



OPEN

Mechanism of *Astragalus membranaceus* in the treatment of laryngeal cancer based on gene co-expression network and molecular docking

Kai Feng Dong^{1,5}, Meng Qi Huo^{4,5}, Heng Ya Sun², Tian Ke Li³ & Dan Li¹✉

Astragalus membranaceus (HUANG QI, HQ) is a kind of traditional Chinese medicine. Researchers have widely concerned its antitumor effect. At present, there is still a lack of research on the treatment of laryngeal cancer with HQ. In this study, we integrated data from the weighted gene co-expression network of laryngeal cancer samples and the components and targets of HQ. A new method for dividing PPI network modules is proposed. Important targets of HQ treatment for laryngeal cancer were obtained through the screening of critical modules. These nodes performed differential expression analysis and survival analysis through external data sets. GSEA enrichment analysis reveals pathways for important targets participation. Finally, molecular docking screened active ingredients in HQ that could interact with important targets. Combined with the laryngeal cancer gene co-expression network and HQ PPI network, we obtained the critical module related to laryngeal cancer. Among them, MMP1, MMP3, and MMP10 were chosen as important targets. External data sets demonstrate that their expression in tumor samples is significantly higher than in normal samples. The survival time of patients with high expression group was significantly shortened, which is a negative factor for prognosis. GSEA enrichment analysis found that they are mainly involved in tumor-related pathways such as ECM receptor interaction and Small cell lung cancer. The docking results show that the components that can well bind to important targets of HQ are quercetin, rutin, and Chlorogenic acid, which may be the primary mechanism of the anti-cancer effect of HQ. These findings provide a preliminary research basis for Chinese medicine treatment of laryngeal cancer and offer ideas to related drug design.

Laryngeal cancer is one of the common malignant tumors of the head and neck, most of which are squamous cell carcinoma¹. It ranks second in the incidence of head and neck cancer (HNSC). The affected population is mainly middle-aged and older men over 40 years old, with a slight trend of youth². The occurrence and development of laryngeal cancer are related to many factors, including smoking, drinking, HPV infection, radiation, low intake of vegetables and fruits, insufficient trace elements, and disorders of sex hormone metabolism³⁻⁵. In the early stage of laryngeal cancer, laser treatment has less damage and less postoperative complications⁶. For most patients with laryngeal cancer, open surgery is still the main treatment. Radiotherapy and chemotherapy were used as auxiliary means⁷. Targeted treatment of tumor can increase drug sensitivity, improve curative effect and reduce side effects⁸. However, not all patients can benefit from it⁹. Clearing tumors, preserving function, controlling recurrence, and improving quality of life are the goals of treating laryngeal cancer⁵.

¹Department of Otolaryngology, The First Hospital of Hebei Medical University, Shijiazhuang 050000, China. ²Department of Otolaryngology, The Third Hospital of Shijiazhuang, Shijiazhuang 050011, China. ³Department of Stomatology, The Fourth Hospital of Hebei Medical University, Shijiazhuang 050011, China. ⁴School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing 102488, China. ⁵These authors contribute equally: Kai Feng Dong and Meng Qi Huo. ✉email: 16264688@qq.com

Astragalus membranaceus (HUANG QI, HQ) is one of the most commonly used Chinese medicine in clinical practice, which is the dried rhizome of *Astragalus membranaceus*. It has been reported that HQ can enhance immunity and protect cardiovascular system¹⁰. In recent years, the antitumor effect of HQ has attracted widespread attention from researchers¹¹. On the one hand, it can enhance cellular and humoral immune functions, improve the killing and inhibiting effects on tumor cells, and promote tumor cell apoptosis¹². On the other hand, it can also reduce the damage to the immune function of radiotherapy and chemotherapy, and play a role in the adjuvant treatment of tumors¹³. However, there are few related studies on HQ in treating laryngeal cancer.

In recent years, researchers have studied the pathogenesis of laryngeal cancer at the molecular level to find more effective treatments¹⁴. Network pharmacology is a new discipline based on systems biology theory, selecting specific signal nodes for multi-target drug design¹⁵. It emphasizes multi-channel regulation of signal pathways, improves the therapeutic effect of drugs, and reduces toxic and side effects¹⁶. This method can improve the success rate of clinical trials of new drugs and save drug development costs¹⁷. The protein–protein interaction (PPI) network is a common technique to analyze the mechanism of traditional Chinese medicine and disease from the molecular level¹⁸. The functional modules in the protein–interaction network are composed of interacting proteins, which together participate in specific biological processes¹⁹. Therefore, analyzing the functional modules in the network is of great significance for studying the mechanism of HQ treatment of laryngeal cancer. In previous studies, people often use the MCODE algorithm, FAG-EC algorithm, etc. to divide modules in PPI networks. Most of these algorithms divide modules based on the heterogeneity of scale-free networks^{20,21}. They can obtain modules with tighter connectivity on the network topology. However, we prefer to get modules that are more closely connected to biological functions²². In this study, we designed a method for obtaining modules that integrate disease gene expression data and clinical information. Based on the gene expression data of 109 laryngeal cancer samples, we constructed a weighted gene co-expression network of laryngeal cancer. The modules are divided according to the correlation between genes, and then they are mapped in the PPI network of related targets of HQ. In this way, modules with a high correlation in biological functions can be received, and redundant PPI data in the protein interaction network can be removed. This method makes the selection process of critical nodes more concise and accurate²³.

In this study, we propose a new method for obtaining PPI network modules. Based on the gene expression information in the GSE27020 dataset, a weighted gene co-expression network was constructed. The co-expressed gene modules were mapped into the PPI network of HQ, and critical genes and corresponding targets were obtained by screening. These genes were confirmed to be differentially expressed in the tumor group and the normal group, and the survival time of patients with high expression group was significantly shortened. Based on the docking, the active ingredients in HQ were obtained. These findings may provide a new perspective for the treatment of laryngeal cancer with HQ.

Results

Acquisition of laryngeal cancer samples. The GSE27020 data set contains a total of 109 laryngeal cancer samples. After removing the genes with too many missing values from samples, an expression matrix containing 13,512 genes was generated. The 3,378 genes with the top 25% of the variance (that is, the genes that change significantly in each sample) are selected for subsequent operations. Clinical phenotypic data included survival time, recurrence, age, and tumor stage. As shown in Fig. S1A. According to the distance of sample clustering, the outliers GSM665579, GSM665597, and GSM665566 were eliminated. Finally, we selected 106 samples to construct a weighted gene co-expression network.

