Network and pathway-based analysis of genes associated with esophageal squamous cell carcinoma

Wenwu He[1,2#^], Kun Yuan[3#], Jinlan He[1], Chenghao Wang[2], Lin Peng[2], Yongtao Han[2], Nianyong Chen[1]

[1]Department of Head and Neck Oncology and Department of Radiation Oncology, Cancer Center and State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, China; [2]Department of Thoracic Surgery, Sichuan Cancer Hospital and Research Institute, Sichuan Cancer Center, Cancer Hospital Affiliated to University of Electronic Science and Technology of China, Chengdu, China; [3]Department of Anesthesiology, The First Hospital of China Medical University, Shenyang, China

Contributions: (I) Conception and design: W He, N Chen; (II) Administrative support: N Chen, Y Han; (III) Provision of study materials: K Yuan, J He; (IV) Collection and assembly of data: C Wang, K Yuan; (V) Data analysis and interpretation: L Peng, W He; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Nianyong Chen. Department of Head and Neck Oncology and Department of Radiation Oncology, Cancer Center and State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, 37 Guoxue Alley, Chengdu 610041, China. Email: n_ychen@hotmail.com.

Background: Although diagnostic methods and treatments have improved over the last few years, the 5-year survival rate of esophageal squamous cell carcinoma (ESCC) patients remains generally poor. The development of high-throughput technology has facilitated great achievements in localization of ESCC-related genes. To take a further step toward a thorough understanding of ESCC at a molecular level, the potential pathogenesis of ESCC needs to be deciphered.

Methods: The interaction of ESCC-related genes was explored by collecting genes associated with ESCC and then performing gene enrichment assays, pathway enrichment assays, pathway crosstalk analysis, and extraction of ESCC-specific subnetwork to describe the relevant biochemical processes.

Results: Through Gene Ontology (GO) enrichment analysis, many molecular functions related to response to chemical, cellular response to stimulus, and cell proliferation were found to be significantly enriched in ESCC-related genes. The results of pathway and pathway crosstalk analysis showed that pathways associated with multiple malignant tumors, the immune system, and metabolic processes were significantly enriched in ESCC-related genes. Through the analysis of specific subnetworks, we obtained some novel ESCC-related potential genes, such as *MUC13*, *GSTO1*, *FIN*, *GRB2*, *CDC25C*, and others.

Conclusions: The molecular mechanism of ESCC is extremely complex. Some inducing factors change the expression status of many genes. The abnormal expression of genes mediates the biological processes involved in immunity and metabolism, apoptosis, and cell proliferation, leading to the occurrence of tumors. The genes *MUC13*, *RYK*, and *FIN* may be potential prognostic indicators of ESCC; *GRB2* and *CDC25C* may be potential targets of ESCC in proliferation. Our work may provide valuable information for further understanding the molecular mechanisms for the development of ESCC.

Keywords: Esophageal squamous cell carcinoma (ESCC); oncogene; network analysis; pathway crosstalk analysis

Submitted Dec 12, 2022. Accepted for publication Jan 10, 2023. Published online Jan 31, 2023.
doi: 10.21037/atm-22-6512
View this article at: https://dx.doi.org/10.21037/atm-22-6512

^ ORCID: 0000-0002-2477-9587.

Introduction

Esophageal squamous cell carcinoma (ESCC), originating from the esophageal mucosa or gland, is malignant and aggressive with a poor prognosis, and accounts for 90% of all cases of esophageal cancer globally. In some eastern Asian and African countries, the incidence of ESCC is extremely high (1,2). An extensive body of research has demonstrated that alcohol and smoking are major risk factors for ESCC. Meanwhile, environmental factors such as the intake of hot beverages, nutritional deficiencies, and limited intake of fruits and vegetables also play a role in the development of this cancer (3,4).

ESCC is clinically challenging and requires multidisciplinary approaches to diagnosis and treatment. The early detection of ESCC is difficult, and most patients are diagnosed with advanced ESCC when major symptoms such as progressive dysphagia and pain behind the sternum occur. Most patients with ESCC require extensive treatment, including chemotherapy, chemoradiotherapy, and/or surgical resection. Palliative treatments are adopted by patients with advanced or metastatic ESCC. However, the limited clinical approaches for the early diagnosis and treatment of ESCC lead to a 10% 5-year survival rate for patients (1,5,6).

Currently, countless genomic analyses have increased the notoriety of these 6 ESCC-implicated genes (*TP53*, *RB1*, *CDKN2A*, *PIK3CA*, *NOTCH1*, *NFE2L2*). *TP53* Pro72 allele increases the risk of ESCC. In previous research, *CDKN2A/RB1* was not expressed in the esophageal mucosa of patients without risk factors whereas p16/pRb expression increased in patients exposed to risk factors or with ESCC (7,8). Although numerous single genetic studies and genetic pathway studies have provided many important insights of the development and prognosis of ESCC, they have not provided clues from the perspective of systems biology. In this study, we conducted a comprehensive collection of ESCC-associated genes from the current studies on ESCC genetic association. Then, we detected the significant biological themes in these genetic factors by performing functional enrichment analyses. Moreover, to ulteriorly reveal the mechanisms of ESCC in a more effective manner, we analyzed the interactions between these biochemical pathways through pathway crosstalk and examined the topological characteristics of these ESCC-associated genes based on human protein-protein interaction (PPI) network. Besides, the ESCC-specific molecular network was deduced and evaluated using the Steiner minimal tree algorithm. This study should provide clues and directions for understanding the molecular mechanism of ESCC from the perspective of systems biology. We present the following article in accordance with the STREGA reporting checklist (available at https://atm.amegroups.com/article/view/10.21037/atm-22-6512/rc).

Highlight box

Key findings

- At the molecular functional level, genes related to chemical response, response to stimuli, and cell proliferation were significantly enriched in ESCC.
- Pathway crosstalk analysis showed that pathways related to multiple malignancies, immune system, and metabolic processes were enriched among ESCC genes.

What is known and what is new?

- The 321 extracted genes, of which aberrant expression mediates biological processes such as immunometabolism, apoptosis, and cell proliferation, are involved in the development and progression of ESCC.
- By ESCC-specific network analysis, we discovered 39 new genes, most of which have never been linked to ESCC. The genes *MUC13*, *RYK*, and *FIN* may be potential prognostic indicators of ESCC. GRB2 and CDC25C may be potential targets of ESCC in proliferation.

What is the implication, and what should change now?

- The molecular mechanisms underlying the development and progression of ESCC require more extensive and in-depth studies.

Methods

Identification of ESCC-related genes

Candidate genes related to ESCC (the standardized term found through medical subject headings [MeSH]) were curated by retrieving the genetic association studies in the PubMed database. We conducted literature retrieval related to ESCC with the terms (Esophageal Squamous Cell Carcinoma AND polymorphism) OR (Esophageal Squamous Cell Carcinoma AND genotype) OR (Esophageal Squamous Cell Carcinoma AND alleles). By 15 July 2022, a total of 1,131 publications were retrieved. After reviewing all 1,131 publications, only the genetic association studies were selected. From the selected publications, we narrowed our selection by focusing on those that reported 1 or more genes significantly associated with ESCC. In addition, the associated genes from several genome-wide association studies (GWAS), showing genetic association at a genome-

wide significance level, were selected. To reduce the number of false-positive findings, we excluded studies that included negative or unrelated associations. We carefully read the full reports of selected studies to ensure that the conclusion was consistent with the content. In this way, we eventually screened out 321 genes from 736 articles. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Functional enrichment analysis of ESCC-related genes

Applying the software the Database for Annotation, Visualization and Integrated Discovery 6.8 (DAVID 6.8; https://david.ncifcrf.gov/), we converted the names of 321 ESCC-related genes obtained from the literature into Entrez Gene IDs. To investigate the functional features of ESCC-related genes, Gene Ontology (GO) term analysis and pathway analysis were applied for functional enrichment analysis. The GO resource (http://www.geneontology.org/) is an international standard classification system of gene function. We performed gene enrichment analysis using GOseq in R Studio 3.4.2 (R Foundation for Statistical Computing, Vienna, Austria). Based on the Wallenius non-central hyper-geometric distribution, this method allows the estimation of gene length preferences and allows us to accurately calculate the probability of a GO term being enriched by the host gene. The enriched P values were corrected by the Benjamini-Hochberg (BH), procedure and we retained terms with P values less than 0.05 as significantly enriched terms. We used directed acyclic graphs (DAGs) as a graphical representation of the GO enrichment analysis of protein genes.

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis

Upon analyzing the entirety of the KEGG database, the pathways in which the study genes were involved were identified and a network of these pathways was established with the aid of the KEGG database's pathway topology. A gene network specific to a single pathway was created through pathway topology analysis. This network was then mapped to the reference gene network from KEGG, resulting in the study gene network. The Entrez Genes IDs were uploaded into the software KOBAS 2.1.1 (http://kobas.cbi.pku.edu.cn/) and were compared with the genes included in each canonical pathway based on the KEGG pathway database. Then, we obtained the corresponding KEGG pathway and P values for each pathway. KEGG pathway analysis was performed using tools in KOBAS, corrected P value [false discovery rate (FDR)] obtained by Fisher's exact test, which were based on hypergeometric distribution, and the enriched P values were corrected by the BH method and a threshold of FDR <0.05 was used to select significantly enriched pathways.

Pathway crosstalk analysis

To further observe the interconnections and interactions between pathways, we conducted a pathway crosstalk analysis of the above-mentioned pathways of significant enrichment. Herein, to test for the overlap between any pair of given pathways, we imported two measurements: the $\text{Jaccard coefficient}(\text{JC}) = \frac{|\text{A} \cap \text{B}|}{|\text{A} \cup \text{B}|}$ and the $\text{overlap coefficient}(OC) = \frac{|\text{A} \cap \text{B}|}{\min(|\text{A}|,|\text{B}|)}$, where A and B are the lists of genes included in the two tested pathways. To construct the pathway crosstalk, we implemented the following procedure:

(I) Select the pathways with PBH ≤0.05. Meanwhile, the number of candidate genes contained in each pathway was restricted to be bigger than or equal to 3, because pathways with too few candidate genes may insufficiently reflect the biological information.

(II) Calculate the number of candidate genes that overlap between any pair of pathways. Delete the pathway pair with less than 3 overlapping genes.

(III) Calculate the overlap of all the pathway pairs that meet the above conditions and rank them by $\text{Score} = \frac{(JC + OC)}{2}$.

(IV) Visualize the selected pathway crosstalk with the software Cytoscape 3.5.1 (9). Indicate the degree of pathways by the size of node, whereby the bigger the degree, the greater the size of the node. Use the thickness of the edge to indicate the score between the pathway pairs, whereby the rougher the edge, the higher the score.

Subnetwork extraction

We merged human protein interaction data downloaded from the Protein Interaction Network Analysis (PINA) platform (updated 21 May 2014) and protein interaction data reported in the recent literature. After removing redundant relationship and self-paired relationship we obtained the final proteins and relationship pairs (10). With these network relationships as a background, we applied

Klein-Ravi algorithm in GenRev, a network-based software package to explore functional relevance of genes via the Steiner minimal tree algorithm that uses a greedy search strategy to merge the smaller trees into larger ones until only one tree connecting all input seeds is built, to extract a subnetwork from the human PPI network by using the 321 ESCC-related genes as seeds (11). In order to verify that this subnetwork was a non-random network, we used Erdos-Renyi algorithm in the igraph R package to generate 100 random background networks with the same number of nodes and edges, and combined the seed gene with the 100 random background networks to generate subnetworks by analyses. Then, we calculated the average values of the shortest-path distance and clustering coefficient. We calculated the number of the average shortest path in the random subnetwork that was shorter than the ESCC networks, denoted the number as n_L; calculated the number of the average clustering coefficient in random subnetwork that was higher than the ESCC network, denoted the number as n_C; calculated the values of p_L and $\frac{p_C}{p_L}=\frac{n_L}{100}$, $p_C=\frac{n_C}{100}$.

Results

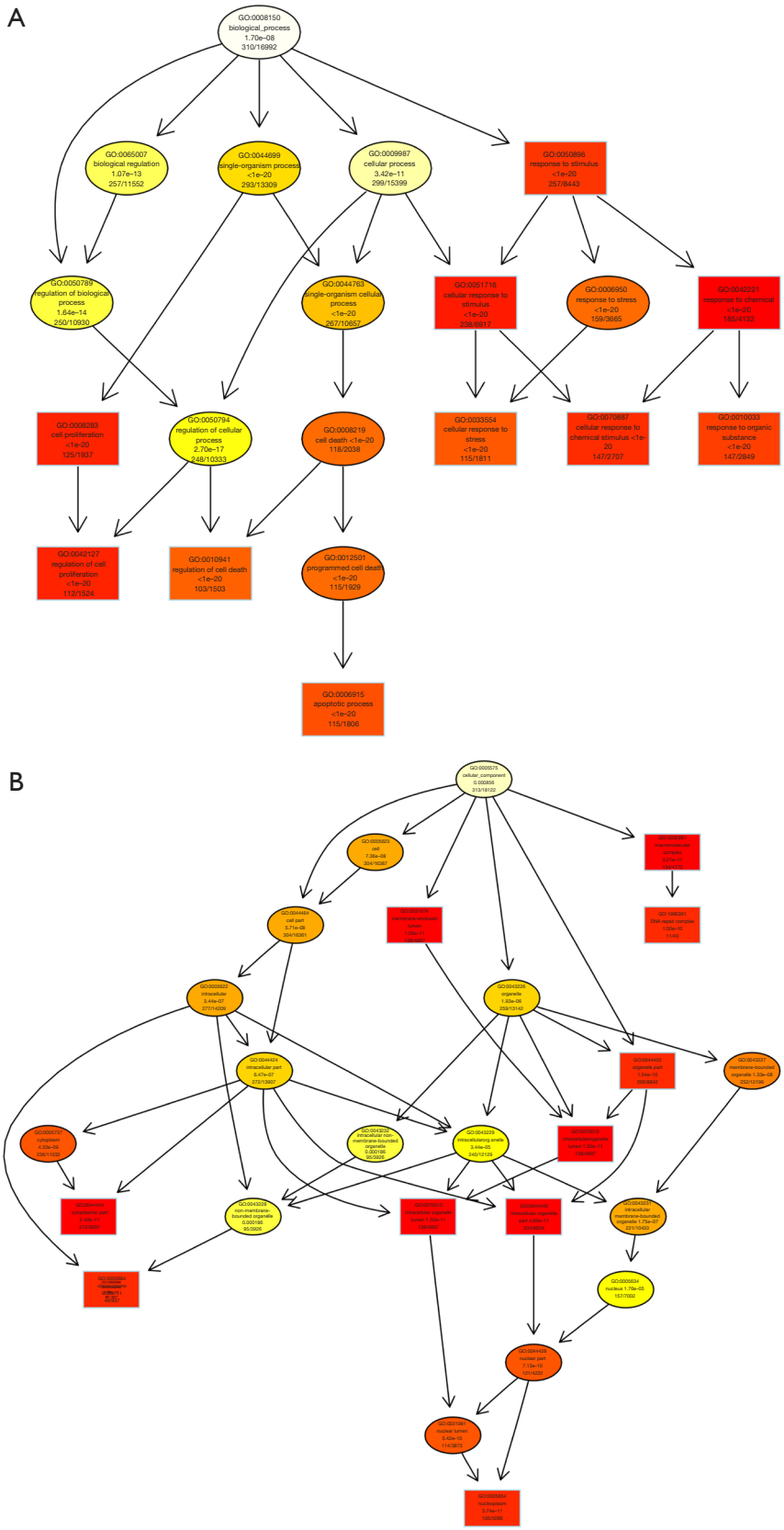
Identification of Genes Reported to be Associated with ESCC

By searching PubMed, we only selected genetic association studies and related genes from several GWAS, which showed a significant degree of genome-wide genetic association. We excluded studies that included either negative or unrelated outcomes to reduce the number of false-positive results.

