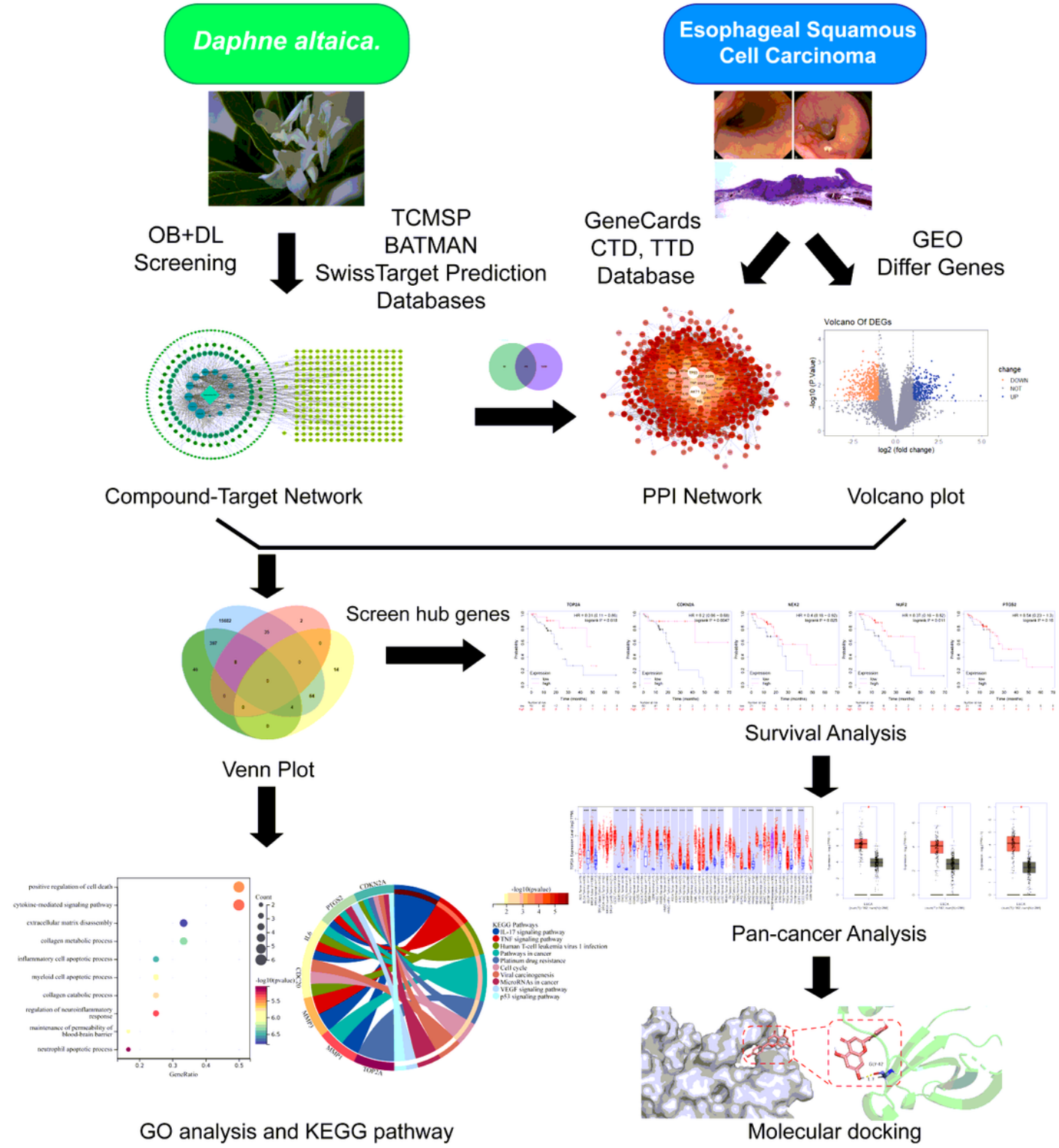

Figure 1

The flow chart of this study

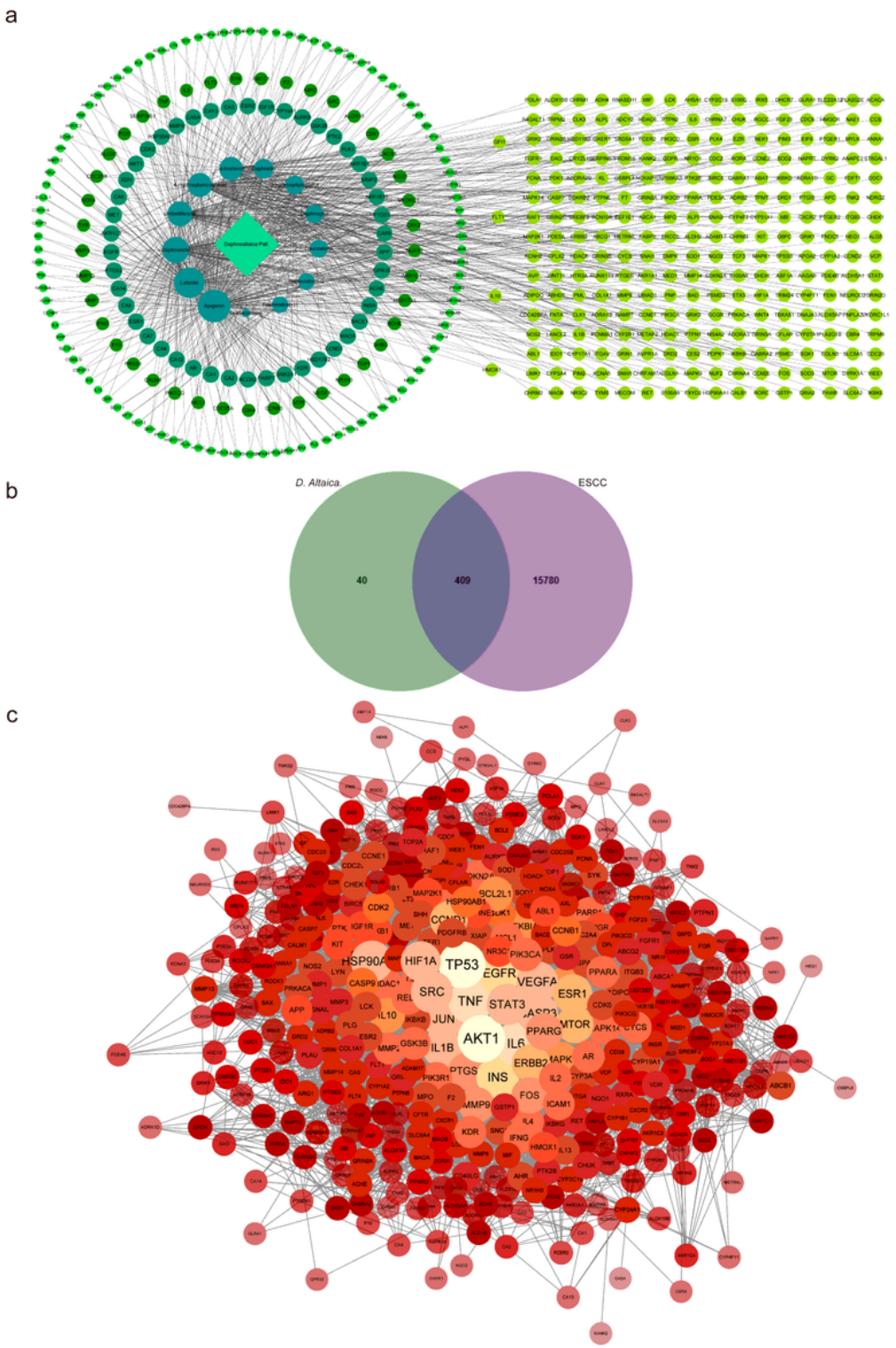

Figure 2

a Compound-target network, the size of the compounds was sorted by Degree value, and the color of the targets was sorted by Degree value. **b** Venn diagram, screening for overlapping genes of *D. altaica*and ESCC. **c** PPI network, the protein size according to the Degree value, AKT1 and TP53 were the two genes with high relevance to the treatment