Identification of critical modules. We use the WGCNA package in R software to build a weighted co-expression network²⁴. When $R^2 > 0.9$, the network meets the scale-free condition. The calculation result is shown in Fig. 1A, where $SFT.R.sq$ is R^2 . When the power value is set to 5, the connectivity between genes in the network satisfies the scale-free network distribution. We use a dynamic cutting algorithm to obtain ten modules, of which the gray module is a set of genes that cannot be aggregated into other modules (Fig. 1B). Then, the correlation between each module and the phenotype (survival time, recurrence, age, and grade) was calculated (Fig. 1C). A heat map showed a significant positive correlation between the green module and recurrence ($cor = 0.32$, $P < 0.001$) (Fig. 1D). The number of genes in each module is shown in Fig. 1E. The GS and MM scatter plots of the green module also illustrate the significant correlation between the green module and laryngeal cancer. Therefore, the green module was selected as the critical module.

Enrichment analysis of the critical module. We used GO enrichment analysis and KEGG pathway analysis to find the biological function of genes in the green module (Fig. 2)^{25,26}. In the biological process classification, genes in the module are significantly enriched in extracellular matrix organization, hemidesmosome assembly, cell adhesion, cellular response to zinc ion, negative regulation of growth, extracellular matrix disassembly, cell–matrix adhesion, cell adhesion mediated by integrin, cellular response to cadmium ion, leukocyte migration. In the cellular component classification, the genes in modules are significantly enriched in focal adhesion, cell surface, extracellular space, basement membrane, integrin complex, perinuclear region of cytoplasm, extracellular matrix, extracellular region, proteinaceous extracellular matrix, extracellular exosome. In the molecular function classification, the genes in modules are significantly enriched in integrin binding, L-ascorbic acid binding, protein binding, procollagen-lysine 5-dioxygenase activity, laminin binding, cadherin binding involved in cell–cell adhesion, zinc ion binding, protease binding, ion channel binding, glycoprotein binding. In the KEGG pathway classification, the genes in modules are significantly enriched in focal adhesion, ECM-receptor interaction, PI3K-Akt signaling pathway, regulation of actin cytoskeleton, mineral absorp-

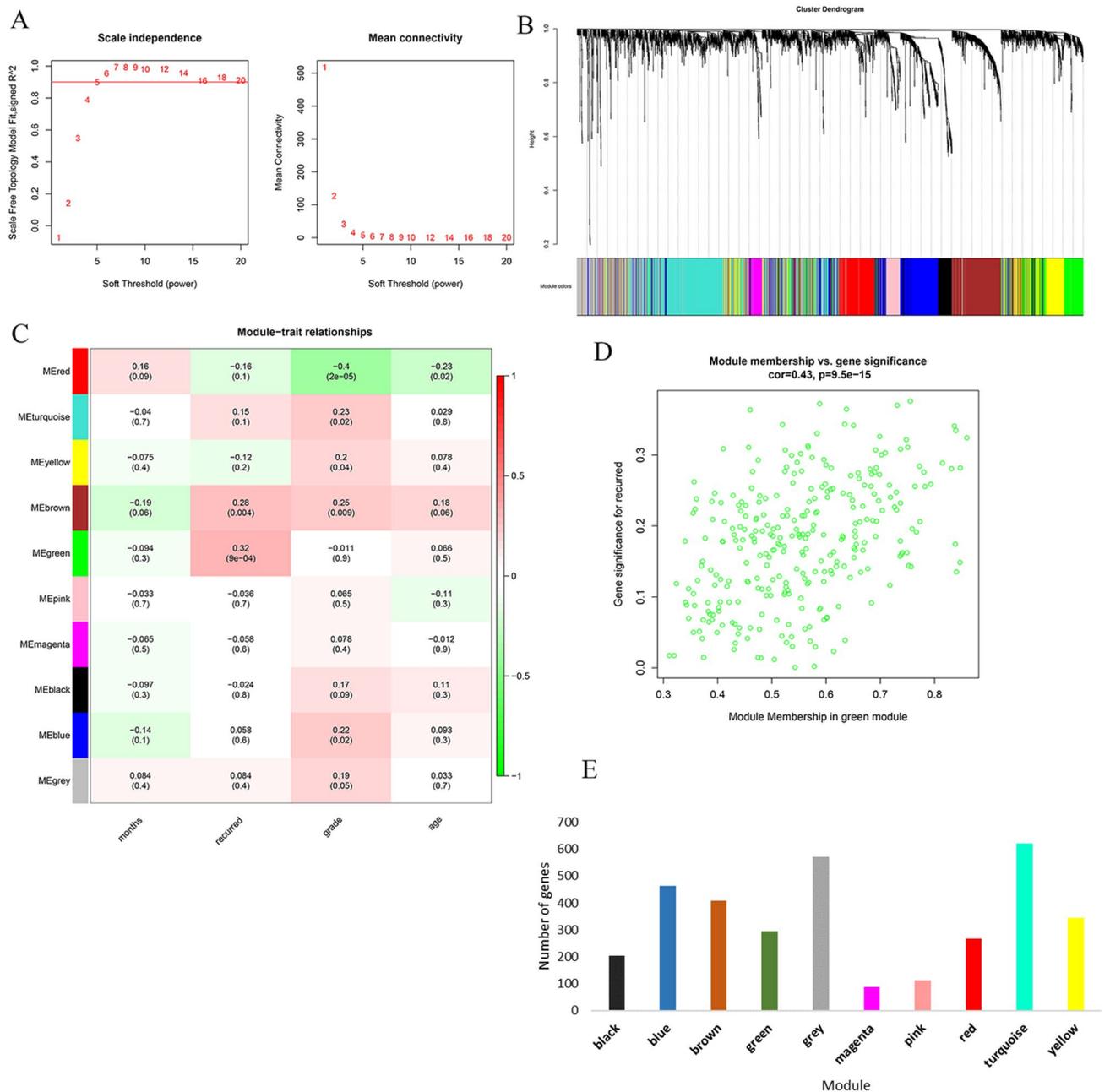


Figure 1. Drawing weighted gene co expression network of laryngeal cancer based on WGCNA package of R 3.5.1 software (<https://www.r-project.org/>) (A) Analysis of the scale-free fit index for various soft-thresholding powers (β), 5 was the most fit power value. (B) The cluster dendrogram of genes in GSE27020. Each branch in the figure represents one gene, and every color below represents one co-expression module. (C) Heat map of the relationship between the characteristic genes of modules and the incidence of laryngeal cancer based on R 3.5.1 software. Correlation coefficient along with P value in parenthesis underneath. Color-coded according to correlation coefficient (legend at right). (D) Scatter plot of gene significance and module membership in the green module. (E) The number of genes in each module.

tion, small cell lung cancer, arrhythmogenic right ventricular cardiomyopathy (ARVC), proteoglycans in cancer, hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy. The detailed results are shown in Table S1.