Altogether, we screened 321 genes reported to be significantly related with ESCC out of 736 articles, which had a range of diverse biological functions. For example, some genes were related to tumor necrosis factor (*TNF*) signaling (e.g., *MMP9*, *CASP3*, *MMP3*, *CASP8*, *AKT1*, *FADD*, *IL6*, *LTA*, *PIK3CA*, *PIK3CB*, *ITCH*, *FAS*, *PTGS2*, *EDN1*, *IL1B*, *TNF*, *TNFRSF1A*, *IL18R1*, and *NFKBIA*), and some genes were related to nucleotide excision repair (e.g., *GTF2H3*, *ERCC2*, *XPA*, *ERCC5*, *ERCC4*, *XPC*, and *ERCC1*), and some genes were related to choline metabolism in cancer (e.g., *mTOR*, *Akt1*, *PIK3CA*, *RHEB*, *EGFR*, *EGF*, *CHKA*, *PIK3CB*, and *WASL*). A detailed list of all genes we found to be associated with ESCC is provided in Table S1.

GO enrichment analysis in ESCC-related genes

The GO database provides a standardized description of the gene products from the function, the participating biological pathway, and cellular localization, which comprises the simple annotation of the gene products. Through GO enrichment analysis, we can gather a rough understanding of the biological functions, pathways, or cellular localization of the differential genes. DAGs of biological process (BP; *Figure 1A*), cellular component (CC; *Figure 1B*), and molecular function (MF; *Figure 1C*) were used to show the GO annotation results. In the GO DAG, annotation moved from more general to more specific as one moved from parent nodes to child nodes. Consequently, a DAG approach was used to provide a clearer understanding of which GO terms were specifically enriched and how these affected other GO terms through upper hierarchies. The top 10 terms with the lowest P value and their parent terms were shown in a GO DAG, the terms with pale marks were significantly enriched, and those with deeper red marks were more significantly enriched. In the DAG of BP, significant enrichment terms, namely, response to chemical (P_{BH}=5.28E-47), cellular response to stimulus (P_{BH}=1.56E-43), and cell proliferation (P_{BH}=1.42E-42) were identified. An example of a significantly enriched term in the DAG of CC is the GO term "intracellular organelle part". This GO term was enriched at a very low FDR (1.46E-08), and two GO terms at upper hierarchies: "nuclear part" and "nuclear lumen" were enriched as a result. These two GO terms together enriched the term "nucleoplasm." The results were consistent with the pathophysiological background of ESCC, indicating that candidate genes were relatively reliable for subsequent bioinformatics analysis. In addition, we used a dot plot to visualize the P values enriched in the first 10 terms in BP, CC, and MF as well as the number of genes involved in each term. At the same time, intuitively, we also found that the genes previously found are in the GO term, which more distinctly showed the correlation. For example, in BP's dot plot (*Figure 2A*), the term "response to stimulus" contained the largest number of genes but had the lowest rich factor. In CC's dot plot (*Figure 2B*), the term "organelle" contained many genes and its p-value was minimal. In the MF's dot plot (*Figure 2C*), we found that the term "damaged DNA binding" had the highest rich factor. Combining all dot plots, we found that there was a trend that the more genes ascribed to a term, the lower the rich factor. This may be because those terms had too many gene numbers. Moreover, we used the



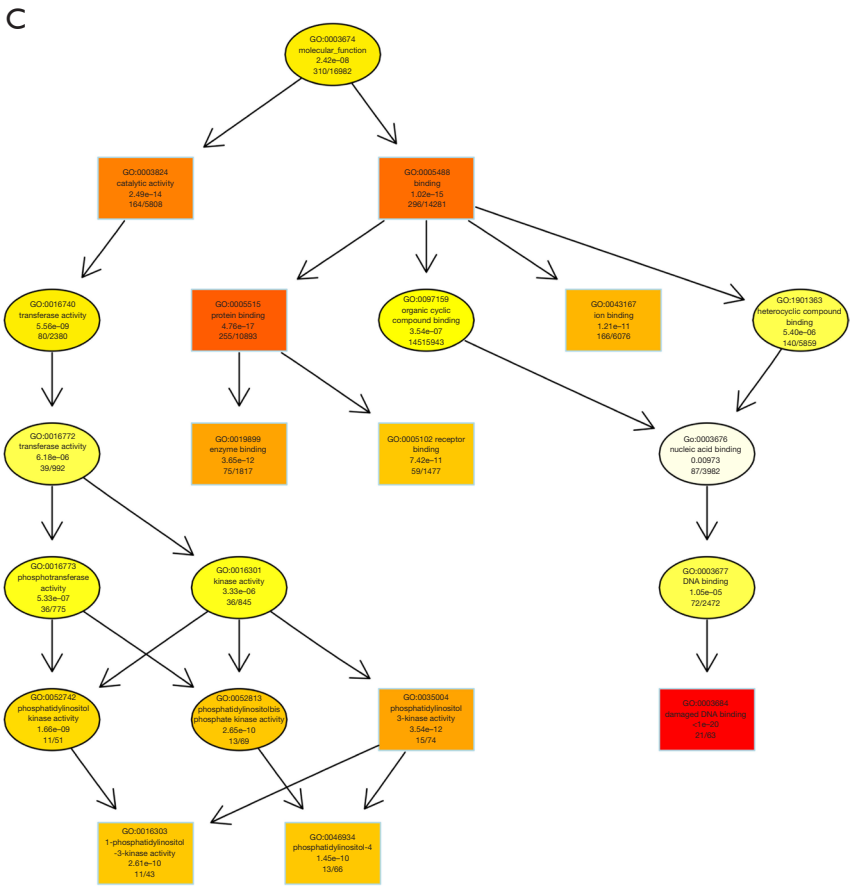


Figure 1 DAG: arrows represent the inclusion of GO terms. The range of functions defined is getting smaller and smaller from top to bottom. The main nodes are the top 10 GO enrichment analysis results, represented by rectangles. Other related GO terms are shown by inclusion relationships, the deeper the color, the higher the degree of enrichment. DAG, directed acyclic graph; GO, Gene Ontology.

ggplot2 package in the Goseq software to graph the top 10 GO terms in BP, CC, and MF (*Figure 3*).

KEGG pathway enrichment analysis in ESCC-related genes

Pathway analysis enables researchers to map the genes, proteins, or molecules onto a particular class of metabolic or regulatory networks, or according to an individuals' molecular set of functions, in order to form their own specific pathway. This is very important for elucidating the molecular mechanism of action and identifying biomarkers. We uploaded the differential genes into KOBAS 2.1.1 software and compared them to the genes contained in the canonical approaches based on the KEGG pathway database, and we enriched 240 pathways. A total of 31 pathways with corrected P value (FDR) <0.05 were retained as the significantly enriched pathways (Table S2). Most of the pathways were related to pathophysiological process of cancer. Among them, several pathways are closely related to certain types of malignant tumors, including bladder cancer (ranked 2nd), melanoma (ranked 4th), prostate cancer (ranked 5th), colorectal cancer (ranked 9th), non-small cell lung cancer (NSCLC; ranked 11th), endometrial cancer (ranked 14th), and glioma (ranked 20th). Many studies have clarified the role of p53 and TNF in the development of cancer. p53 is a well-known tumor suppressor gene. However, in ESCC, the p53 pathway is inactivated by TP53 mutations, so it has carcinogenic effects (12-14). Numerous studies have shown that TNF has antitumor activity and is also an endogenous tumor promoter (15). TNF is down-regulated in ESCC tissues and multivariate analysis has shown that down-regulation of TNF is independently associated

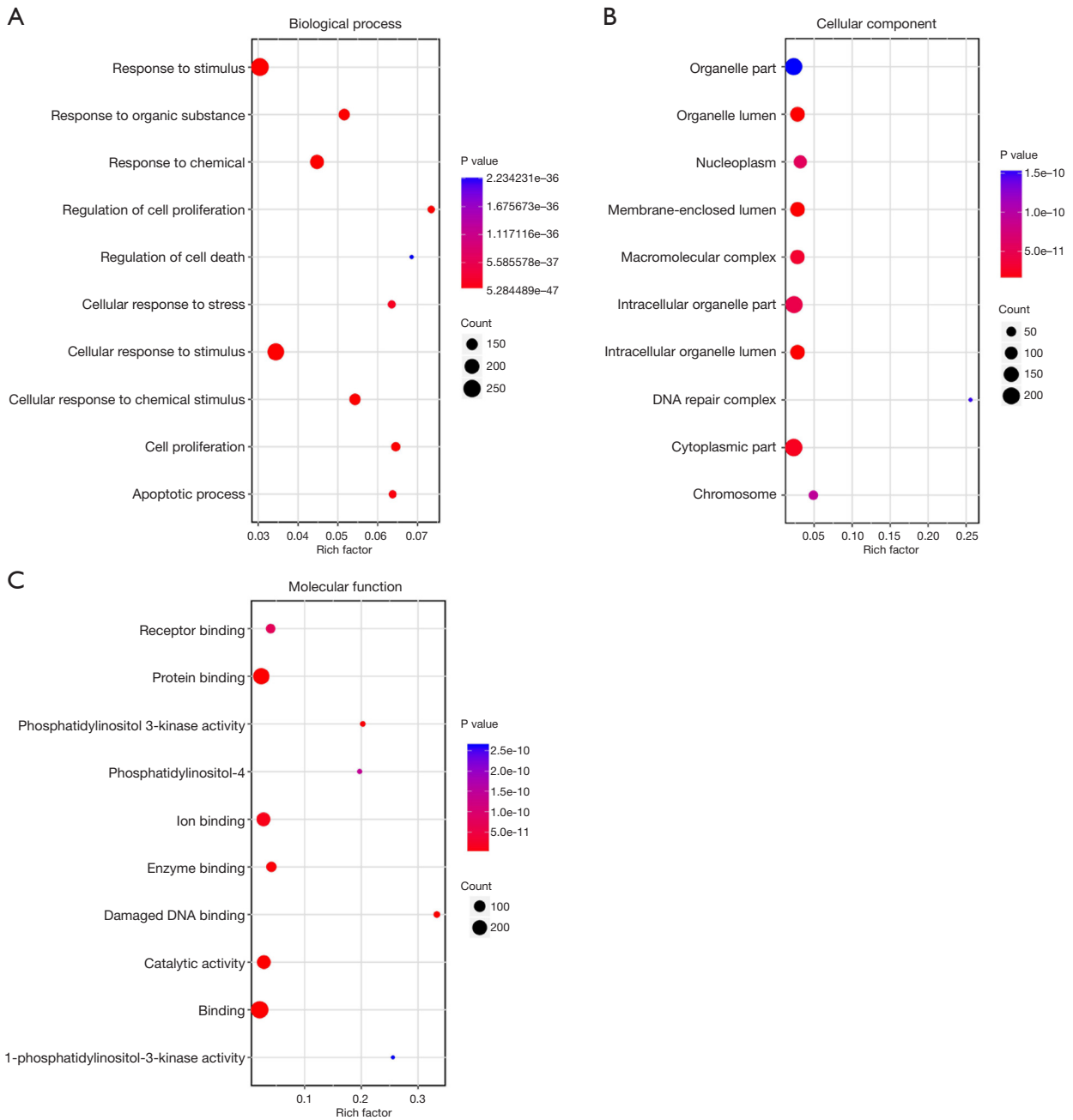


Figure 2 Dot plot: the graph shows the rich factor and P values for the top 10 GO terms. The size of the solid dot indicates the number of ESCC-related genes in this term. GO, Gene Ontology; ESCC, esophageal squamous cell carcinoma.

with early local tumor recurrence (16,17). In cellular metabolism, we obtained pathways such as proteoglycans in cancer, endocrine resistance, HIF-1 signaling pathway, and microRNAs in cancer. It was worth mentioning that in the microRNAs in cancer pathway, MYC binding protein (MYCBP) was identified as a direct target of miR-26 (18). In drug metabolism, we found pathways such as platinum drug resistance (ranked 1st), metabolism of xenobiotics by cytochrome P450, and drug metabolism-cytochrome P450. Cisplatin is an important part of chemotherapy for esophageal cancer (19). However, in the use of platinum drugs, some challenges such as patients' partial