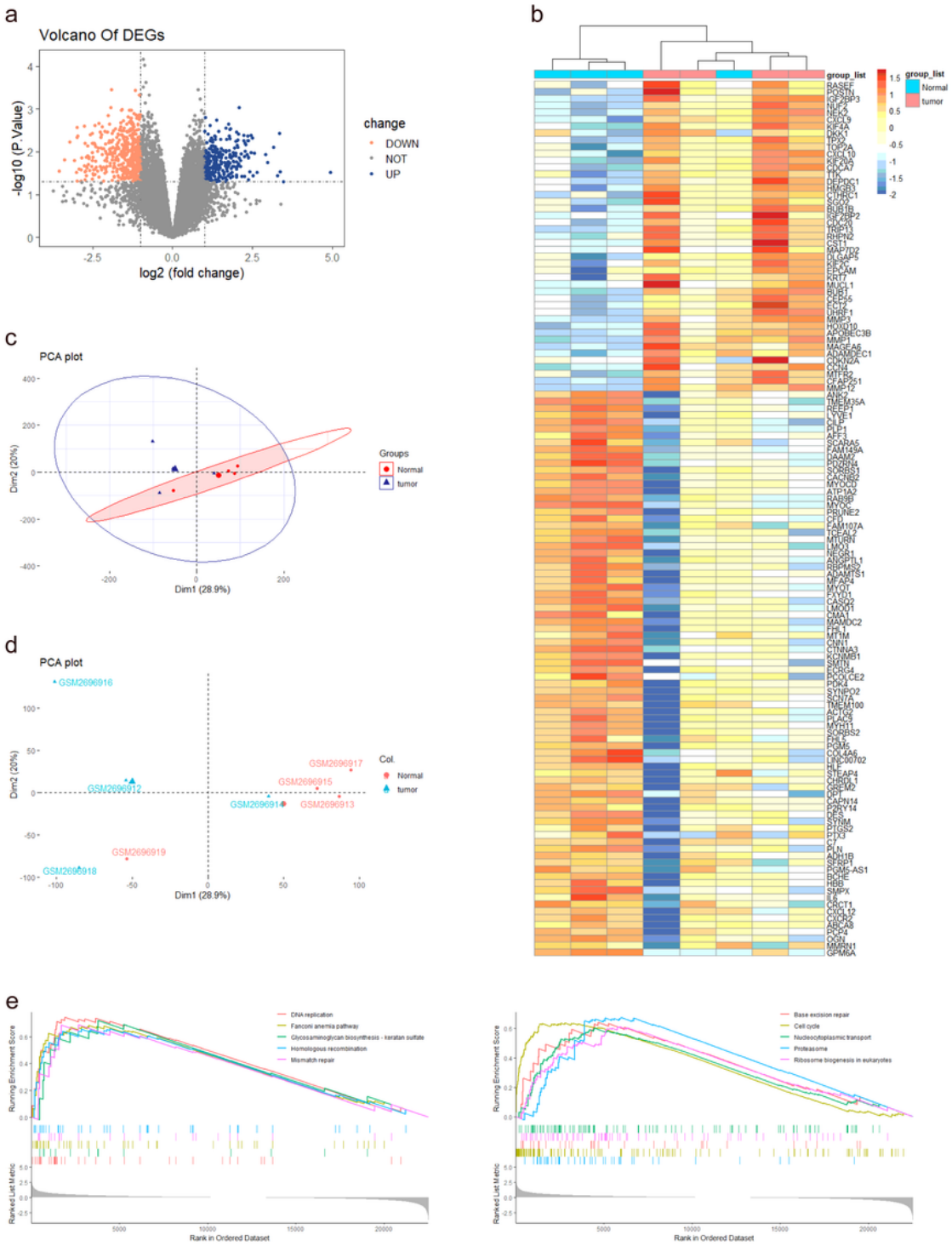

Figure 3

a Volcano plot of DEGs, the red color represented up-regulated genes and the blue represented up-regulated genes. **b** Heatmap, 127 DEGs in normal and tumor groups. **c-d**PCA plots, distribution of DEGs in normal and tumor groups. **e-f** GSEA plots, closely related to the DNA replication, Fanconi anemia pathway,