Identification of important targets. A total of 77 components of HQ and 953 corresponding targets were obtained from the TCMSP database. The target's PPIs were introduced into Cytoscape to obtain the PPI network of HQ. The network contains 471 nodes and 2,983 edges, as shown in Fig. S2. We take HQ-PPI network as the background. Gene interactions in the weighted gene co expression network of laryngeal cancer was mapped to the HQ-PPI network. Common nodes and interactions between gene modules and PPI network are preserved. Finally, we got laryngeal cancer-HQ-PPI network. The network consists of 64 nodes and 249 edges, which are composed of 9 modules (with the gray modules removed), as shown in Fig. 3A. Therefore, the cor-

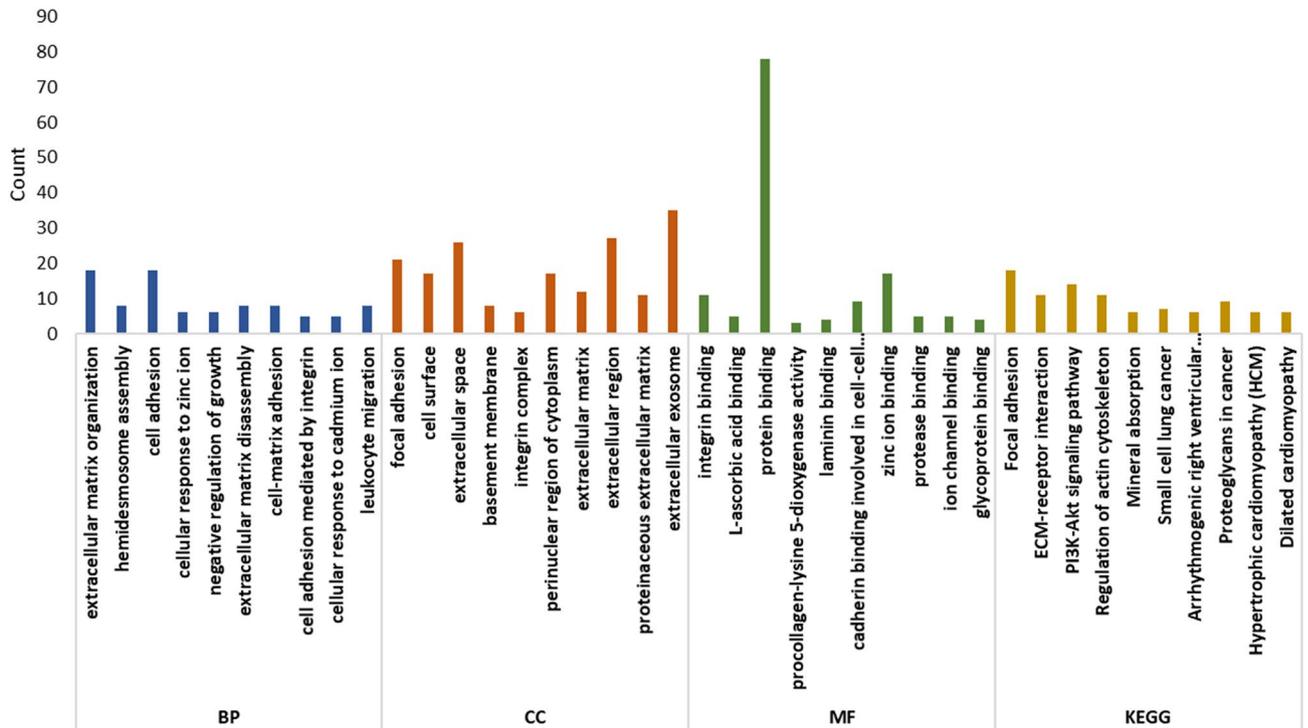


Figure 2. GO and KEGG enrichment analysis of green module based on DAVID 6.8 database (<https://david.ncicrf.gov/>). X-axis refers to the ID of the enriched terms, and Y-axis refers to the number of the enriched genes. Brown bars show the results of the biological process classification. Blue bars show the results of the cellular component classification. Yellow bars show the results of the molecular function classification. Green bars show the results of the KEGG pathway classification.

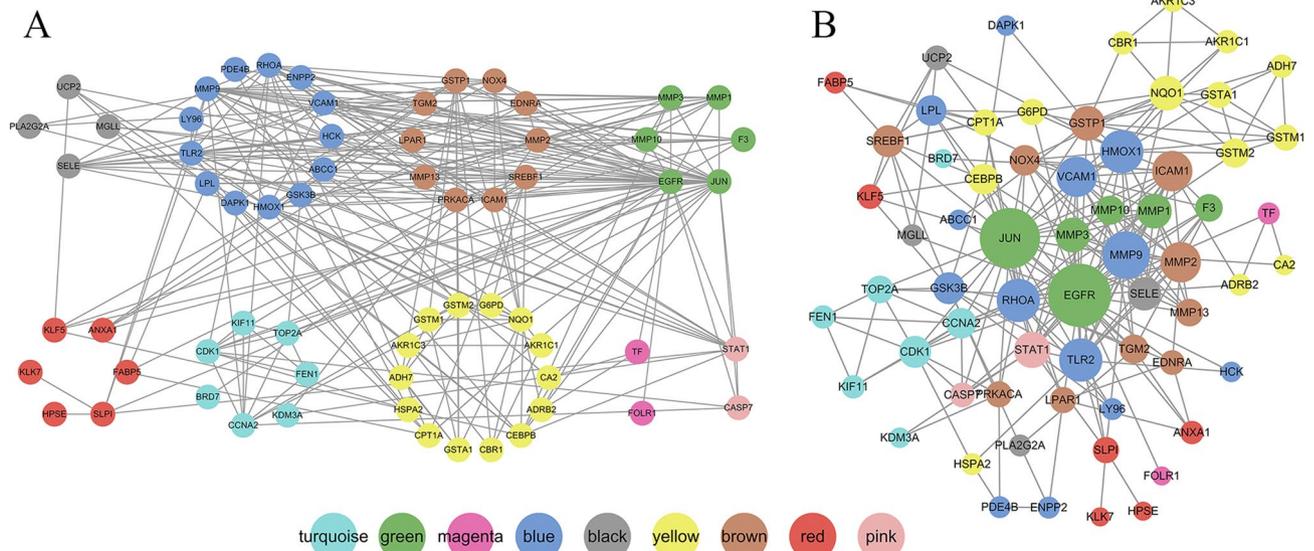


Figure 3. Constructing network based on Cytoscape 3.7.0. (<https://cytoscape.org/>) (A) Module diagram of laryngeal cancer weighted gene co-expression network mapped to HQ PPI network. The different colors in the figure represent the corresponding modules. The black line represents the PPI relationship between the nodes. (B) A larger node indicates a greater degree of the node.

responding green module after the mapping is considered to be a critical module for HQ treatment of laryngeal cancer. Nodes in the green module are considered candidate important targets. By calculating the topology parameters of the network nodes, the degree value of each node is obtained. As shown in Fig. 3B, compared with the nodes of other modules, the nodes of the green module have a higher degree value and are in the core position of the network.