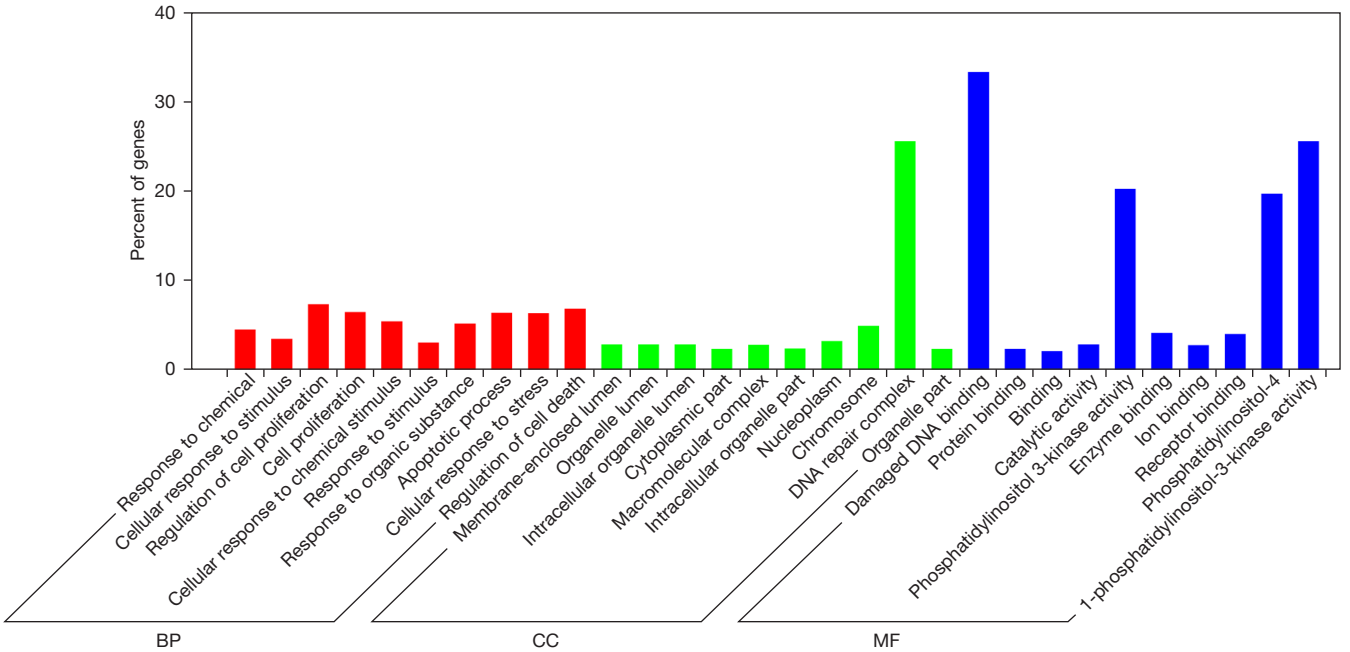


Figure 3 Ggplot2 of top 10 GO terms in BP, CC, and MF. Vertical coordinate indicates: the proportion of obtained ESCC-related genes in all genes of each pathway. BP, biological process; CC, cellular component; MF, molecular function; GO, Gene Ontology; ESCC, esophageal squamous cell carcinoma.

antitumor response, drug resistance development, and tumor recurrence limit the patient's life expectancy (20). Furthermore, immune-associated biological processes such as graft-versus-host disease and allograft rejection were also significantly enriched.

Pathway crosstalk analysis among significantly enriched pathways

To further observe the interconnections and interactions between significantly enriched pathways, we conducted a pathway crosstalk analysis of the 31 significantly enriched pathways mentioned above. The approach was based on the assumption that two pathways were considered to crosstalk if they shared a proportion of ESCC-related genes (21). A total of 30 of the above 31 significantly enriched pathways were in accordance with the crosstalk analysis criteria, namely, each pathway contained 3 or more genes, each of which had at least two genes sharing with 1 or more other pathways. The base excision repair pathway was the only non-compliant pathway due to the number of genes overlapping with other pathways being less than or equal to 2. All paths formed by these approaches were used to construct pathway crosstalk and the level of overlap between the two pathways was measured on the basis of the average scores of the coefficients JC and OC. Based on their crosstalk, pathways could be broadly divided into three main parts, each containing more interactions than other pathways and might involve in the same or similar biological processes (*Figure 4*). The first part consisted of 16 pathways, including 9 of the top 10 significantly enriched pathways of KEGG. Most of these 16 pathways were specific cancer pathways, such as prostate cancer and NSCLC. Three of these pathways shared some genes with the platinum drug resistance pathway, and these three pathways were linked to drug metabolism and chemical carcinogenesis. The second part contained seven pathways, which connected the other two parts of pathways that shared no genes. These seven pathways were mainly involved in the regulation of many kinds of receptors and cytokines in cells. At the same time, they played an important role in the regulation of cell proliferation, apoptosis, inflammation in vivo, infection, and other physiological processes. The third pathway was mainly associated with immunity and inflammation. As indicated above, path crosstalk analysis can provide important clues to the development and prognosis of ESCC and provide insights into the ECSS mechanism.

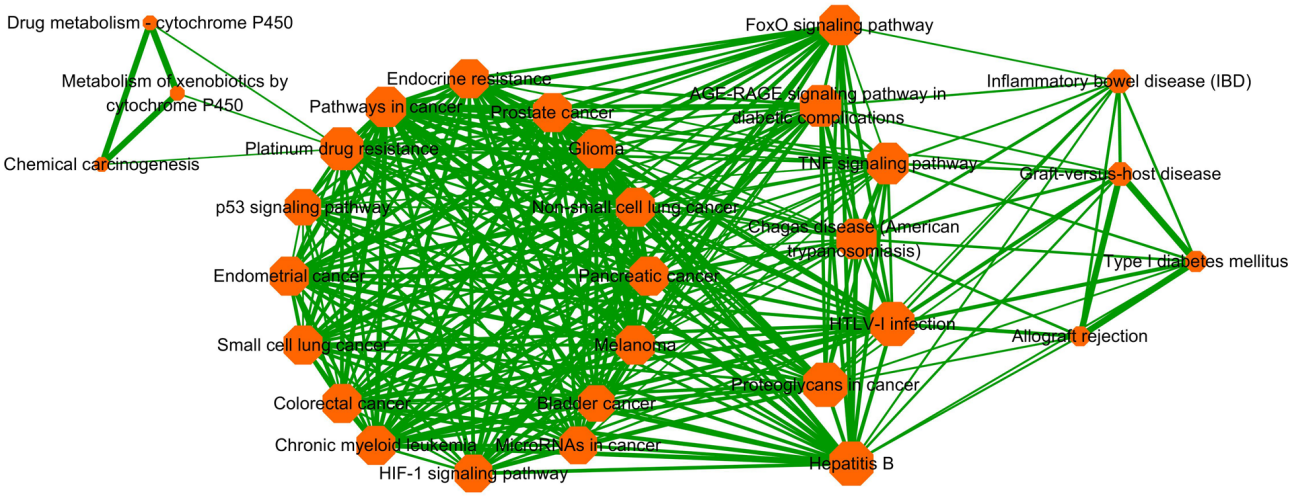


Figure 4 Pathway crosstalk. Nodes represent pathways and edges represent crosstalk between pathways. Node size corresponds to the number of ESCC-related genes found in the corresponding pathway. Edge width corresponds to the score of the related pathways. ESCC, esophageal squamous cell carcinoma.

Subnetwork extraction

To refine the ESCC-related interactions, we used the Steiner's minimal tree algorithm to extract the ESCC subnetwork from the human PPI network and recently published protein interaction pairs (232,866 pairs of 17,379 proteins were involved). This approach attempts to connect a maximum number of input nodes with a minimum number of linked nodes (we used 308 nodes). As shown in *Figure 5*, 347 nodes and 1,339 edges were contained. A total of 308 out of the 321 ESCC-related genes were included in the ESCC-specific network, accounting for about 95.95% of the candidate genes and 88.76% of the genes in the ESCC-specific network, demonstrating a high coverage of ESCC-related genes in the subnetwork. It was noteworthy that we found 39 additional proteins, some of which had been reported as associated with ESCC (*Table 1*). In addition, to verify that the sub-network was a non-random network, we used the Erdos-Renyi model in the igraph R package to generate 100 random sub-networks. We calculated the arithmetic average values of the shortest path distance and clustering coefficients in the random sub-networks. The average shortest path distance value and average clustering coefficient were compared with the corresponding values of the ESCC-specific network. For these random background networks, the average shortest path distance was 2.86, bigger than the shortest path distance of ESCC specific network (shortest path distance was 2.60; $P_L<0.01$). The average clustering coefficient of random sub-networks was 0.19, which was significantly smaller than that of ESCC-specific networks (clustering coefficient was 0.33, $P_C<0.01$). Therefore, the ESCC-specific network extracted from the entire PPI network was a non-random network.

Discussion

As a highly malignant and aggressive tumor, ESCC is not sensitive to radiotherapy and chemotherapy. In the past few decades, studies of human participants, animals, or cell models have gained insight into the molecular mechanisms of ESCC. Although increasing numbers of genes/proteins are believed to be associated with this disease, with the development of high-throughput technologies, many genes/proteins are considered potential targets for detection or treatment. However, the understanding of the biological processes associated with the pathogenesis of ESCC at the molecular level is far from complete. Therefore, there is a need to decode the underlying pathogenesis of ESCC at the level of systems biology. In this study, we explored the interactions of these genes by collecting genes related to ESCC as well as utilizing pathway and network analysis systems. We have provided a comprehensive and systematic framework for mapping related biochemical processes.

Through biological function enrichment analysis, we obtained specific biological processes caused by ESCC-related genes. The GO enrichment analysis provided us with a rough interpretation of the biological functions,

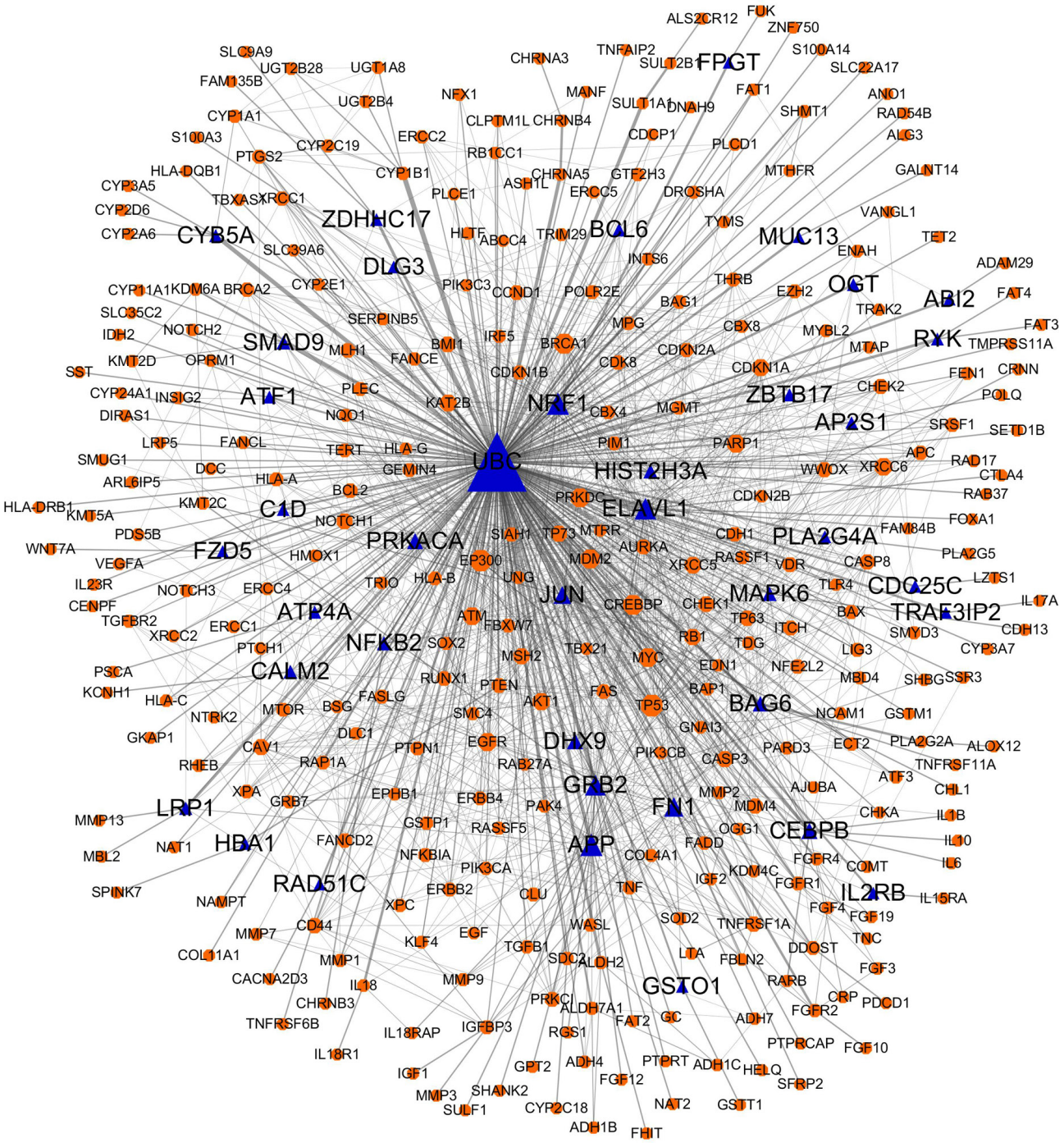


Figure 5 ESCC-specific network. The orange dots represent ESCC-related genes that we entered, and the blue triangles represent newly discovered genes. The size of the node is related to the node's degree in the ESCC-specific network. ESCC, esophageal squamous cell carcinoma.