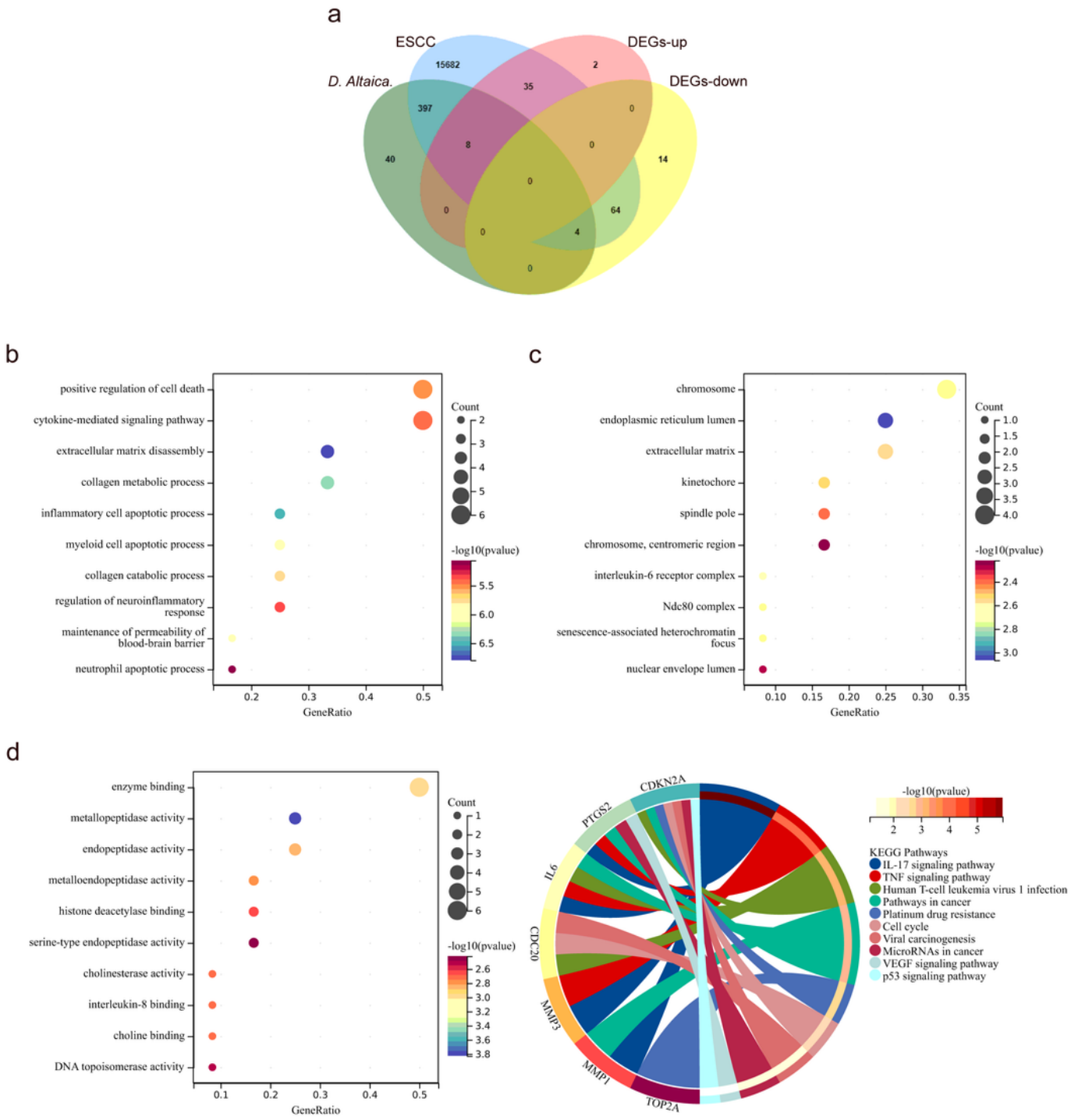

Figure 4

aVenn diagram, screening for overlapping genes of *D. altaica*, ESCC and DEGs. The green color represented *D. altaica*, the blue color represented ESCC, the pink color represented up-regulated genes, and the yellow color represented down-regulated genes. **b** GO analysis - BP terms. **c** GO analysis - CC terms. **d** GO analysis - MF terms. **e** KEGG pathway plot