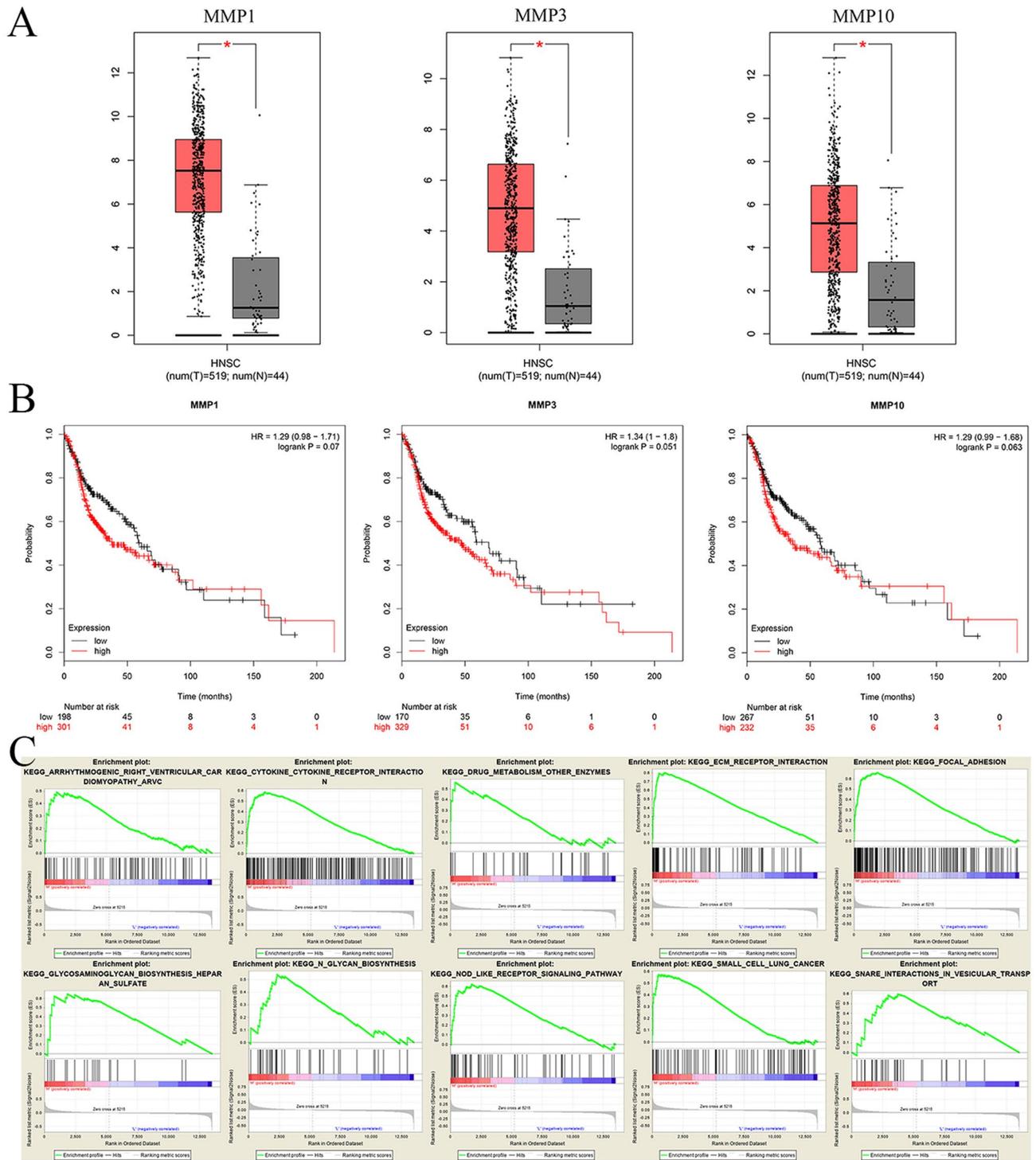


Figure 4. (A) Differential expression analysis of important targets based on TCGA database and GTEx database. (B) Survival analysis of important targets. Notes: three important targets in green module with significant results of survival analysis ($P < 0.1$ was regarded as significant). They were MMP1, MMP3, MMP10 respectively. (C) Enrichment analysis of important targets in high expression group based on GSEA 3.0 software (<https://software.broadinstitute.org/gsea/downloads.jsp>).

Differential expression analysis, survival analysis, and GSEA analysis of important targets. Based on the sample information of TCGA and GTEx projects in the GEPIA database, it was found that some candidate important targets had differential expression between normal and laryngeal cancer. The Wilcoxon rank test was performed on the normal group and the laryngeal cancer group (Fig. 4A). Among them, the expressions of MMP1, MMP3, and MMP10 in tumor tissues were higher than those in normal tissues ($p < 0.05$). After that, we used the Kaplan Meier-plotter website for survival analysis. The results are shown

MMP1	Affinity (kcal/mol)	MMP3	Affinity (kcal/mol)	MMP10	Affinity (kcal/mol)
MOL000415	-9.6	MOL000098	-9.6	MOL000437	-8.9
MOL001955	-9	MOL000438	-9.6	MOL000415	-8.8
MOL000098	-8.8	MOL000436	-9.4	MOL000418	-8.8
MOL000417	-8.8	MOL000409	-9.4	MOL000098	-8.7
MOL000442	-8.8	MOL000415	-9.3	MOL000442	-8.4
MOL000433	-8.8	MOL000417	-9.3	MOL001955	-8.3
MOL000373	-8.8	MOL000418	-9	MOL000422	-8.3
MOL000436	-8.7	MOL001955	-9	MOL000251	-8.3
MOL000354	-8.7	MOL000422	-9	MOL000423	-8.3
MOL000416	-8.7	MOL000251	-9	MOL000239	-8.3
MOL000033	-8.7	MOL000376	-9	MOL000439	-8.3

Table 1. Top ten compounds with free binding energy for each target.

in Fig. 4B, in which MMP1, MMP3, and MMP10 were significantly correlated with patient prognosis ($P < 0.1$). As their expression levels increase, the overall survival time decreases significantly. The above results suggest that MMP1, MMP3, and MMP10 are poor prognostic factors, and their high expression is significantly associated with shortened patient survival. Therefore, they are considered to be important targets closely related to laryngeal cancer, which may be the primary mechanism of HQ treatment of laryngeal cancer. To understand the pathways involved in these important targets (MMP1, MMP3, MMP10), a GSEA analysis was performed using the data set GSE27020. There were ten pathways enriched in the high expression group (Fig. 4C), which are the snare interactions in vesicular transport, focal adhesion, ECM receptor interaction, glycosaminoglycan biosynthesis heparan sulfate, cytokine cytokine receptor interaction, N glycan biosynthesis, small cell lung cancer, arrhythmogenic right ventricular cardiomyopathy (ARVC), Nod-like receptor signaling pathway, drug metabolism other enzymes.