Table 1 39 genes associated with ESCC discovered by ESCC-specific network

Gene abbreviations	Gene ID	Species	Gene name
RAD51C	5889	Homo sapiens	RAD51 paralog C (RAD51C)
GRB2	2885	Homo sapiens	Growth factor receptor bound protein 2 (GRB2)
AP2S1	1175	Homo sapiens	Adaptor related protein complex 2 sigma 1 subunit (AP2S1)
ABI2	10152	Homo sapiens	Abl interactor 2 (ABI2)
NFKB2	4791	Homo sapiens	Nuclear factor kappa B subunit 2 (NFKB2)
ZBTB17	7709	Homo sapiens	Zinc finger and BTB domain containing 17 (ZBTB17)
ATF1	466	Homo sapiens	Activating transcription factor 1 (ATF1)
APP	351	Homo sapiens	Amyloid beta precursor protein (APP)
BAG6	7917	Homo sapiens	BCL2 associated athanogene 6 (BAG6)
DLG3	1741	Homo sapiens	Discs large MAGUK scaffold protein 3 (DLG3)
BCL6	604	Homo sapiens	B-cell CLL/lymphoma 6 (BCL6)
PRKACA	5566	Homo sapiens	Protein kinase camp-activated catalytic subunit alpha (PRKACA)
OGT	8473	Homo sapiens	O-linked N-acetylglucosamine (glcnac) transferase (OGT)
GSTO1	9446	Homo sapiens	Glutathione S-transferase omega 1 (GSTO1)
MUC13	56667	Homo sapiens	Mucin 13, cell surface associated (MUC13)
FN1	2335	Homo sapiens	Fibronectin 1 (FN1)
HIST2H3A	333932	Homo sapiens	Histone cluster 2 H3 family member a (HIST2H3A)
DHX9	1660	Homo sapiens	Dexh-box helicase 9 (DHX9)
IL2RB	3560	Homo sapiens	Interleukin 2 receptor subunit beta (IL2RB)
CEBPB	1051	Homo sapiens	CCAAT/enhancer binding protein beta (CEBPB)
SMAD9	4093	Homo sapiens	SMAD family member 9 (SMAD9)
ATP4A	495	Homo sapiens	ATPase H+/K+ transporting alpha subunit (ATP4A)
RYK	6259	Homo sapiens	Receptor-like tyrosine kinase (RYK)
ELAVL1	1994	Homo sapiens	ELAV like RNA binding protein 1 (ELAVL1)
CYB5A	1528	Homo sapiens	Cytochrome b5 type A (CYB5A)
HBA1	3039	Homo sapiens	Hemoglobin subunit alpha 1 (HBA1)
FZD5	7855	Homo sapiens	Frizzled class receptor 5 (FZD5)
CDC25C	995	Homo sapiens	Cell division cycle 25C (CDC25C)
TRAF3IP2	10758	Homo sapiens	TRAF3 interacting protein 2 (TRAF3IP2)
NRF1	4899	Homo sapiens	Nuclear respiratory factor 1 (NRF1)
PLA2G4A	5321	Homo sapiens	Phospholipase A2 group IVA (PLA2G4A)
ZDHHC17	23390	Homo sapiens	Zinc finger DHHC-type containing 17 (ZDHHC17)
LRP1	4035	Homo sapiens	LDL receptor related protein 1 (LRP1)
MAPK6	5597	Homo sapiens	Mitogen-activated protein kinase 6 (MAPK6)
JUN	3725	Homo sapiens	Jun proto-oncogene, AP-1 transcription factor subunit (JUN)
UBC	7316	Homo sapiens	Ubiquitin C (UBC)
FPGT	8790	Homo sapiens	Fucose-1-phosphate guanylyltransferase (FPGT)
CALM2	805	Homo sapiens	Calmodulin 2 (CALM2)
C1D	10438	Homo sapiens	C1D nuclear receptor corepressor (C1D)

ESCC, esophageal squamous cell carcinoma.

pathways, or cellular localization of the differential genes, convincingly demonstrating that these genes involved in immune response, metabolic regulation, cell proliferation, apoptosis, and drug response. In cluster analysis, GO terms relative to apoptosis, cell metabolism, regulation of cell development, DNA damage repair, angiogenesis, and xenobiotic metabolism turned out high enrichment scores, which was consistent with our perception of ESCC so far. For example, in GO MF, key terms such as catalytic activity, binding, protein binding, and damaged DNA binding, along with other enriched terms delineated that the cancer cell has an increased damage repair function and can fully repair the DNA damage caused by chemotherapeutic drugs, which at least partially provides a reason for the unsatisfactory effect of chemotherapy in the treatment of ESCC.

The results of pathway analysis showed that genes related to esophageal cancer contain multiple tumor pathways, revealing its commonality with other malignancies. We found that 7 pathways: bladder cancer, prostate cancer, melanoma, NSCLC, endometrial cancer, glioma, and pancreatic cancer, were in the same functional annotation clustering, and these pathways all contain the genes *TP53*, *RB1*, *EGFR*, *EGF*, and *CCND1*, which was highly consistent with our understanding of the mechanisms of cancer development. *TP53* encodes a tumor suppressor protein containing transcriptional activation, DNA binding, and oligomerization domains. The encoded protein responds to diverse cellular stresses to regulate expression of target genes, thereby inducing cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism. In ESCC, over 90% of *TP53* mutations and inactivation of the p53 pathway are associated with patient prognosis and resistance to radiotherapy and chemotherapy. *CCND1* mutations exceeding 20%, deletions of *CDKN2A*, *RB1*, and *CDKN2A*, and abnormal methylation impair the cell cycle. *EP300*, *CREBBP*, and *NOTCH* (*NOTCH1*, *NOTCH2*, *NOTCH3*) are associated with epigenetic processes (12,22,23). Moreover, *CCND1*, *EGFR*, and *ERBB2* are likely to be driver genes for the development of ESCC (24,25). Previous studies have shown that *RASSF1* is a very promising tumor suppressor gene, and the extent of its transcript expression and methylation is related to tumor progression and survival of ESCC patients (26,27). In the RAS signaling pathway and the RAP1 signaling pathway, *RASSF1* interacts with and influences *PIK3CA*, *AKT1*, *EGF*, and *EGFR*. These genes are involved in the formation, proliferation, migration, and angiogenesis of tumors. *AKT1* is significantly associated with local recurrence of ESCC (28). Tumor inhibitors of the *RASSF* family act as *RAS* apoptosis and senescence effectors. It is speculated that inactivation of the *RASSFlA* tumor suppressor promotes *K-RAS*-mediated transformation by uncoupling it with apoptotic pathways such as the Hippo pathway (29,30). In addition, we observed that in the cell cycle, proliferation, the *PI3K-AKT* signaling pathway, allograft rejection reaction, *MYC*, *COX*-2, *p53*, and *RB1* also showed associations. As indicated by these results, a variety of genes and signaling molecules serve as bridges and connect to each other, join to many signaling pathways and form a complex network of ESCC molecules.

In crosstalk analysis, we identified three major modules. The first module contains 16 pathways, most of which are specific cancer pathways, such as prostate cancer, NSCLC, pancreatic cancer, and the like. Among these 16 pathways, 3 share some genes with the platinum drug resistance pathway. These three pathways are related to drug metabolism and chemical carcinogenesis: drug metabolism—cytochrome *P450*, metabolism of xenobiotics by cytochrome P450, and chemical carcinogenesis. The *CYP1B1* mutation was significantly associated with ESCC risk (31,32). Studies have shown that the most important issue associated with ESCC treatment is the intrinsic resistance of chemotherapeutic drugs, and multidrug resistance is increasingly common in patients with ESCC (33,34). Apoptosis is one of the most critical processes in cell proliferation and a key molecular mechanism for anticancer therapy. The *PI3K/Akt/mTOR* pathway is involved in cell proliferation, differentiation, survival, apoptosis, and metastasis (35-37). Sensitized drugs promote the action of oxaliplatin by significantly inducing apoptosis and modulating the *PI3K/Akt/mTOR* pathway, counteracting resistance to chemotherapy (38,39). The second module contains 7 pathways which connect other pathways that do not share genes. These 7 pathways are mainly involved in the regulation of various receptors and cytokines in cells, and they also play an important role in the regulation of cell proliferation, apoptosis, inflammation, infection, and other physiological processes. By analysis of module II and module III, pathways related to immune response, cell adhesion (i.e., *HTLV-I* infection), allograft rejection, graft-versus-host disease, the *FoxO* signaling pathway, shared genes such as *IL6*, *FAS*, *END1*, *IL1B*, *TNF*, *TNFRSF1A*, *NFKBIA*, *HLA-A*, *HLA-B*, and *HLA-C*. In these three pathways (graft-versus-host disease, type I diabetes mellitus, and allograft rejection), we found similar processes and common genes such as the *HLA* family (*HLA-A*, *HLA-B*,

HLA-C, *HAL-G*, *HLA- DQB1*, *HLA-DRB1*), *FAS*, *FASLG*, and some cytokines (*IL6*, *IL1B*). By targeting the *FASLG* gene, miR-21 down-regulates cell growth, invades, and induces apoptosis, and by combining with *FASLG* treatment methods, the efficacy of radiotherapy can be enhanced, and unnecessary angiogenesis-promoting effects can be reduced (40-42). By modulating *HLA-G* expression, miR-148a is indicated to be involved in the carcinogenesis of primary ESCC. The current results indicate that miR-148a is a potential biomarker for ESCC (43). *HLA* class I is critical for tumor immunity, and its degree of expression is an independent prognostic factor for relapse-free survival (44,45). Studies have shown that *HLA-G* is highly expressed in ESCC and can be used as a novel tumor marker (43,46-48). *HLA-II* molecules are mainly encoded by *DP*, *DQ*, and *DR* genes, expressed in immune cells, and are responsible for presenting antigenic peptides to $CD4^+$ T cells to trigger the expansion and differentiation of these T cells and induce a series of antigen-specific immunity responses. Studies have shown that aberrant methylation of *HLA-DQB1* and *HLA-DAB1* at different sites is significantly associated with ESCC differentiation and late stages, and their methylation is conducive to the abnormal expression of *HLA-II* (49-51). Therefore, abnormal expression of *HLA-II* may lead to immune response or autoimmunity deficiency in some diseases. Their abnormal expression inhibits the function of immune cells such as natural killer (NK) cells and T cells, which promotes the development of ESCC. These pathways related to the immune response and cell adhesion movement profoundly influence the occurrence and development of ESCC, and they also shed a new light for understanding diseases through molecular mechanisms.

Remarkably, in the ESCC-specific network, we discovered 39 new genes, most of which have never been linked to ESCC, such as *APP*, *TRAF3IP2*, *ZBTB17*, *NRF1*, *CEBPB*, *HBA1*, *ABI2*, *CALM2*, *RAD51C*, and *AP2S1* (*Table 1*). Studies have shown that RYK may be an important predictor of ESCC clinical stage staging and prognosis (52). The combination of *MUC13/MUC20* expression is a potential prognostic indicator of neoadjuvant chemotherapy and postoperative ESCC (53,54). *GSTO1* may also be a potential biomarker for early detection (55). The drug resistance/response mechanism of ESCC is very complicated. A study found that differential genes are also key cancer-promoting genes related to chemotherapy resistance through PPI, such as *BMP1* and *DBF4* (56); prognosis-related genes, such as *TTK* and *KIF18A*. These studies can help improve the survival rate of patients (57).

JUN is a chromosome region of human malignancy translocations and deletions associated with various signal transduction pathways (58,59). In ESCC, *BCL6* is up-regulated, and its acylation is controlled by *HDAC* and SIRT-dependent mechanisms (60). Interference with this process may lead to cell cycle arrest and apoptosis. Its biological function may indicate that it can become a new regulator in esophageal cancer cells (61,62). *GRB2* is a prominent node involved in the naturally regulated cell proliferation-related pathways of the tyrosine kinase signaling pathway and *Erk1/Erk2* Mapk signaling pathway. *UBC* participates in cell cycle progression and is upregulated in more than 70% of ESCC tissues (63-65). *UEB2C* directly interferes with the level of cyclin B1 protein and alters the proliferation and cell cycle characteristics of ESCC (66). Due to activation of the *p53/p21* pathway, *CDC25C* is thus reduced, leading to cell cycle arrest (67,68). Although the quantity and quality of PPI data have greatly improved in recent years, the human PPI network is far from being completed. In addition, due to current technology limitations, there may be some false positive results in the PPI data. These potential deviations associated with human PPI networks may have affected our results.

Conclusions

The mechanism of occurrence and development of ESCC is extremely complex and involves a variety of factors, such as genetic and environmental factors. In this study, we used a systems biology framework to select candidate genes and performed a variety of analyses on ESCC. By integrating information from tandem analysis of GO, pathways, and pathway crosstalk, we found that ESCC was associated with a variety of cancer pathways, cell proliferation, apoptosis, immune responses, and drug metabolism. In addition, in ESCC-specific network, some of these additional genes have been reported to be associated with ESCC. In order to reveal the molecular mechanism of ESCC, we initially constructed its molecular network. Our research facilitates a more in-depth understanding of the mechanism of ESCC.

Acknowledgments

The primary results of the abstract in this study were presented as a meeting poster (No. PS02.037) in the 2018 International Society for Diseases of the Esophagus.
Funding: This work was funded by Bethune Charitable

Foundation (No. HZB-20190528-19).

Footnote

Reporting Checklist: The authors have completed the STREGA reporting checklist. Available at https://atm.amegroups.com/article/view/10.21037/atm-22-6512/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://atm.amegroups.com/article/view/10.21037/atm-22-6512/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

References

1. Smyth EC, Lagergren J, Fitzgerald RC, et al. Oesophageal cancer. Nat Rev Dis Primers 2017;3:17048.
2. Tao Y, Chai D, Ma L, et al. Identification of distinct gene expression profiles between esophageal squamous cell carcinoma and adjacent normal epithelial tissues. Tohoku J Exp Med 2012;226:301-11.
3. Testa U, Castelli G, Pelosi E. Esophageal Cancer: Genomic and Molecular Characterization, Stem Cell Compartment and Clonal Evolution. Medicines (Basel) 2017;4:67.
4. Holmes RS, Vaughan TL. Epidemiology and pathogenesis of esophageal cancer. Semin Radiat Oncol 2007;17:2-9.
5. Lagergren J, Smyth E, Cunningham D, et al. Oesophageal cancer. Lancet 2017;390:2383-96.
6. Song Y, Li L, Ou Y, et al. Identification of genomic alterations in oesophageal squamous cell cancer. Nature 2014;509:91-5.
7. Müller LB, Meurer L, Lopes AB, et al. Stepwise expression of CDKN2A and RB1 proteins in esophageal mucosa from patients at high risk for squamous cell carcinoma. Appl Immunohistochem Mol Morphol 2014;22:669-73.
8. Ni X, Wu P, Wu J, et al. Hyperthermic intraperitoneal perfusion chemotherapy and response evaluation in patients with gastric cancer and malignant ascites. Oncol Lett 2017;14:1691-6.
9. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498-504.
10. Cowley MJ, Pinese M, Kassahn KS, et al. PINA v2.0: mining interactome modules. Nucleic Acids Res 2012;40:D862-5.
11. Zheng S, Zhao Z. GenRev: exploring functional relevance of genes in molecular networks. Genomics 2012;99:183-8.
12. Kishino T, Niwa T, Yamashita S, et al. Integrated analysis of DNA methylation and mutations in esophageal squamous cell carcinoma. Mol Carcinog 2016;55:2077-88.
13. Wu Z, Zhou J, Zhang X, et al. Reprogramming of the esophageal squamous carcinoma epigenome by SOX2 promotes ADAR1 dependence. Nat Genet 2021;53:881-94.
14. Li L, Xu J, Qiu G, et al. Epigenomic characterization of a p53-regulated 3p22.2 tumor suppressor that inhibits STAT3 phosphorylation via protein docking and is frequently methylated in esophageal and other carcinomas. Theranostics 2018;8:61-77.
15. Scott KA, Moore RJ, Arnott CH, et al. An anti-tumor necrosis factor-alpha antibody inhibits the development of experimental skin tumors. Mol Cancer Ther 2003;2:445-51.
16. Xie Y, Wang B. Downregulation of TNFAIP2 suppresses proliferation and metastasis in esophageal squamous cell carcinoma through activation of the Wnt/β-catenin signaling pathway. Oncol Rep 2017;37:2920-8.
17. Chen Y, Wang Y, Song H, et al. Expression profile of apoptosis-related genes potentially explains early recurrence after definitive chemoradiation in esophageal squamous cell carcinoma. Tumour Biol 2014;35:4339-46.
18. Li J, Liang Y, Lv H, et al. miR-26a and miR-26b inhibit esophageal squamous cancer cell proliferation through suppression of c-MYC pathway. Gene 2017;625:1-9.
19. Lebwohl D, Canetta R. Clinical development of platinum complexes in cancer therapy: an historical perspective and an update. Eur J Cancer 1998;34:1522-34.