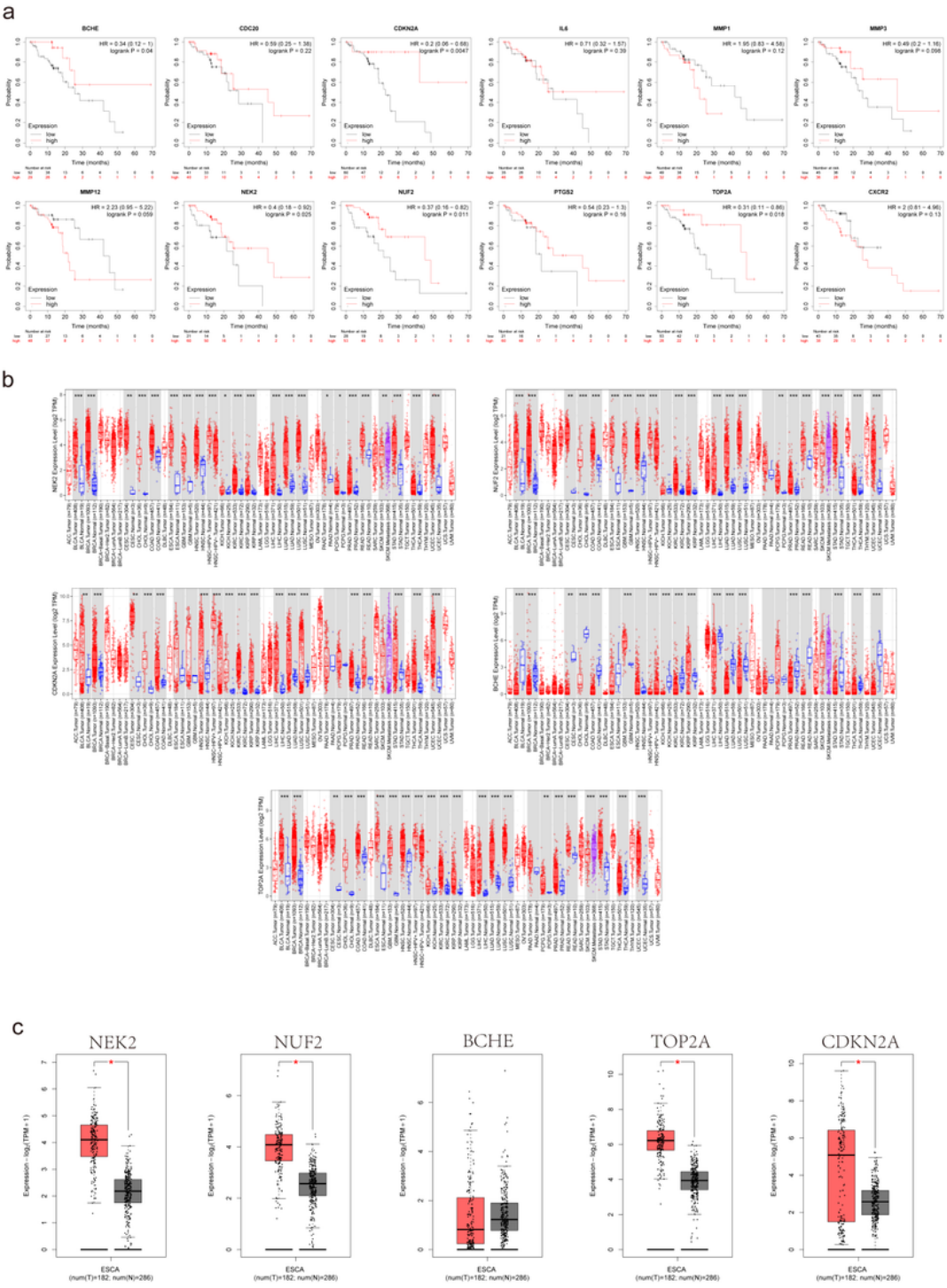

Figure 5

a Screening hub genes by prognostic analysis, and p < 0.05 was significant in survival analysis. **b** Pan-cancer analysis. The statistical significance computed by the Wilcoxon test is annotated by the number of stars (*: p.value<0.05; **: p.value<0.01; ***: p.value<0.001). **c** The expression level of hub genes in tumor tissues (n=182) and normal tissues (n=286) in the TCGA database