Molecular docking. A total of 77 HQ components were obtained through the TCMSP database, and the component information is shown in Table S2. The crystal structure with the PDB code of 3AYK was selected as the docking model of MMP1, and the central axis was adjusted to 3.816, -4.348, -0.55 for MMP1. The crystal structure with the PDB code of 4G9L was selected as the docking model of MMP3, and the central axis was adjusted to 21.36, 68.883, 106.078 for MMP3. The crystal structure with the PDB code of 1Q3A was selected as the docking model of MMP10, and the central axis was adjusted to 25.304, -10.345, 24.278 for MMP10. All components are sequentially docked with MMP1, MMP3, and MMP10 to obtain the corresponding binding free energy. The calculation results of compounds and targets in the top ten (including all the 10th in parallel) are shown in Table 1 and Fig. 5. The docking results of other compounds are shown in Table S3.

As shown in the figure, among the top 10 ingredients, the best combination with MMP10 is isomucronulatol-7,2'-di-O-glucoside, Jaranol, rhamnocitrin-3-O-glucoside, Hirsutrin, the best combination with MMP3 is (3R)-3-(2-hydroxy-3,4-dimethoxyphenyl)chroman-7-ol,7,2'-dihydroxy-3',4'-dimethoxyisoflavone-7-O- β -D-glucoside, Astragaloside IV, and the best combination with MMP1 is (2S)-2-[[4-[(2-Amino-4-oxo-1H-pteridin-6-yl)methylamino]benzoyl]amino]pentanedioic acid, (3S,8S,9S,10R,13R,14S,17R)-10,13-dimethyl-17-[(2R,5S)-5-propan-2-yl]octan-2-yl]-2,3,4,7,8,9,11,12,14,15,16,17-dodecahydro-1H-cyclopenta[a]phenanthren-3-ol, Lariciresinol, Isorhamnetin, (2S)-4-methoxy-7-methyl-2-[1-methyl-1-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-methylol-tetrahydropyran-2-yl]oxy-ethyl]-2,3-dihydrofuro[3,2-g]chromen-5-one. Some ingredients have stronger binding effects with three targets, such as quercetin, rutin, Chlorogenic acid.

Discussion

In this study, we obtained ten modules based on 109 laryngeal cancer expression data in the GSE27020 dataset. The green module is the critical module we found. The results of enrichment of GO and KEGG also showed the mechanism of laryngeal cancer. Among them, cell adhesion mediated by integrin, leukocyte migration, ECM-receptor interaction, PI3K-Akt signaling pathway etc. have been proved to be closely related to the occurrence of cancer. Integrin mediates the interaction between cell and cell, cell and extracellular matrix, and participates in the information transmission inside and outside cell²⁷. It plays an important role in the proliferation, differentiation, adhesion, migration and apoptosis of cancer cells. Some studies have shown that integrin is the key to cancer cell proliferation. When cancer cells float, integrin will change its function from adhesion to communication²⁸. The dysfunction of ECM-receptor interaction may lead to the abnormal activation of signal transduction pathways, including Ras/ Raf/MAPK, Raf/JNK, Rho/RAC/Pak and PI3K/Akt/mTOR. So as to create a tumor microenvironment that can promote tumor survival, angiogenesis and invasion²⁹. PI3K-Akt signaling pathway is the most common dysregulated pathway in many tumor cells. When it is activated abnormally, many tumor suppressor genes, including PIK3CA, PIK3R1, PTEN, Akt, TSC1, TSC2, LKB1, mTOR, will be affected³⁰. The results of the enrichment analysis may be the main mechanism of laryngeal cancer.

The green module was mapped on the HQ PPI network, and finally, the potential critical targets MMP1, MMP3, and MMP10 for the treatment of laryngeal cancer were obtained. Gene expression data from the external

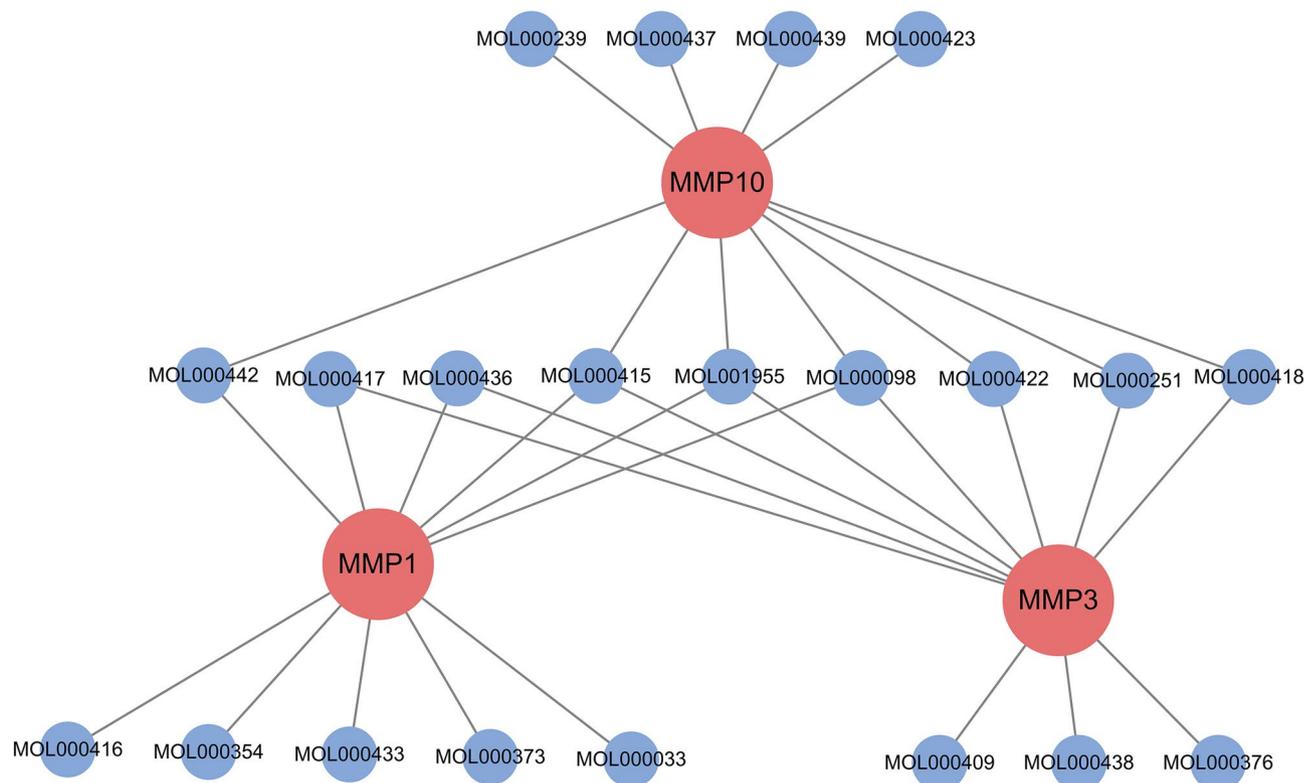


Figure 5. Constructing compound-target interaction network diagram based on Cytoscape 3.7.0. (<https://cytoscape.org/>). The blue nodes are top ten compounds with free binding energy for each target. The red nodes are important targets. The black lines indicate the relationship between compounds and targets.