20. Oberoi HS, Nukolova NV, Kabanov AV, et al. Nanocarriers for delivery of platinum anticancer drugs. Adv Drug Deliv Rev 2013;65:1667-85.
21. Jia P, Kao CF, Kuo PH, et al. A comprehensive network and pathway analysis of candidate genes in major depressive disorder. BMC Syst Biol 2011;5 Suppl 3:S12.
22. Gao YB, Chen ZL, Li JG, et al. Genetic landscape of esophageal squamous cell carcinoma. Nat Genet 2014;46:1097-102.
23. Huang YH, Cai K, Xu PP, et al. CREBBP/EP300 mutations promoted tumor progression in diffuse large B-cell lymphoma through altering tumor-associated macrophage polarization via FBXW7-NOTCH-CCL2/CSF1 axis. Signal Transduct Target Ther 2021;6:10.
24. Sawada G, Niida A, Hirata H, et al. An Integrative Analysis to Identify Driver Genes in Esophageal Squamous Cell Carcinoma. PLoS One 2015;10:e0139808.
25. Cui D, Zhu Y, Yan D, et al. Dual inhibition of cMET and EGFR by microRNA-338-5p suppresses metastasis of esophageal squamous cell carcinoma. Carcinogenesis 2021;42:995-1007.
26. Perez-Janices N, Blanco-Luquin I, Torrea N, et al. Differential involvement of RASSF2 hypermethylation in breast cancer subtypes and their prognosis. Oncotarget 2015;6:23944-58.
27. Haluskova J, Lachvac L, Nagy V. The investigation of GSTP1, APC and RASSF1 gene promoter hypermethylation in urine DNA of prostate-diseased patients. Bratisl Lek Listy 2015;116:79-82.
28. Tian B, Liu J, Zhang N, et al. Oncogenic SNORD12B activates the AKT-mTOR-4EBP1 signaling in esophageal squamous cell carcinoma via nucleus partitioning of PP-1α. Oncogene 2021;40:3734-47.
29. Schmidt ML, Hobbing KR, Donninger H, et al. RASSF1A Deficiency Enhances RAS-Driven Lung Tumorigenesis. Cancer Res 2018;78:2614-23.
30. Yu W, Chu L, Zhao K, et al. A nomogram based on phosphorylated AKT1 for predicting locoregional recurrence in patients with oesophageal squamous cell carcinoma. J Cancer 2017;8:3755-63.
31. Hyland PL, Freedman ND, Hu N, et al. Genetic variants in sex hormone metabolic pathway genes and risk of esophageal squamous cell carcinoma. Carcinogenesis 2013;34:1062-8.
32. Malik MA, Upadhyay R, Mittal RD, et al. Association of xenobiotic metabolizing enzymes genetic polymorphisms with esophageal cancer in Kashmir Valley and influence of environmental factors. Nutr Cancer 2010;62:734-42.
33. Zhou P, Zhang R, Wang Y, et al. Cepharanthine hydrochloride reverses the mdr1 (P-glycoprotein)-mediated esophageal squamous cell carcinoma cell cisplatin resistance through JNK and p53 signals. Oncotarget 2017;8:111144-60.
34. Chi Y, Wang F, Zhang Y, et al. Apatinib inhibits tumour progression and promotes antitumour efficacy of cytotoxic drugs in oesophageal squamous cell carcinoma. J Cell Mol Med 2022;26:1905-17.
35. Li T, Wang G. Computer-aided targeting of the PI3K/Akt/mTOR pathway: toxicity reduction and therapeutic opportunities. Int J Mol Sci 2014;15:18856-91.
36. Furusawa S, Wu J. The effects of biscoclaurine alkaloid cepharanthine on mammalian cells: implications for cancer, shock, and inflammatory diseases. Life Sci 2007;80:1073-9.
37. Wang L, Zhang Z, Yu X, et al. SOX9/miR-203a axis drives PI3K/AKT signaling to promote esophageal cancer progression. Cancer Lett 2020;468:14-26.
38. Lin JJ, Su JH, Tsai CC, et al. 11-epi-Sinulariolide acetate reduces cell migration and invasion of human hepatocellular carcinoma by reducing the activation of ERK1/2, p38MAPK and FAK/PI3K/AKT/mTOR signaling pathways. Mar Drugs 2014;12:4783-98.
39. Zhou J, Jiang YY, Chen H, et al. Tanshinone I attenuates the malignant biological properties of ovarian cancer by inducing apoptosis and autophagy via the inactivation of PI3K/AKT/mTOR pathway. Cell Prolif 2020;53:e12739.
40. Blaes J, Thomé CM, Pfenning PN, et al. Inhibition of CD95/CD95L (FAS/FASLG) Signaling with APG101 Prevents Invasion and Enhances Radiation Therapy for Glioblastoma. Mol Cancer Res 2018;16:767-76.
41. Wang N, Zhang CQ, He JH, et al. MiR-21 down-regulation suppresses cell growth, invasion and induces cell apoptosis by targeting FASL, TIMP3, and RECK genes in esophageal carcinoma. Dig Dis Sci 2013;58:1863-70.
42. Du X, Hong L, Sun L, et al. miR-21 induces endothelial progenitor cells proliferation and angiogenesis via targeting FASLG and is a potential prognostic marker in deep venous thrombosis. J Transl Med 2019;17:270.
43. Chen Q, Luo G, Zhang X. MiR-148a modulates HLA-G expression and influences tumor apoptosis in esophageal squamous cell carcinoma. Exp Ther Med 2017;14:4448-52.
44. Ito S, Okano S, Morita M, et al. Expression of PD-L1 and HLA Class I in Esophageal Squamous Cell Carcinoma:

Prognostic Factors for Patient Outcome. Ann Surg Oncol 2016;23:508-15.
45. Jhunjhunwala S, Hammer C, Delamarre L. Antigen presentation in cancer: insights into tumour immunogenicity and immune evasion. Nat Rev Cancer 2021;21:298-312.
46. Golyan FF, Abbaszadegan MR, Forghanifard MM. TWIST1, MMP-21, and HLAG-1 co-overexpression is associated with ESCC aggressiveness. J Cell Biochem 2019;120:14838-46.
47. Zheng J, Xu C, Chu D, et al. Human leukocyte antigen G is associated with esophageal squamous cell carcinoma progression and poor prognosis. Immunol Lett 2014;161:13-9.
48. Hu J, Li L, Liu Y, et al. Overexpression of HLA-G Is positively associated with Kazakh esophageal squamous cell carcinoma in Xinjiang, China. Viral Immunol 2013;26:180-4.
49. Baccar A, Ferchichi I, Troudi W, et al. CD99 and HLA-II immunostaining in breast cancer tissue and their correlation with lymph node metastasis. Dis Markers 2013;34:363-71.
50. De Lerma Barbaro A, De Ambrosis A, Banelli B, et al. Methylation of CIITA promoter IV causes loss of HLA-II inducibility by IFN-gamma in promyelocytic cells. Int Immunol 2008;20:1457-66.
51. Jo YS, Lee JC, Li S, et al. Significance of the expression of major histocompatibility complex class II antigen, HLA-DR and -DQ, with recurrence of papillary thyroid cancer. Int J Cancer 2008;122:785-90.
52. Lv J, Cao XF, Ji L, et al. Association of Wnt1/beta-catenin with clinical pathological characteristics and prognosis of esophageal squamous cell carcinoma. Genet Test Mol Biomarkers 2010;14:363-9.
53. Shen LY, Wang H, Dong B, et al. Possible prediction of the response of esophageal squamous cell carcinoma to neoadjuvant chemotherapy based on gene expression profiling. Oncotarget 2016;7:4531-41.
54. Wang H, Shen L, Lin Y, et al. The expression and prognostic significance of Mucin 13 and Mucin 20 in esophageal squamous cell carcinoma. J Cancer Res Ther 2015;11 Suppl 1:C74-9.
55. Li Y, Zhang Q, Peng B, et al. Identification of glutathione S-transferase omega 1 (GSTO1) protein as a novel tumor-associated antigen and its autoantibody in human esophageal squamous cell carcinoma. Tumour Biol 2014;35:10871-7.
56. Xie F, Zhang D, Qian X, et al. Analysis of cancer-promoting genes related to chemotherapy resistance in esophageal squamous cell carcinoma. Ann Transl Med 2022;10:92.
57. Yang Y, Sun Z, Shi Y, et al. Integrated bioinformatics analysis to identify key genes related to the prognosis of esophageal squamous cell carcinoma. Transl Cancer Res 2021;10:1679-91.
58. Li LY, Jiang H, Xie YM, et al. Macrolide analog F806 suppresses esophageal squamous cell carcinoma (ESCC) by blocking β1 integrin activation. Oncotarget 2015;6:15940-52.
59. Luo A, Yu X, Li G, et al. Differentiation-associated genes regulated by c-Jun and decreased in the progression of esophageal squamous cell carcinoma. PLoS One 2014;9:e96610.
60. Jiang YY, Jiang Y, Li CQ, et al. TP63, SOX2, and KLF5 Establish a Core Regulatory Circuitry That Controls Epigenetic and Transcription Patterns in Esophageal Squamous Cell Carcinoma Cell Lines. Gastroenterology 2020;159:1311-1327.e19.
61. Ahrens TD, Timme S, Hoeppner J, et al. Selective inhibition of esophageal cancer cells by combination of HDAC inhibitors and Azacytidine. Epigenetics 2015;10:431-45.
62. Chen J, Guo L, Peiffer DA, et al. Genomic profiling of 766 cancer-related genes in archived esophageal normal and carcinoma tissues. Int J Cancer 2008;122:2249-54.
63. Li X, Jiang Z, Feng J, et al. 2-Acetylamino-3-[4-(2-acetylamino-2-carboxyethylsulfanylcarbonylamino) phenyl carbamoylsulfanyl] propionic acid, a glutathione reductase inhibitor, induces G2/M cell cycle arrest through generation of thiol oxidative stress in human esophageal cancer cells. Oncotarget 2017;8:61846-60.
64. Li LY, Zhang K, Jiang H, et al. Quantitative proteomics reveals the downregulation of GRB2 as a prominent node of F806-targeted cell proliferation network. J Proteomics 2015;117:145-55.
65. Lu J, Xue L, Jin M, et al. Expression profiling of metastasis-related microRNAs in early esophageal squamous cell carcinoma. Zhonghua Bing Li Xue Za Zhi 2014;43:313-7.
66. Palumbo A Jr, Da Costa NM, De Martino M, et al. UBE2C is overexpressed in ESCC tissues and its abrogation attenuates the malignant phenotype of ESCC cell lines. Oncotarget 2016;7:65876-87.
67. Wang J, Zhang Z, Che Y, et al. Rabdocoestin B exhibits antitumor activity by inducing G2/M phase arrest and apoptosis in esophageal squamous cell carcinoma. Cancer

Chemother Pharmacol 2018;81:469-81.
68. Yin X, Zhang R, Feng C, et al. Diallyl disulfide induces G2/M arrest and promotes apoptosis through the p53/p21 and MEK-ERK pathways in human esophageal squamous cell carcinoma. Oncol Rep 2014;32:1748-56.

(English Language Editor: J. Jones)

Cite this article as: He W, Yuan K, He J, Wang C, Peng L, Han Y, Chen N. Network and pathway-based analysis of genes associated with esophageal squamous cell carcinoma. Ann Transl Med 2023;11(2):102. doi: 10.21037/atm-22-6512