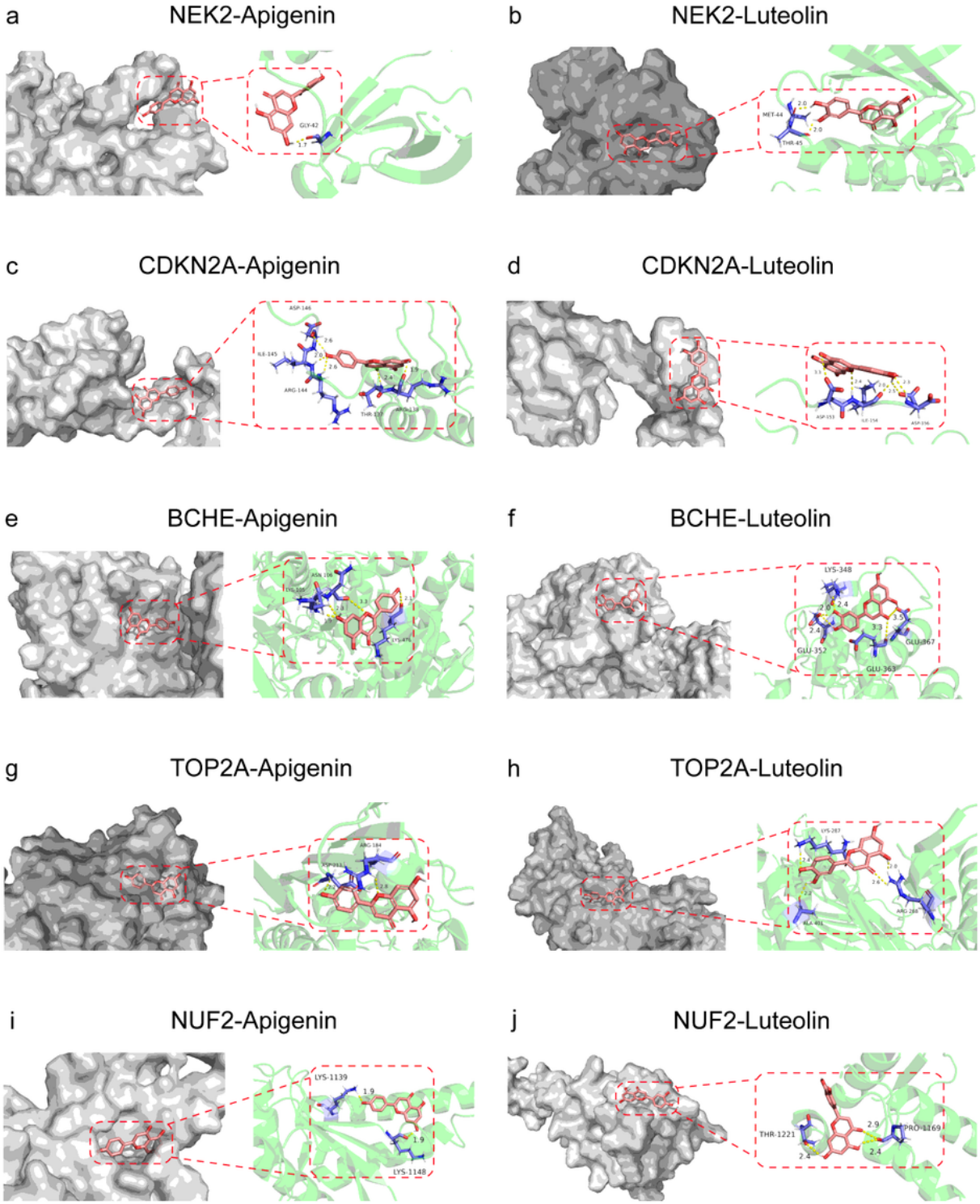

Figure 6

Molecular docking maps of the Hub targets with active compounds, and the yellow dotted line is the hydrogen bond. **a** NEK2-Apigenin, affinity = -4.22kcal·mol^{-1}; **b** NEK2-Luteolin, affinity = -3.64kcal·mol^{-1}; **c** CDKN2A-Apigenin, affinity = -4.04kcal·mol^{-1}; **d** CDKN2A-Luteolin, affinity = -3.68kcal·mol^{-1}; **e** BCHE-Apigenin, affinity = -3.90kcal·mol^{-1}; **f** BCHE-Luteolin, affinity = -2.54kcal·mol^{-1}; **g** TOP2A-Apigenin, affinity =

-3.11kcal·mol^{-1}; **h** TOP2A-Luteolin, affinity = -2.53kcal·mol^{-1}; **i** NUF2-Apigenin, affinity = -2.79kcal·mol^{-1}; **j** NUF2-Luteolin, affinity = -2.08kcal·mol^{-1}

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable.docx