data set showed that the expression of these three genes in the tumor group was significantly higher than in the normal group. Survival analysis showed that high expressions of MMP1, MMP3, and MMP10 would lead to a reduction in patient survival. They are detrimental to prognosis. GSEA analysis showed that the MMP1, MMP3, and MMP10 high-expression groups participated in tumor-related ECM receptor interaction, small cell lung cancer, and other pathways. MMP1, MMP3, and MMP10 belong to the matrix metalloproteinases (MMPs) family³¹. The MMPs family contains multiple types of proteases. They can break down most of the extracellular matrix components and some proteins, producing growth factors essential for tumorigenesis and progression. Therefore, the MMP family has become a biomarker for some tumors³². MMP1 expression increases in malignant tumor cells, which is helpful for tumor cell invasion and metastasis³³. Studies have shown that the genetic characteristics of MMP1 may become an independent prognostic factor for laryngeal cancer. And the expression of MMP1 in most of the laryngeal cancer tissues with lymphatic metastasis was higher than that of the laryngeal cancer tissues without lymphatic metastasis. MMP1 also increased with the increase of infiltration³⁴. MMP3 is one of the most active members in MMPs, and it is expressed in tumor cells such as gastrointestinal tumors and liver tumors³⁵. Gobin Emily's research found that MMP3 and MMP10 were significantly up-regulated in at least ten cancer types³⁶. Alaseem Ali's research shows that the MMPs family plays a vital role in the invasion and metastasis of malignant tumors and the formation of new blood vessels. Inhibiting the expression and activity of MMPs can provide new ideas for the comprehensive treatment of tumors³⁷. In addition, The secretion of MMP1 can stimulate the secretion of other MMPs (e.g., MMP3, MMP10), which may be the reason why they together become a critical gene of laryngeal cancer³⁸.

The docking results show that quercetin, rutin, and Chlorogenic acid can be combined with MMP1, MMP3, and MMP10. Previous studies have shown that quercetin induces apoptosis in human Hep2 cells by inhibiting Akt / PKB phosphorylation. The effect of 40 μ M quercetin on the induction of apoptosis in Hep2 cells increased from 18.7 to 42.2%³⁹. Rutin can inhibit the activity of NF-kappa B and P38 and regulate the occurrence and development of lung cancer⁴⁰. It can also induce mitochondrial apoptosis in colon cancer cells through a caspase-dependent mechanism⁴¹. In breast cancer, rutin inhibits P-GP and BCRP pumps non-selectively. It can effectively reverse the resistance of breast cancer cells and restore sensitivity to cyclophosphamide⁴². Chlorogenic acid inhibits tumorigenesis of breast cancer cells by causing mitochondrial dysfunction. It can increase the sensitivity of tumour cells to chemotherapy⁴³. Besides, chlorogenic acid can also affect the expression of apoptosis-related genes in A549 human lung cancer cells⁴⁴. In addition to quercetin has been reported to be associated with laryngeal cancer, rutin and chlorogenic acid have been reported to be associated with a variety of cancers. But whether they can treat laryngeal cancer has not been reported, and it may become a potential therapeutic drug.

Our method is different from the common PPI-based network module division method. In the past, network modules were divided based on scale-free topological characteristics to obtain the modules with the best connectivity. In this study, we integrated gene expression data for laryngeal cancer. Modules are divided by co-expressed

genes in laryngeal cancer so that they have a better degree of connection in biological functions. After this, we performed a topology analysis on the mapped network and found that important targets are in the center of the network. Therefore, our method not only integrates the patient's disease information but also ensures the high connectivity of the important targets. As the number of patient samples in TCGA, GEO and other public databases increases, more potential targets and active ingredients can be discovered in the future. Besides, our research lacks the validation of animal or cell experiments. In future research, we will perform pharmacodynamic, and toxicity experiments on these candidate components, and in vivo and in vitro experiments such as RT-PCR and western blot will perform biological verification of potential critical genes.

In this study, we analyzed the gene expression data of 109 laryngeal cancer samples and screened to obtain three important targets (MMP1, MMP3, and MMP10). Their expression in tumor samples was significantly higher than in normal samples, and patients in the high expression group had significantly shorter survival times. The components that can be combined with critical targets in HQ are quercetin, rutin, and Chlorogenic acid. These components may be the primary mechanism of the anti-cancer effect of HQ. These findings provide a preliminary research basis for Chinese medicine treatment of laryngeal cancer and offer ideas to related drug design.

Methods

Data download and preprocessing. The laryngeal cancer GSE27020 data set was obtained from the GEO database and contained a total of 109 laryngeal cancer samples. The downloaded data needs to be pre-processed before analysis. The clinical information of the samples was extracted based on the Affymetrix Human Genome U133A Array platform, and the probe names were converted into gene names. Finally, a gene expression matrix with row names of gene names and column names of sample names was obtained for subsequent analysis.

Construction of weighted gene co-expression network. In this study, we use WGCNA package in R software to construct the weighted gene coexpression network of laryngeal cancer. Before constructing network, we need to remove outliers. In this study, the goodsamplesgenes algorithm was used to remove the genes with too many missing values. Then, the hclust algorithm was used to cluster the tissue samples. The top 25% of genes with variance in expression were selected. The remaining samples were used to build a weighted gene co-expression network. The weighting coefficient β is selected according to the criteria of the scale-free network, and the correlation matrix is converted into an adjacency matrix, which is hierarchical clustering based on the value of topology matrix (TOM) value. The dis-TOM is used as the distance measurement of the modules. The minimum module size is set as 10. We use a dynamic hybrid cutting algorithm to identify the module and draw the gene tree⁴⁵.

Then, we calculate the eigenvalue of the module, which is the first principal component obtained by principal component analysis (PCA) of all genes in the module. Screen the modules that are significantly related to clinical characteristics as critical modules. Calculate gene significance (GS) and module membership (MM) to measure the correlation between genes and clinical information⁴⁶.

Enrichment analysis of critical modules. In order to understand the biological functions of candidate genes, this study maps the genes of critical modules to the online website DAVID (<https://david-d.ncifcrf.gov/>) for gene ontology (GO) enrichment analysis and KEGG pathway analysis. The items with $P < 0.05$ were statistically significant.