Supplementary

Table S1 321 genes related to ESCC

Gene abbreviations	Gene ID	Species	Gene Name
SLC9A9	285195	Homo sapiens	Solute carrier family 9 member A9 (SLC9A9)
DLC1	10395	Homo sapiens	DLC1 Rho GTPase activating protein (DLC1)
XRCC5	7520	Homo sapiens	X-ray repair cross complementing 5 (XRCC5)
S100A3	6274	Homo sapiens	S100 calcium binding protein A3 (S100A3)
CYP24A1	1591	Homo sapiens	Cytochrome P450 family 24 subfamily A member 1 (CYP24A1)
CYP3A5	1577	Homo sapiens	Cytochrome P450 family 3 subfamily A member 5 (CYP3A5)
FHIT	2272	Homo sapiens	Fragile histidine triad (FHIT)
CYP3A7	1551	Homo sapiens	cytochrome P450 family 3 subfamily A member 7 (CYP3A7)
XRCC2	7516	Homo sapiens	X-ray repair cross complementing 2 (XRCC2)
PTGS2	5743	Homo sapiens	Prostaglandin-endoperoxide synthase 2 (PTGS2)
XRCC6	2547	Homo sapiens	X-ray repair cross complementing 6 (XRCC6)
CYP2D6	1565	Homo sapiens	Cytochrome P450 family 2 subfamily D member 6 (CYP2D6)
EDN1	1906	Homo sapiens	Endothelin 1 (EDN1)
AURKA	6790	Homo sapiens	Aurora kinase A (AURKA)
XRCC1	7515	Homo sapiens	X-ray repair cross complementing 1 (XRCC1)
SLC52A3	113278	Homo sapiens	solute carrier family 52 member 3 (SLC52A3)
CLPTM1L	81037	Homo sapiens	CLPTM1 like (CLPTM1L)
DIRAS1	148252	Homo sapiens	DIRAS family GTPase 1 (DIRAS1)
CD44	960	Homo sapiens	CD44 molecule (Indian blood group) (CD44)
TRAK2	66008	Homo sapiens	Trafficking kinesin protein 2 (TRAK2)
INSIG2	51141	Homo sapiens	Insulin induced gene 2 (INSIG2)
HOTAIR	100124700	Homo sapiens	HOX transcript antisense RNA (HOTAIR)
SULT1A1	6817	Homo sapiens	Sulfotransferase family 1A member 1 (SULT1A1)
PIK3C3	5289	Homo sapiens	Phosphatidylinositol 3-kinase catalytic subunit type 3 (PIK3C3)
INTS6	26512	Homo sapiens	Integrator complex subunit 6 (INTS6)
CHRNA5	1138	Homo sapiens	Cholinergic receptor nicotinic alpha 5 subunit (CHRNA5)
PIK3CA	5290	Homo sapiens	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (PIK3CA)
CHRNA6	8973	Homo sapiens	Cholinergic receptor nicotinic alpha 6 subunit (CHRNA6)
FAS	355	Homo sapiens	Fas cell surface death receptor (FAS)
RARB	5915	Homo sapiens	Retinoic acid receptor beta (RARB)
GPT2	84706	Homo sapiens	Glutamic--pyruvic transaminase 2 (GPT2)
WWOX	51741	Homo sapiens	WW domain containing oxidoreductase (WWOX)
CHRNA3	1136	Homo sapiens	Cholinergic receptor nicotinic alpha 3 subunit (CHRNA3)
RAB27A	5873	Homo sapiens	RAB27A, member RAS oncogene family (RAB27A)
BSG	682	Homo sapiens	Basigin (Ok blood group) (BSG)
CYP1A1	1543	Homo sapiens	cytochrome P450 family 1 subfamily A member 1 (CYP1A1)
VANGL1	81839	Homo sapiens	VANGL planar cell polarity protein 1 (VANGL1)
PIK3CB	5291	Homo sapiens	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit beta (PIK3CB)
PIM1	5292	Homo sapiens	Pim-1 proto-oncogene, serine/threonine kinase (PIM1)
LIG3	3980	Homo sapiens	DNA ligase 3 (LIG3)
GSTT1	2952	Homo sapiens	Glutathione S-transferase theta 1 (GSTT1)
PTPRT	11122	Homo sapiens	Protein tyrosine phosphatase, receptor type T (PTPRT)
CYP2E1	1571	Homo sapiens	Cytochrome P450 family 2 subfamily E member 1 (CYP2E1)
VEGFA	7422	Homo sapiens	Vascular endothelial growth factor A (VEGFA)
PLA2G2A	5320	Homo sapiens	Phospholipase A2 group IIA (PLA2G2A)
TMPRSS11A	339967	Homo sapiens	Transmembrane protease, serine 11A (TMPRSS11A)
TNFRSF6B	8771	Homo sapiens	TNF receptor superfamily member 6b (TNFRSF6B)
GC	2638	Homo sapiens	GC, vitamin D binding protein (GC)
CHKA	1119	Homo sapiens	Choline kinase alpha (CHKA)
GNAI3	2773	Homo sapiens	G protein subunit alpha i3 (GNAI3)
CYP1B1	1545	Homo sapiens	Cytochrome P450 family 1 subfamily B member 1 (CYP1B1)
ERBB4	2066	Homo sapiens	erb-b2 receptor tyrosine kinase 4 (ERBB4)
HLA-DRB1	3123	Homo sapiens	Major histocompatibility complex, class II, DR beta 1 (HLA-DRB1)
ERBB2	2064	Homo sapiens	Erb-b2 receptor tyrosine kinase 2 (ERBB2)
NFKBIA	4792	Homo sapiens	NFKB inhibitor alpha (NFKBIA)
CHEK1	1111	Homo sapiens	Checkpoint kinase 1 (CHEK1)
CHEK2	11200	Homo sapiens	Checkpoint kinase 2 (CHEK2)
MYBL2	4605	Homo sapiens	MYB proto-oncogene like 2 (MYBL2)
EPHB1	2047	Homo sapiens	EPH receptor B1 (EPHB1)
MANF	7873	Homo sapiens	Mesencephalic astrocyte derived neurotrophic factor (MANF)
IL17A	3605	Homo sapiens	Interleukin 17A (IL17A)
IDH2	3418	Homo sapiens	Isocitrate dehydrogenase (NADP (+)) 2, mitochondrial (IDH2)
POLQ	10721	Homo sapiens	DNA polymerase theta (POLQ)
IL23R	149233	Homo sapiens	Interleukin 23 receptor (IL23R)
CREBBP	1387	Homo sapiens	CREB binding protein (CREBBP)
TGFBR2	7048	Homo sapiens	transforming growth factor beta receptor 2 (TGFBR2)
SMYD3	64754	Homo sapiens	SET and MYND domain containing 3 (SMYD3)
TRIO	7204	Homo sapiens	Trio Rho guanine nucleotide exchange factor (TRIO)
S100A14	57402	Homo sapiens	S100 calcium binding protein A14 (S100A14)
KCNK2	3776	Homo sapiens	Potassium two pore domain channel subfamily K member 2 (KCNK2)
ATM	472	Homo sapiens	ATM serine/threonine kinase (ATM)
NOTCH3	4854	Homo sapiens	Notch 3 (NOTCH3)
NOTCH2	4853	Homo sapiens	Notch 2 (NOTCH2)
NOTCH1	4851	Homo sapiens	Notch 1 (NOTCH1)
RGS1	5996	Homo sapiens	Regulator of G-protein signaling 1 (RGS1)
SFRP2	6423	Homo sapiens	Secreted frizzled related protein 2 (SFRP2)
NTRK2	4915	Homo sapiens	Neurotrophic receptor tyrosine kinase 2 (NTRK2)
CHRNB4	1143	Homo sapiens	Cholinergic receptor nicotinic beta 4 subunit (CHRNB4)
RHEB	6009	Homo sapiens	Ras homolog enriched in brain (RHEB)
CHRNB3	1142	Homo sapiens	Cholinergic receptor nicotinic beta 3 subunit (CHRNB3)
RAD54B	25788	Homo sapiens	RAD54 homolog B (S. cerevisiae) (RAD54B)
FUK	197258	Homo sapiens	Fucokinase (FUK)
PARP1	142	Homo sapiens	Poly (ADP-ribose) polymerase 1 (PARP1)
CHRNE	1145	Homo sapiens	cholinergic receptor nicotinic epsilon subunit (CHRNE)
KLF4	9314	Homo sapiens	Kruppel like factor 4 (KLF4)
OPRM1	4988	Homo sapiens	opioid receptor mu 1 (OPRM1)
KCNH1	3756	Homo sapiens	potassium voltage-gated channel subfamily H member 1 (KCNH1)
SRSF1	6426	Homo sapiens	serine and arginine rich splicing factor 1 (SRSF1)
HELQ	113510	Homo sapiens	helicase, POLQ-like (HELQ)
LZTS1	11178	Homo sapiens	leucine zipper tumor suppressor 1 (LZTS1)
CYP2C19	1557	Homo sapiens	cytochrome P450 family 2 subfamily C member 19 (CYP2C19)
CYP2C18	1562	Homo sapiens	cytochrome P450 family 2 subfamily C member 18 (CYP2C18)
TBX21	30009	Homo sapiens	T-box 21 (TBX21)
EZH2	2146	Homo sapiens	enhancer of zeste 2 polycomb repressive complex 2 subunit (EZH2)
TP63	8626	Homo sapiens	tumor protein p63 (TP63)
MLH1	4292	Homo sapiens	mutL homolog 1 (MLH1)
CDCP1	64866	Homo sapiens	CUB domain containing protein 1 (CDCP1)
PTEN	5728	Homo sapiens	phosphatase and tensin homolog (PTEN)
SPINK7	84651	Homo sapiens	serine peptidase inhibitor, Kazal type 7 (putative) (SPINK7)
FANCL	55120	Homo sapiens	Fanconi anemia complementation group L (FANCL)
TNFRSF11A	8792	Homo sapiens	TNF receptor superfamily member 11a (TNFRSF11A)
BAG1	573	Homo sapiens	BCL2 associated athanogene 1 (BAG1)
FANCE	2178	Homo sapiens	Fanconi anemia complementation group E (FANCE)
CDKN2B-AS1	100048912	Homo sapiens	CDKN2B antisense RNA 1 (CDKN2B-AS1)
GKAP1	80318	Homo sapiens	G kinase anchoring protein 1 (GKAP1)
COL11A1	1301	Homo sapiens	collagen type XI alpha 1 chain (COL11A1)
MYC	4609	Homo sapiens	v-myc avian myelocytomatosis viral oncogene homolog (MYC)
FGF3	2248	Homo sapiens	fibroblast growth factor 3 (FGF3)
FGF4	2249	Homo sapiens	fibroblast growth factor 4 (FGF4)
TP53	7157	Homo sapiens	tumor protein p53 (TP53)
GTF2H3	2967	Homo sapiens	general transcription factor IIH subunit 3 (GTF2H3)
MBD4	8930	Homo sapiens	methyl-CpG binding domain 4, DNA glycosylase (MBD4)
MMP13	4322	Homo sapiens	matrix metallopeptidase 13 (MMP13)
MTRR	4552	Homo sapiens	5-methyltetrahydrofolate-homocysteine methyltransferase reductase (MTRR)
CCND1	595	Homo sapiens	cyclin D1 (CCND1)
SERPINB5	5268	Homo sapiens	serpin family B member 5 (SERPINB5)
FANCD2	2177	Homo sapiens	Fanconi anemia complementation group D2 (FANCD2)
ASH1L	55870	Homo sapiens	ASH1 like histone lysine methyltransferase (ASH1L)
MTAP	4507	Homo sapiens	methylthioadenosine phosphorylase (MTAP)
SIAH1	6477	Homo sapiens	siah E3 ubiquitin protein ligase 1 (SIAH1)
PLA2G5	5322	Homo sapiens	phospholipase A2 group V (PLA2G5)
CHL1	10752	Homo sapiens	cell adhesion molecule L1 like (CHL1)
ALOX12	239	Homo sapiens	arachidonate 12-lipoxygenase, 12S type (ALOX12)
HLA-DQB1	3119	Homo sapiens	major histocompatibility complex, class II, DQ beta 1 (HLA-DQB1)
PARD3	56288	Homo sapiens	par-3 family cell polarity regulator (PARD3)
POLR2E	5434	Homo sapiens	RNA polymerase II subunit E (POLR2E)
UNG	7374	Homo sapiens	uracil DNA glycosylase (UNG)
CYP27B1	1594	Homo sapiens	cytochrome P450 family 27 subfamily B member 1 (CYP27B1)
FAT3	120114	Homo sapiens	FAT atypical cadherin 3 (FAT3)
FAT4	79633	Homo sapiens	FAT atypical cadherin 4 (FAT4)
RB1CC1	9821	Homo sapiens	RB1 inducible coiled-coil 1 (RB1CC1)
BCL2	596	Homo sapiens	BCL2, apoptosis regulator (BCL2)
FAT1	2195	Homo sapiens	FAT atypical cadherin 1 (FAT1)
FAT2	2196	Homo sapiens	FAT atypical cadherin 2 (FAT2)
ZNF750	79755	Homo sapiens	zinc finger protein 750 (ZNF750)
FEN1	2237	Homo sapiens	flap structure-specific endonuclease 1 (FEN1)
PLEC	5339	Homo sapiens	plectin (PLEC)
GALNT14	79623	Homo sapiens	polypeptide N-acetylgalactosaminyltransferase 14 (GALNT14)
IL18R1	8809	Homo sapiens	interleukin 18 receptor 1 (IL18R1)
SHMT1	6470	Homo sapiens	serine hydroxymethyltransferase 1 (SHMT1)
IL6	3569	Homo sapiens	interleukin 6 (IL6)
TBXAS1	6916	Homo sapiens	thromboxane A synthase 1 (TBXAS1)
PDS5B	23047	Homo sapiens	PDS5 cohesin associated factor B (PDS5B)
NAT1	9	Homo sapiens	N-acetyltransferase 1 (NAT1)
TRIM29	23650	Homo sapiens	tripartite motif containing 29 (TRIM29)
NAT2	10	Homo sapiens	N-acetyltransferase 2 (NAT2)
CTLA4	1493	Homo sapiens	cytotoxic T-lymphocyte associated protein 4 (CTLA4)
IGF1	3479	Homo sapiens	insulin like growth factor 1 (IGF1)
IGF2	3481	Homo sapiens	insulin like growth factor 2 (IGF2)
TP73	7161	Homo sapiens	tumor protein p73 (TP73)
SOD2	6648	Homo sapiens	superoxide dismutase 2, mitochondrial (SOD2)
CDH13	1012	Homo sapiens	cadherin 13 (CDH13)
RASSF5	83593	Homo sapiens	Ras association domain family member 5 (RASSF5)
RAB37	326624	Homo sapiens	RAB37, member RAS oncogene family (RAB37)
FBLN2	2199	Homo sapiens	fibulin 2 (FBLN2)
RASSF1	11186	Homo sapiens	Ras association domain family member 1 (RASSF1)
ABCC4	10257	Homo sapiens	ATP binding cassette subfamily C member 4 (ABCC4)
PTCH1	5727	Homo sapiens	patched 1 (PTCH1)
OGG1	4968	Homo sapiens	8-oxoguanine DNA glycosylase (OGG1)
WNT7A	7476	Homo sapiens	Wnt family member 7A (WNT7A)