PPI network of potential targets of HQ in the treatment of laryngeal cancer. Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform (TCMSP, <https://tcmspw.com/tcmsp.php>) includes 499 Chinese herbal medicines registered in the Chinese Pharmacopoeia, 29,384 ingredients, 3,311 targets, and 837 related diseases⁴⁷. From this, all components of HQ and corresponding targets can be obtained. String (version 11.0, <https://string-db.org>) is a database integrating known and predicted protein-protein interactions, and has the function of visualizing protein-protein interaction network⁴⁸. The PPIs of the HQ target were obtained with medium confidence of 0.4 and then imported into Cytoscape to construct the PPI network of HQ target⁴⁹.

If a node is a potential target of HQ, and it is also in a critical module of a weighted gene co-expression network, this node may be a potential target for HQ treatment of laryngeal cancer⁵⁰. Therefore, we take the HQ PPI network as the background and map the gene modules in the weighted gene co-expression network to the HQ PPI network. The Uniprot database (<https://www.uniprot.org/>) is used to convert gene names and target names⁴¹. Therefore, the module based on the weighted gene co-expression network automatically divides the PPI network into multiple modules. The nodes in the critical module are candidate important targets.

Verification of important targets based on external data sets. GEPIA (<https://gepia.cancer-pku.cn/>) is an interactive web server for integrated analysis of cancer expression profiling data⁵². It contains 33 malignant tumor RNA sequencing expression data from TCGA and GTEx. In this study, gene expression data from the TCGA and GTEx databases were used as external test sets to verify important targets obtained in the previous screening. Run the "BoxPlots" module to analyze whether important targets have been differentially expressed between laryngeal cancer and normal samples. Kaplan Meier-plotter (<https://kmplot.com/analysis/index.php?p=service>) is a website for online survival analysis⁵³. The survival analysis of important targets was carried out. Under different expression levels, whether the survival time of patients has a significant difference. Enrichment analysis of pathways associated with high or low expression of important targets was performed using GSEA 3.0 software⁵⁴. They are grouped by the expression level of important targets in the GSE27020 data-

set. The data set of c2.cp.kegg.v7.0.symbols.gmt in the msigdb database of the GSEA website was selected as the reference gene set. The number of random combinations was set to 1,000.

Verification of important targets based on molecular docking. Download 3D structures of all molecules in HQ based on the PubChem database⁵⁵. The 3D structure of the target was searched and downloaded from the PDB protein database (<https://www.rcsb.org/pdb/home/home.do>)⁵⁶. We prefer to select high-resolution crystal structures complexed with corresponding biologically active ligands and use Autodock software for docking⁵⁷. All ligands are removed from the protein receptor complex, plus polar hydrogen atoms and charges. Finally, the compounds in HQ dock with protein crystals in a semi-flexible manner.

Received: 7 January 2020; Accepted: 4 June 2020

Published online: 07 July 2020

References

- Philippe, G. A contemporary review of evidence for transoral robotic surgery in laryngeal cancer. *Front. Oncol.* **8**, 121 (2018).
- Du, L., Li, H., Zhu, C., Zheng, R. & Chen, W. J. Incidence and mortality of laryngeal cancer in China. *Chin. J. Cancer Res.* **27**, 52–58 (2015).
- Elwany, S., Radi, S., Khalil, H., Talaat, I. & Belasy, K. Cluster of differentiation 8 T-cell population in the laryngeal mucosa of smokers with laryngeal cancer. *J. Laryngol. Otol.* **132**, 1134–1137 (2018).
- Starska, K. *et al.* Fibroblast growth factor receptor 1 and 3 expression is associated with regulatory PI3K/AKT kinase activity, as well as invasion and prognosis, in human laryngeal cancer. *Cell. Oncol.* **41**, 253–268 (2018).
- Baird, B. J., Sung, C. K., Beadle, B. M. & Divi, V. J. O. O. Treatment of early-stage laryngeal cancer: a comparison of treatment options. *Oral Oncol.* **87**, 8–16 (2018).
- Horwich, P. *et al.* Laryngeal recurrence sites in patients previously treated with transoral laser microsurgery for squamous cell carcinoma. *J. Otolaryngol. Head Neck Surg.* **47**, 14 (2018).
- Dotto, J. E. Solitary fibrous tumor of the larynx: a case report and review of the literature. *Arch. Pathol. Lab. Med.* **130**, 213 (2006).
- Tao, Y. *et al.* Identification of novel enriched recurrent chimeric COL7A1-UCN2 in human laryngeal cancer samples using deep sequencing. *BMC Cancer* **18**, 1 (2018).
- Eskander, A., Blakaj, D. M. & Dziegielewski, P. T. J. O. O. Decision making in advanced larynx cancer: an evidenced based review. *Oral Oncol.* **86**, 195–199 (2018).
- Zhou, R. *et al.* Extract from Astragalus membranaceus inhibit breast cancer cells proliferation via PI3K/AKT/mTOR signaling pathway. *BMC Complement. Altern. Med.* **18**, 83 (2018).
- Auyeung, K. K., Han, Q. B. & Ko, J. K. Astragalus membranaceus : A review of its protection against inflammation and gastrointestinal cancers. *Am. J. Chin. Med.* **44**, 1–22 (2016).
- Liu, A. J. *et al.* Extraction of a novel cold-water-soluble polysaccharide from Astragalus membranaceus and its antitumor and immunological activities. *Molecules* **23**, 62 (2017).
- Yu, J., Ji, H. Y. & Liu, A. J. Alcohol-soluble polysaccharide from Astragalus membranaceus: preparation, characteristics and anti-tumor activity. *Int. J. Biol. Macromol.* **118**, 2057–2064 (2018).
- Liu, H. H. *et al.* Characterization of long non-coding RNA and messenger RNA profiles in laryngeal cancer by weighted gene co-expression network analysis. *Aging-Us* **11**, 10074–10099 (2019).
- Guo, Y. *et al.* Multiscale modeling of inflammation-induced tumorigenesis reveals competing oncogenic and onco-protective roles for inflammation. *Can. Res.* **2017**, 1662 (2017).
- Yucheng, G. *et al.* Network-Based Combinatorial CRISPR-Cas9 Screens Identify Synergistic Modules in Human Cells. *ACS Synth. Biol.* (2019).
- Zheng, J. *et al.* Network pharmacology to unveil the biological basis of health-strengthening herbal medicine in cancer treatment. *Cancers* **10**, 1 (2018).
- Peyvandi, H., Peyvandi, A. A., Safaei, A., Azodi, M. Z. & Rezaei-Tavirani, M. Introducing potential key proteins and pathways in human laryngeal cancer: a system biology approach. *Iran. J. Pharmaceut. Res.* **17**, 415–425 (2018).
- Sikandar, A. *et al.* Decision tree based approaches for detecting protein complex in protein protein interaction network (PPI) via link and sequence analysis. *IEEE Access* **6**, 22108–22120 (2018).
- Gan, Y., Zheng, S., Baak, J. P. A., Zhao, S. & Fu, C. Prediction of the anti-inflammatory mechanisms of curcumin by module-based protein interaction network analysis. *Acta Pharmaceutica Sinica B* **5**, 590–595 (2015).
- Wang, J. *et al.* ClusterViz: A Cytoscape APP for Cluster Analysis of Biological Network. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **12**, 815–822 (2015).
- Zhi, W. & Biology, Z. Correction: In Search of the Biological Significance of Modular Structures in Protein Networks. *PLoS Computational Biology* **3**, e146(2007).
- Liu, G., Wong, L. & Chua, H. N. J. B. Complex discovery from weighted PPI networks. *Bioinformatics* **25**, 1891–1897 (2009).
- Peter, L. & Steve, H. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* **9**, 559 (2008).
- Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2016).
- Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
- Yu, H., Gao, M. & Ma, Y. Inhibition of cell migration by focal adhesion kinase Time-dependent difference in integrin-induced signaling between endothelial and hepatoblastoma cells. *Int. J. Mol. Med.* **41**, 1 (2018).
- Huan, L. *et al.* PRDM4 mediates YAP-induced cell invasion by activating leukocyte-specific integrin b2 expression. *EMBO Rep.* **19**, e45180 (2018).
- Fang, E. *et al.* Identification of breast cancer hub genes and analysis of prognostic values using integrated bioinformatics analysis. *Cancer Biomark.* **1**, 1–9 (2020).
- Luo, H. J., Manning, S. D., Cantley, L. C. Targeting the PI3K-Akt pathway in human cancer: rationale and promise **4**, 257–262 (2003).
- Kai, K., Plaks, V. & Werb, Z. Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell* **141**, 67 (2010).
- Overall, C. M. & Kleinfeld, O. Validating matrix metalloproteinases as drug targets and anti-targets for cancer therapy. *Nat. Rev. Cancer* **6**, 227–239 (2006).
- Deryugina, E. I. & Rev, J. Matrix metalloproteinases and tumor metastasis. *Cancer Metastasis Rev.* **25**, 9–34 (2006).
- Olejniczak, I. *et al.* Polymorphism of metalloproteinases MMP-1 and MMP-2 in risk of laryngeal cancer. *Otolaryngologia Polska the Polish Otolaryngology* **66** (2012).
- Zhang, C. *et al.* Meta-analysis of MMP2, MMP3, and MMP9 promoter polymorphisms and head and neck cancer risk. *PLoS ONE* **8**, e62023 (2020).