SSR3	6747	Homo sapiens	signal sequence receptor subunit 3 (SSR3)
LRP5	4041	Homo sapiens	LDL receptor related protein 5 (LRP5)
SLC22A17	51310	Homo sapiens	solute carrier family 22 member 17 (SLC22A17)
FGF19	9965	Homo sapiens	fibroblast growth factor 19 (FGF19)
ENAH	55740	Homo sapiens	enabled homolog (Drosophila) (ENAH)
MALAT1	378938	Homo sapiens	metastasis associated lung adenocarcinoma transcript 1 (non-protein coding) (MALAT1)
THRB	7068	Homo sapiens	thyroid hormone receptor beta (THRB)
MMP9	4318	Homo sapiens	matrix metallopeptidase 9 (MMP9)
IL18	3606	Homo sapiens	interleukin 18 (IL18)
MMP7	4316	Homo sapiens	matrix metallopeptidase 7 (MMP7)
ADH1C	126	Homo sapiens	alcohol dehydrogenase 1C (class I), gamma polypeptide (ADH1C)
CBX4	8535	Homo sapiens	chromobox 4 (CBX4)
FGF10	2255	Homo sapiens	fibroblast growth factor 10 (FGF10)
ADH1B	125	Homo sapiens	alcohol dehydrogenase 1B (class I), beta polypeptide (ADH1B)
TLR4	7099	Homo sapiens	toll like receptor 4 (TLR4)
FGF12	2257	Homo sapiens	fibroblast growth factor 12 (FGF12)
MMP3	4314	Homo sapiens	matrix metallopeptidase 3 (MMP3)
CBX8	57332	Homo sapiens	chromobox 8 (CBX8)
MMP2	4313	Homo sapiens	matrix metallopeptidase 2 (MMP2)
MMP1	4312	Homo sapiens	matrix metallopeptidase 1 (MMP1)
TGFB1	7040	Homo sapiens	transforming growth factor beta 1 (TGFB1)
IL10	3586	Homo sapiens	interleukin 10 (IL10)
H19	283120	Homo sapiens	H19, imprinted maternally expressed transcript (non-protein coding) (H19)
GSTM1	2944	Homo sapiens	glutathione S-transferase mu 1 (GSTM1)
CDKN2A	1029	Homo sapiens	cyclin dependent kinase inhibitor 2A (CDKN2A)
CDKN2B	1030	Homo sapiens	cyclin dependent kinase inhibitor 2B (CDKN2B)
PGLYRP2	114770	Homo sapiens	peptidoglycan recognition protein 2 (PGLYRP2)
IL1B	3553	Homo sapiens	interleukin 1 beta (IL1B)
RBMS3	27303	Homo sapiens	RNA binding motif single stranded interacting protein 3 (RBMS3)
IL15RA	3601	Homo sapiens	interleukin 15 receptor subunit alpha (IL15RA)
ITCH	83737	Homo sapiens	itchy E3 ubiquitin protein ligase (ITCH)
NQO1	1728	Homo sapiens	NAD (P)H quinone dehydrogenase 1 (NQO1)
DDOST	1650	Homo sapiens	dolichyl-diphosphooligosaccharide--protein glycosyltransferase non-catalytic subunit (DDOST)
EGFR	1956	Homo sapiens	epidermal growth factor receptor (EGFR)
HLA-A	3105	Homo sapiens	major histocompatibility complex, class I, A (HLA-A)
FADD	8772	Homo sapiens	Fas associated via death domain (FADD)
HLA-C	3107	Homo sapiens	major histocompatibility complex, class I, C (HLA-C)
HLA-B	3106	Homo sapiens	major histocompatibility complex, class I, B (HLA-B)
HLA-G	3135	Homo sapiens	major histocompatibility complex, class I, G (HLA-G)
NCAM1	4684	Homo sapiens	neural cell adhesion molecule 1 (NCAM1)
PLCE1	51196	Homo sapiens	phospholipase C epsilon 1 (PLCE1)
EP300	2033	Homo sapiens	E1A binding protein p300 (EP300)
PSCA	8000	Homo sapiens	prostate stem cell antigen (PSCA)
NFE2L2	4780	Homo sapiens	nuclear factor, erythroid 2 like 2 (NFE2L2)
TNFAIP2	7127	Homo sapiens	TNF alpha induced protein 2 (TNFAIP2)
RAD17	5884	Homo sapiens	RAD17 checkpoint clamp loader component (RAD17)
FGFR2	2263	Homo sapiens	fibroblast growth factor receptor 2 (FGFR2)
FGFR1	2260	Homo sapiens	fibroblast growth factor receptor 1 (FGFR1)
CAV1	857	Homo sapiens	caveolin 1 (CAV1)
FGFR4	2264	Homo sapiens	fibroblast growth factor receptor 4 (FGFR4)
SETD1B	23067	Homo sapiens	SET domain containing 1B (SETD1B)
SOX2	6657	Homo sapiens	SRY-box 2 (SOX2)
CLU	1191	Homo sapiens	clusterin (CLU)
ORAOV1	220064	Homo sapiens	oral cancer overexpressed 1 (ORAOV1)
PTPRCAP	5790	Homo sapiens	protein tyrosine phosphatase, receptor type C associated protein (PTPRCAP)
VDR	7421	Homo sapiens	vitamin D (1,25- dihydroxyvitamin D3) receptor (VDR)
FBXW7	55294	Homo sapiens	F-box and WD repeat domain containing 7 (FBXW7)
EGF	1950	Homo sapiens	epidermal growth factor (EGF)
RUNX1	861	Homo sapiens	runt related transcription factor 1 (RUNX1)
ARL6IP5	10550	Homo sapiens	ADP ribosylation factor like GTPase 6 interacting protein 5 (ARL6IP5)
UGT2B28	54490	Homo sapiens	UDP glucuronosyltransferase family 2 member B28 (UGT2B28)
ADAM29	11086	Homo sapiens	ADAM metallopeptidase domain 29 (ADAM29)
FOXA1	3169	Homo sapiens	forkhead box A1 (FOXA1)
BRCA2	675	Homo sapiens	BRCA2, DNA repair associated (BRCA2)
CACNA2D3	55799	Homo sapiens	calcium voltage-gated channel auxiliary subunit alpha2delta 3 (CACNA2D3)
TET2	54790	Homo sapiens	tet methylcytosine dioxygenase 2 (TET2)
SHANK2	22941	Homo sapiens	SH3 and multiple ankyrin repeat domains 2 (SHANK2)
BRCA1	672	Homo sapiens	BRCA1, DNA repair associated (BRCA1)
HULC	728655	Homo sapiens	hepatocellular carcinoma up-regulated long non-coding RNA (HULC)
CDKN1A	1026	Homo sapiens	cyclin dependent kinase inhibitor 1A (CDKN1A)
ATF3	467	Homo sapiens	activating transcription factor 3 (ATF3)
CDKN1B	1027	Homo sapiens	cyclin dependent kinase inhibitor 1B (CDKN1B)
THSD1	55901	Homo sapiens	thrombospondin type 1 domain containing 1 (THSD1)
ALDH2	217	Homo sapiens	aldehyde dehydrogenase 2 family (mitochondrial) (ALDH2)
TDG	6996	Homo sapiens	thymine DNA glycosylase (TDG)
RAP1A	5906	Homo sapiens	RAP1A, member of RAS oncogene family (RAP1A)
PTPN1	5770	Homo sapiens	protein tyrosine phosphatase, non-receptor type 1 (PTPN1)
MTOR	2475	Homo sapiens	mechanistic target of rapamycin (MTOR)
FAM84B	157638	Homo sapiens	family with sequence similarity 84 member B (FAM84B)
GRB7	2886	Homo sapiens	growth factor receptor bound protein 7 (GRB7)
SST	6750	Homo sapiens	somatostatin (SST)
BMI1	648	Homo sapiens	BMI1 proto-oncogene, polycomb ring finger (BMI1)
NAMPT	10135	Homo sapiens	nicotinamide phosphoribosyltransferase (NAMPT)
KDM6A	7403	Homo sapiens	lysine demethylase 6A (KDM6A)
ALG3	10195	Homo sapiens	ALG3, alpha-1,3- mannosyltransferase (ALG3)
ANO1	55107	Homo sapiens	anoctamin 1 (ANO1)
CRP	1401	Homo sapiens	C-reactive protein (CRP)
FASLG	356	Homo sapiens	Fas ligand (FASLG)
SULT2B1	6820	Homo sapiens	sulfotransferase family 2B member 1 (SULT2B1)
BAP1	8314	Homo sapiens	BRCA1 associated protein 1 (BAP1)
PDCD1	5133	Homo sapiens	programmed cell death 1 (PDCD1)
SDC2	6383	Homo sapiens	syndecan 2 (SDC2)
AKT1	207	Homo sapiens	AKT serine/threonine kinase 1 (AKT1)
CASP3	836	Homo sapiens	caspase 3 (CASP3)
HMOX1	3162	Homo sapiens	heme oxygenase 1 (HMOX1)
PAK4	10298	Homo sapiens	p21 (RAC1) activated kinase 4 (PAK4)
CASP8	841	Homo sapiens	caspase 8 (CASP8)
ALS2CR12	130540	Homo sapiens	amyotrophic lateral sclerosis 2 chromosome region candidate 12 (ALS2CR12)
LTA	4049	Homo sapiens	lymphotoxin alpha (LTA)
KMT5A	387893	Homo sapiens	lysine methyltransferase 5A (KMT5A)
TERT	7015	Homo sapiens	telomerase reverse transcriptase (TERT)
NFX1	4799	Homo sapiens	nuclear transcription factor, X-box binding 1 (NFX1)
SHBG	6462	Homo sapiens	sex hormone binding globulin (SHBG)
IL18RAP	8807	Homo sapiens	interleukin 18 receptor accessory protein (IL18RAP)
CYP11A1	1583	Homo sapiens	cytochrome P450 family 11 subfamily A member 1 (CYP11A1)
CDK8	1024	Homo sapiens	cyclin dependent kinase 8 (CDK8)
PRKCI	5584	Homo sapiens	protein kinase C iota (PRKCI)
RB1	5925	Homo sapiens	RB transcriptional corepressor 1 (RB1)
HLTF	6596	Homo sapiens	helicase like transcription factor (HLTF)
ECT2	1894	Homo sapiens	epithelial cell transforming 2 (ECT2)
XPA	7507	Homo sapiens	XPA, DNA damage recognition and repair factor (XPA)
ALDH7A1	501	Homo sapiens	aldehyde dehydrogenase 7 family member A1 (ALDH7A1)
XPC	7508	Homo sapiens	XPC complex subunit, DNA damage recognition and repair factor (XPC)
MDM2	4193	Homo sapiens	MDM2 proto-oncogene (MDM2)
CYP2A6	1548	Homo sapiens	cytochrome P450 family 2 subfamily A member 6 (CYP2A6)
WASL	8976	Homo sapiens	Wiskott-Aldrich syndrome like (WASL)
MDM4	4194	Homo sapiens	MDM4, p53 regulator (MDM4)
GSTP1	2950	Homo sapiens	glutathione S-transferase pi 1 (GSTP1)
DCC	1630	Homo sapiens	DCC netrin 1 receptor (DCC)
KMT2D	8085	Homo sapiens	lysine methyltransferase 2D (KMT2D)
MBL2	4153	Homo sapiens	mannose binding lectin 2 (MBL2)
DNAH9	1770	Homo sapiens	dynein axonemal heavy chain 9 (DNAH9)
TNF	7124	Homo sapiens	tumor necrosis factor (TNF)
TNC	3371	Homo sapiens	tenascin C (TNC)
KMT2C	58508	Homo sapiens	lysine methyltransferase 2C (KMT2C)
MGMT	4255	Homo sapiens	O-6-methylguanine-DNA methyltransferase (MGMT)
PRKDC	5591	Homo sapiens	protein kinase, DNA-activated, catalytic polypeptide (PRKDC)
CDH1	999	Homo sapiens	cadherin 1 (CDH1)
ADH7	131	Homo sapiens	alcohol dehydrogenase 7 (class IV), mu or sigma polypeptide (ADH7)
COMT	1312	Homo sapiens	catechol-O-methyltransferase (COMT)
CRNN	49860	Homo sapiens	cornulin (CRNN)
SMUG1	23583	Homo sapiens	single-strand-selective monofunctional uracil-DNA glycosylase 1 (SMUG1)
DROSHA	29102	Homo sapiens	drosha ribonuclease III (DROSHA)
TYMS	7298	Homo sapiens	thymidylate synthetase (TYMS)
TNFRSF1A	7132	Homo sapiens	TNF receptor superfamily member 1A (TNFRSF1A)
ERCC5	2073	Homo sapiens	ERCC excision repair 5, endonuclease (ERCC5)
MTHFR	4524	Homo sapiens	methylenetetrahydrofolate reductase (MTHFR)
UGT1A8	54576	Homo sapiens	UDP glucuronosyltransferase family 1 member A8 (UGT1A8)
FAM135B	51059	Homo sapiens	family with sequence similarity 135 member B (FAM135B)
ADH4	127	Homo sapiens	alcohol dehydrogenase 4 (class II), pi polypeptide (ADH4)
PLCD1	5333	Homo sapiens	phospholipase C delta 1 (PLCD1)
SLC39A6	25800	Homo sapiens	solute carrier family 39 member 6 (SLC39A6)
ERCC4	2072	Homo sapiens	ERCC excision repair 4, endonuclease catalytic subunit (ERCC4)
GEMIN4	50628	Homo sapiens	gem nuclear organelle associated protein 4 (GEMIN4)
ERCC1	2067	Homo sapiens	ERCC excision repair 1, endonuclease non-catalytic subunit (ERCC1)
APC	324	Homo sapiens	APC, WNT signaling pathway regulator (APC)
ERCC2	2068	Homo sapiens	ERCC excision repair 2, TFIIH core complex helicase subunit (ERCC2)
COL4A1	1282	Homo sapiens	collagen type IV alpha 1 chain (COL4A1)
KAT2B	8850	Homo sapiens	lysine acetyltransferase 2B (KAT2B)
MSH2	4436	Homo sapiens	mutS homolog 2 (MSH2)
CENPF	1063	Homo sapiens	centromere protein F (CENPF)
SMC4	10051	Homo sapiens	structural maintenance of chromosomes 4 (SMC4)
AJUBA	84962	Homo sapiens	ajuba LIM protein (AJUBA)
SLC35C2	51006	Homo sapiens	solute carrier family 35 member C2 (SLC35C2)
MPG	4350	Homo sapiens	N-methylpurine DNA glycosylase (MPG)
IRF5	3663	Homo sapiens	interferon regulatory factor 5 (IRF5)
BAX	581	Homo sapiens	BCL2 associated X, apoptosis regulator (BAX)
SULF1	23213	Homo sapiens	sulfatase 1 (SULF1)
KDM4C	23081	Homo sapiens	lysine demethylase 4C (KDM4C)
UGT2B4	7363	Homo sapiens	UDP glucuronosyltransferase family 2 member B4 (UGT2B4)
IGFBP3	3486	Homo sapiens	insulin like growth factor binding protein 3 (IGFBP3)

ESCC, esophageal squamous cell carcinoma

Table S2 Significantly enriched KEGG pathways

Pathway name	Database	ID	Input number	Background number	P value	Corrected P value	Gene ID	Gene name
Platinum drug resistance	KEGG pathway	hsa01524	24	75	7.51E-08	1.80E-05	596|5290|5291|4436|2944|836|7157|207|581|8772|4193|672|2950|2952|1026|841|1029|4292|472|356|355|7507|2064|2067	BCL2|PIK3CA|PIK3CB|MSH2|GSTM1|CASP3|TP53|AKT1|BAX|FADD|MDM2|BRCA1|GSTP1|GSTT1|CDKN1A|CASP8|CDKN2A|MLH1|ATM|FASLG|FAS|XPA|ERBB2|ERCC1
Bladder cancer	KEGG pathway	hsa05219	16	41	1.35E-06	0.000162	4318|4312|1026|11186|7157|999|1029|4193|4313|5925|1956|1950|4609|595|2064|7422	MMP9|MMP1|CDKN1A|RASSF1|TP53|CDH1|CDKN2A|MDM2|MMP2|RB1|EGFR|EGF|MYC|CCND1|ERBB2|VEGFA
Chemical carcinogenesis	KEGG pathway	hsa05204	22	82	3.06E-06	0.000245	54490|6817|131|5743|2944|1545|1543|7363|1548|9|1562|125|126|127|2950|2952|1551|10|1557|54576|1571|1577	UGT2B28|SULT1A1|ADH7|PTGS2|GSTM1|CYP1B1|CYP1A1|UGT2B4|CYP2A6|NAT1|CYP2C18|ADH1B|ADH1C|ADH4|GSTP1|GSTT1|CYP3A7|NAT2|CYP2C19|UGT1A8|CYP2E1|CYP3A5
Melanoma	KEGG pathway	hsa05218	20	71	4.60E-06	0.000276	5728|1026|595|207|9965|7157|999|1029|3479|5290|4193|5925|1956|2248|2249|2257|2255|5291|2260|1950	PTEN|CDKN1A|CCND1|AKT1|FGF19|TP53|CDH1|CDKN2A|IGF1|PIK3CA|MDM2|RB1|EGFR|FGF3|FGF4|FGF12|FGF10|PIK3CB|FGFR1|EGF
Prostate cancer	KEGG pathway	hsa05215	22	89	9.35E-06	0.000359	2475|595|596|5290|5291|5728|7157|3479|5925|2033|2263|2260|207|4193|1956|1950|2950|4792|1027|1026|1387|2064	MTOR|CCND1|BCL2|PIK3CA|PIK3CB|PTEN|TP53|IGF1|RB1|EP300|FGFR2|FGFR1|AKT1|MDM2|EGFR|EGF|GSTP1|NFKBIA|CDKN1B|CDKN1A|CREBBP|ERBB2
p53 signaling pathway	KEGG pathway	hsa04115	19	69	1.04E-05	0.000359	472|5728|1026|836|841|581|7157|6477|4194|7161|1029|4193|3479|1111|11200|355|595|3486|5268	ATM|PTEN|CDKN1A|CASP3|CASP8|BAX|TP53|SIAH1|MDM4|TP73|CDKN2A|MDM2|IGF1|CHEK1|CHEK2|FAS|CCND1|IGFBP3|SERPINB5
Pathways in cancer	KEGG pathway	hsa05200	58	397	1.05E-05	0.000359	2475|4318|596|83593|2773|5290|5291|5743|595|2248|1630|5728|836|9965|7157|3479|5925|2033|2249|324|2263|2260|1030|7476|4312|4313|11186|207|581|8772|675|999|4193|5915|1956|7040|4609|2950|7048|4792|1027|1026|841|1282|861|3569|1029|4292|356|355|4436|7422|2257|2255|1387|2064|5727|1950	MTOR|MMP9|BCL2|RASSF5|GNAI3|PIK3CA|PIK3CB|PTGS2|CCND1|FGF3|DCC|PTEN|CASP3|FGF19|TP53|IGF1|RB1|EP300|FGF4|APC|FGFR2|FGFR1|CDKN2B|WNT7A|MMP1|MMP2|RASSF1|AKT1|BAX|FADD|BRCA2|CDH1|MDM2|RARB|EGFR|TGFB1|MYC|GSTP1|TGFBR2|NFKBIA|CDKN1B|CDKN1A|CASP8|COL4A1|RUNX1|IL6|CDKN2A|MLH1|FASLG|FAS|MSH2|VEGFA|FGF12|FGF10|CREBBP|ERBB2|PTCH1|EGF
Endocrine resistance	KEGG pathway	hsa01522	22	97	2.93E-05	0.00088	2475|595|596|5290|5291|7157|3479|5925|1565|4313|207|581|4318|4193|1956|1027|1026|2064|1029|4854|4853|4851	MTOR|CCND1|BCL2|PIK3CA|PIK3CB|TP53|IGF1|RB1|CYP2D6|MMP2|AKT1|BAX|MMP9|MDM2|EGFR|CDKN1B|CDKN1A|ERBB2|CDKN2A|NOTCH3|NOTCH2|NOTCH1
Colorectal cancer	KEGG pathway	hsa05210	15	62	0.000289	0.007489	1630|836|207|596|7157|581|5290|4292|4436|7040|324|4609|595|5291|7048	DCC|CASP3|AKT1|BCL2|TP53|BAX|PIK3CA|MLH1|MSH2|TGFB1|APC|MYC|CCND1|PIK3CB|TGFBR2
AGE-RAGE signaling pathway in diabetic complications	KEGG pathway	hsa04933	20	101	0.000322	0.007489	1027|3569|3553|207|1282|51196|5292|5290|5291|581|595|5333|4313|7040|1906|836|7124|596|7048|7422	CDKN1B|IL6|IL1B|AKT1|COL4A1|PLCE1|PIM1|PIK3CA|PIK3CB|BAX|CCND1|PLCD1|MMP2|TGFB1|EDN1|CASP3|TNF|BCL2|TGFBR2|VEGFA
Non-small cell lung cancer	KEGG pathway	hsa05223	14	56	0.000343	0.007489	11186|207|2272|7157|83593|1029|5290|5291|5925|1956|1950|5915|595|2064	RASSF1|AKT1|FHIT|TP53|RASSF5|CDKN2A|PIK3CA|PIK3CB|RB1|EGFR|EGF|RARB|CCND1|ERBB2
Hepatitis B	KEGG pathway	hsa05161	25	146	0.000443	0.008802	595|596|5290|5291|5728|836|7157|7099|5925|2033|207|581|4318|7040|4609|7124|4792|1027|1026|841|3569|8772|356|355|1387	CCND1|BCL2|PIK3CA|PIK3CB|PTEN|CASP3|TP53|TLR4|RB1|EP300|AKT1|BAX|MMP9|TGFB1|MYC|TNF|NFKBIA|CDKN1B|CDKN1A|CASP8|IL6|FADD|FASLG|FAS|CREBBP
Metabolism of xenobiotics by cytochrome P450	KEGG pathway	hsa00980	16	73	0.000477	0.008802	1545|1543|54490|2950|7363|131|1548|54576|125|126|127|1571|2944|1577|1565|2952	CYP1B1|CYP1A1|UGT2B28|GSTP1|UGT2B4|ADH7|CYP2A6|UGT1A8|ADH1B|ADH1C|ADH4|CYP2E1|GSTM1|CYP3A5|CYP2D6|GSTT1
Endometrial cancer	KEGG pathway	hsa05213	13	52	0.000553	0.009472	5728|595|207|7157|999|5290|4292|1956|1950|324|4609|5291|2064	PTEN|CCND1|AKT1|TP53|CDH1|PIK3CA|MLH1|EGFR|EGF|APC|MYC|PIK3CB|ERBB2
Proteoglycans in cancer	KEGG pathway	hsa05205	31	205	0.000667	0.010306	2475|595|5290|5291|6383|857|836|7157|51196|3479|7099|4318|5727|2260|7476|4313|207|3481|4193|1956|7040|4609|7124|1026|960|356|355|7422|29102|2064	MTOR|CCND1|PIK3CA|PIK3CB|SDC2|CAV1|CASP3|TP53|PLCE1|IGF1|TLR4|MMP9|PTCH1|FGFR1|WNT7A|MMP2|AKT1|IGF2|MDM2|EGFR|TGFB1|MYC|TNF|CDKN1A|CD44|FASLG|FAS|VEGFA|DROSHA|ERBB2|ERBB4
Base excision repair	KEGG pathway	hsa03410	10	33	0.000687	0.010306	142|4350|6996|2237|7374|7515|23583|4968|8930|3980	PARP1|MPG|TDG|FEN1|UNG|XRCC1|SMUG1|OGG1|MBD4|LIG3
Drug metabolism - cytochrome P450	KEGG pathway	hsa00982	15	69	0.000761	0.010742	1557|54490|2950|7363|131|1548|54576|125|126|127|1571|2944|1577|1565|2952	CYP2C19|UGT2B28|GSTP1|UGT2B4|ADH7|CYP2A6|UGT1A8|ADH1B|ADH1C|ADH4|CYP2E1|GSTM1|CYP3A5|CYP2D6|GSTT1
HTLV-I infection	KEGG pathway	hsa05166	36	259	0.00103	0.013222	595|5290|5291|11200|7015|8850|7132|7157|3105|3106|3107|4049|1111|5925|2033|324|1030|7476|3601|207|581|467|3135|7040|4609|7124|4605|7048|4792|1026|3119|3569|1029|472|3123|1387	CCND1|PIK3CA|PIK3CB|CHEK2|TERT|KAT2B|TNFRSF1A|TP53|HLA-A|HLA-B|HLA-C|LTA|CHEK1|RB1|EP300|APC|CDKN2B|WNT7A|IL15RA|AKT1|BAX|ATF3|HLA-G|TGFB1|MYC|TNF|MYBL2|TGFBR2|NFKBIA|CDKN1A|HLA-DQB1|IL6|CDKN2A|ATM|HLA-DRB1|CREBBP
Graft-versus-host disease	KEGG pathway	hsa05332	11	42	0.001047	0.013222	3107|3553|3119|3123|3106|3569|3135|356|355|3105|7124	HLA-C|IL1B|HLA-DQB1|HLA-DRB1|HLA-B|IL6|HLA-G|FASLG|FAS|HLA-A|TNF
Glioma	KEGG pathway	hsa05214	14	65	0.001216	0.014276	5728|1026|595|207|7157|1029|2475|5290|3479|5925|4193|1956|1950|5291	PTEN|CDKN1A|CCND1|AKT1|TP53|CDKN2A|MTOR|PIK3CA|IGF1|RB1|MDM2|EGFR|EGF|PIK3CB
Chronic myeloid leukemia	KEGG pathway	hsa05220	15	73	0.001249	0.014276	1027|1026|595|207|5925|861|1029|4193|5290|5291|7157|7040|4609|7048|4792	CDKN1B|CDKN1A|CCND1|AKT1|RB1|RUNX1|CDKN2A|MDM2|PIK3CA|PIK3CB|TP53|TGFB1|MYC|TGFBR2|NFKBIA
Pancreatic cancer	KEGG pathway	hsa05212	14	66	0.001379	0.014392	595|207|7157|675|1029|7422|5290|5291|5925|1956|7040|2064|7048|1950	CCND1|AKT1|TP53|BRCA2|CDKN2A|VEGFA|PIK3CA|PIK3CB|RB1|EGFR|TGFB1|ERBB2|TGFBR2|EGF
Inflammatory bowel disease (IBD)	KEGG pathway	hsa05321	14	66	0.001379	0.014392	8809|3553|3605|3586|3123|3569|149233|30009|8807|7040|3119|7099|7124|3606	IL18R1|IL1B|IL17A|IL10|HLA-DRB1|IL6|IL23R|TBX21|IL18RAP|TGFB1|HLA-DQB1|TLR4|TNF|IL18
Type I diabetes mellitus	KEGG pathway	hsa04940	11	44	0.001442	0.014423	3553|3119|3123|3106|3107|4049|3135|356|355|3105|7124	IL1B|HLA-DQB1|HLA-DRB1|HLA-B|HLA-C|LTA|HLA-G|FASLG|FAS|HLA-A|TNF
TNF signaling pathway	KEGG pathway	hsa04668	19	110	0.001867	0.017926	4318|836|4314|841|207|8772|3569|4049|5290|5291|83737|355|5743|1906|3553|7124|7132|8809|4792	MMP9|CASP3|MMP3|CASP8|AKT1|FADD|IL6|LTA|PIK3CA|PIK3CB|ITCH|FAS|PTGS2|EDN1|IL1B|TNF|TNFRSF1A|IL18R1|NFKBIA
Allograft rejection	KEGG pathway	hsa05330	10	39	0.001998	0.01844	3586|3119|3123|3106|3107|3135|356|355|3105|7124	IL10|HLA-DQB1|HLA-DRB1|HLA-B|HLA-C|HLA-G|FASLG|FAS|HLA-A|TNF
Small cell lung cancer	KEGG pathway	hsa05222	16	86	0.00215	0.018586	5728|595|207|1282|7157|596|1027|5290|2272|5925|5743|4609|5915|1030|5291|4792	PTEN|CCND1|AKT1|COL4A1|TP53|BCL2|CDKN1B|PIK3CA|FHIT|RB1|PTGS2|MYC|RARB|CDKN2B|PIK3CB|NFKBIA
HIF-1 signaling pathway	KEGG pathway	hsa04066	18	103	0.002168	0.018586	2475|1026|207|596|2064|3569|3479|5290|5291|7099|1956|2033|1950|1906|1387|3162|1027|7422	MTOR|CDKN1A|AKT1|BCL2|ERBB2|IL6|IGF1|PIK3CA|PIK3CB|TLR4|EGFR|EP300|EGF|EDN1|CREBBP|HMOX1|CDKN1B|VEGFA
MicroRNAs in cancer	KEGG pathway	hsa05206	38	299	0.003152	0.025261	2475|595|10298|596|8626|5292|5290|5743|3162|5728|1545|836|7157|4854|2033|324|11186|4318|4194|4193|672|1956|1591|4609|1027|1026|960|4853|1029|3371|5268|472|2146|7422|648|1387|2064|4851	MTOR|CCND1|PAK4|BCL2|TP63|PIM1|PIK3CA|PTGS2|HMOX1|PTEN|CYP1B1|CASP3|TP53|NOTCH3|EP300|APC|RASSF1|MMP9|MDM4|MDM2|BRCA1|EGFR|CYP24A1|MYC|CDKN1B|CDKN1A|CD44|NOTCH2|CDKN2A|TNC|SERPINB5|ATM|EZH2|VEGFA|BMI1|CREBBP|ERBB2|NOTCH1
FoxO signaling pathway	KEGG pathway	hsa04068	21	134	0.003158	0.025261	5728|1026|6648|3586|207|3569|1027|4193|5290|3479|472|356|1956|2033|7040|595|1030|1387|7048|5291|1950	PTEN|CDKN1A|SOD2|IL10|AKT1|IL6|CDKN1B|MDM2|PIK3CA|IGF1|ATM|FASLG|EGFR|EP300|TGFB1|CCND1|CDKN2B|CREBBP|TGFBR2|PIK3CB|EGF
Chagas disease (American trypanosomiasis)	KEGG pathway	hsa05142	17	104	0.005151	0.039881	7132|3553|207|3586|356|8772|2773|3569|5290|5291|841|355|7040|7124|7099|7048|4792	TNFRSF1A|IL1B|AKT1|IL10|FASLG|FADD|GNAI3|IL6|PIK3CA|PIK3CB|CASP8|FAS|TGFB1|TNF|TLR4|TGFBR2|NFKBIA

KEGG, Kyoto Encyclopedia of Genes and Genomes.

https://dx.doi.org/10.21037/atm-22-6512