36. Gobin, E., Bagwell, K., Wagner, J., Mysona, D. & Cancer, J. A pan-cancer perspective of matrix metalloproteases (MMP) gene expression profile and their diagnostic/prognostic potential. *BMC Cancer* **19**, 1 (2019).
37. Alaseem, A. *et al.* Matrix Metalloproteinases: A challenging paradigm of cancer management. *Semin. Cancer Biol.* **56**, 100–115 (2017).
38. O'Hara, A., Lim, F. L., Mazzatti, D. J. & Trayhurn, P. Microarray analysis identifies matrix metalloproteinases (MMPs) as key genes whose expression is up-regulated in human adipocytes by macrophage-conditioned medium. *Eur. J. Physiol.* **458**, 1103–1114 (2009).
39. Sharma, H., Sen, S. & Singh, N. Molecular pathways in the chemosensitization of cisplatin by quercetin in human head and neck cancer. *Cancer Biol. Ther.* **4**, 949–955 (2005).
40. Li, X. *et al.* Expression of NF-kappaB and p38 under intervention of rutin in lung cancer therapy. *Biomed. Res. India* **28**, 2344–2347 (2017).
41. Guon, T. & Chung, H. Hyperoside and rutin of *Nelumbo nucifera* induce mitochondrial apoptosis through a caspase-dependent mechanism in HT-29 human colon cancer cells. *Oncol. Lett.* **11**, 2463–2470 (2016).
42. Iriti, M. *et al.* Rutin, a quercetin glycoside, restores chemosensitivity in human breast cancer cells. *J. Phytother. Res.* **31**, 1529–1538 (2017).
43. Liao, W. T. & Li, C. J. Chlorogenic acid inhibits tumorigenesis and enhances chemotherapeutic sensitivity of breast cancer cells through mitochondrial dysfunction. *Int. J. Mol. Med.* **40**, S51–S51 (2017).
44. Yamagata, K., Izawa, Y., Onodera, D., Tagami, M. J. M. & Biochemistry, C. Chlorogenic acid regulates apoptosis and stem cell marker-related gene expression in A549 human lung cancer cells. *Mol. Cell. Biochem.* **441**, 9–19 (2018).
45. Zhou, X. G. *et al.* Identifying miRNA and gene modules of colon cancer associated with pathological stage by weighted gene co-expression network analysis. *Oncotargets And Therapy* **11**, 2815–2830 (2018).
46. 46Giulietti, M. *et al.* Emerging Biomarkers in Bladder Cancer Identified by Network Analysis of Transcriptomic Data. *Frontiers in Oncology* **8**, (2018).
47. Ru, J. *et al.* TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. *J. Cheminf.* **6**, 13 (2014).
48. Damian, S. *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447–D452 (2015).
49. Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M. & Research, L. STRING V91: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **41**, D808–D815 (2012).
50. Huo, M., Wang, Z., Wu, D., Zhang, Y. & Qiao, Y. J. Using coexpression protein interaction network analysis to identify mechanisms of danshensu affecting patients with coronary heart disease. *Int. J. Mol. Sci.* **18**, 1 (2017).
51. Research, N. UniProt: a hub for protein information. *Nucleic Acids Res.* **43**, D204–D212 (2015).
52. Tang, Z. *et al.* GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* **45**, W98–W102 (2017).
53. Hou, G. X., Liu, P., Yang, J. & Wen, S. Mining expression and prognosis of topoisomerase isoforms in non-small-cell lung cancer by using OncoPrint and Kaplan-Meier plotter. *PLoS ONE* **12**, e0174515 (2017).
54. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA.* **102**, 15545–15550 (2005).
55. Wang, Y. *et al.* PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res.* **37**, W623–W633 (2009).
56. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
57. Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).

Acknowledgements

This study was supported by Project of medical science research of Hebei Health Committee (no. 20170486).

Author contributions

M.H., D.L. and K.D. conceived the project. M.H., K.D. and H.S. performed the experiments. K.D. and D.L. wrote the paper. T.L. and D.L. were involved in planning and supervising the project. All authors discussed the results, commented on the manuscript and approved the final format.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-68093-0>.

Correspondence and requests for materials should be addressed to D.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020