

Bioinformatics Analysis of lncRNA-mRNA Interaction Network in Different Clinical Stages of Esophageal Squamous Cell Carcinoma

Seyed Navid Goftari¹, Ahmad Reza Bahrami^{1,2}, Maryam M. Matin^{1,3*} 

¹ Department of Biology, Faculty of science, Ferdowsi University of Mashhad, Mashhad, Iran

² Industrial Biotechnology Research Group, Institute of Biotechnology, Ferdowsi University of Mashhad, Mashhad, Iran

³ Novel Diagnostics and Therapeutics Research Group, Institute of Biotechnology, Ferdowsi University of Mashhad, Mashhad, Iran

Received 7 September 2023

Accepted 16 October 2023

Abstract

Esophageal cancer is one of the most aggressive gastrointestinal malignancies, and esophageal squamous cell carcinoma (ESCC) is the most prevalent esophagus neoplastic disease with high mortality rates in some Asian countries. Nonetheless, the etiology of ESCC continues to be vaguely comprehended, and the role of long noncoding RNAs (lncRNAs) in different clinical stages of this type of malignancy remains to be clarified. Here, we aimed to investigate the crucial genes corresponding to various clinical stages of ESCC, determine the hub lncRNAs in these stages, and predict patients' overall survival time. In the current study, The Cancer Genome Atlas (TCGA) RNA-seq public data were analyzed in order to discover novel biomarkers or therapeutic targets implicated in the progression of ESCC. Stage-related genes were analyzed, the protein-protein interaction network for any stage was constructed and the top 5 genes with the most Maximal Clique Centrality score in each network were selected as the hub mRNAs. lncRNAs interacting with each stage hub mRNA were also determined as stage-related hub lncRNAs. Gene set enrichment analysis on stage-associated modules was also carried out. Finally, Cox regression analysis was performed to assess the prognostic significance of identified hub lncRNAs in the survival of patients with ESCC. Finally, hub mRNAs and hub lncRNAs associated with ESCC progression were identified, which may have implications as biomarkers and targets for therapeutic interventions. Six lncRNAs, including AC013391.2, AC104088.1, AC026341.3, AL139023.1, AL583808.1, and LINC01707 were also identified to be significantly correlated with ESCC patients' overall survival time, which could be potential predictors for the survival rate of patients, however, more research is required in order to confirm the results experimentally.

Keywords: ESCC clinical stages, TCGA, WGCNA, hub mRNA, hub lncRNA, prognostic factor

Introduction

Esophageal cancer (EC) accounts for the sixth most prevalent malignancy in men and women worldwide. Poor prognosis and high mortality rates, make it one of the leading causes of cancer-related deaths (Siegel et al., 2023). Moreover, the incidence of esophageal cancer is on the rise (DiSiena et al., 2021). Esophageal squamous cell carcinoma (ESCC) is one of the most prevailing types of EC, comprising 55-65% of cases (Bray et al., 2018). Treatment options range from surgical resection to systemic chemotherapies (Yu, 2016). Despite improvements in ESCC treatments, including immunotherapy or neoadjuvant chemotherapy, the prognosis of ESCC patients remains poor, with five-year overall survival (OS) rates of less than 20% (Sheikh et al., 2023). The development of ESCC is a multistage procedure, comprising many

gene networks and variations in signaling pathways, while their exact contribution is not fully clarified yet (Kadian et al., 2022). These genetic imbalances involve protein-coding genes and noncoding RNAs (Feng et al., 2019). Many key signaling pathways such as RAF/MAPK/ERK, PI3K/AKT/mTOR, WNT- β -catenin and JAK/STAT are known to be involved in the pathogenesis of ESCC (Aravalli et al., 2013; Shin and Chung, 2013). Studies in the last decade suggested that some long noncoding RNAs (lncRNAs) are also implicated in ESCC development, as their dysregulation is associated with tumorigenesis, metastasis or prognosis outcome (Gao et al., 2018; Lin et al., 2017; Zheng et al., 2022). Notably, the exact mechanisms of ESCC development, metastasis, and recurrence remain to be fully elucidated. Thus, there is a pressing need to gain a more comprehensive understanding of the biological mechanisms involved in the pathogenesis

* Corresponding author's e-mail address: matin@um.ac.ir

and progression of ESCC. This understanding is crucial for the identification of novel diagnostic biomarkers and also development of effective therapeutic strategies. Many studies have demonstrated that lncRNAs are extensively expressed in various human tissues and play a significant role in numerous biological processes, such as cell fate determination, differentiation of stem cells, X chromosome inactivation, cancer initiation and progression, metastasis, and drug resistance (Jia et al., 2021; Li et al., 2020; Liu et al., 2020; Luo et al., 2022; Ren et al., 2022; Wu et al., 2022). So, dysregulation of lncRNAs expression is involved in development of several diseases, including cancer (Shi et al., 2013).

Over the last decade, significant progress has been made in high-throughput RNA sequencing techniques and bioinformatics methodologies. These advancements have facilitated the comprehensive evaluation of the entire transcriptome, enabling the analysis of numerous cancer tissue samples, RNA sequencing, and microarray datasets. These investigations exhibited altered expression patterns in a substantial number of lncRNAs in various human cancers (McCabe and Rasmussen, 2021; Tan et al., 2021; Yang et al., 2022). Furthermore, some of these studies have led to the discovery of many new lncRNAs influencing human disease phenotypes (Cui et al., 2017; Li et al., 2019; Sun and Kraus, 2015).

Systems biology approaches can simplify the determination of the mechanisms underlying ESCC transformation as well as the discovery of promising novel biomarkers. Systems biology techniques are of great value due to their comprehensive, and accurate properties, which integrate various biological data (Pinu et al., 2019). As a new research methodology, these approaches have been applied in the diagnosis and discovery of novel biomarkers for different diseases, including different human cancers (Cherubini et al., 2018; Mambetsariev et al., 2023). These methods can be used to predict the treatment outcomes and prognosis in ESCC patients, with potential clinical applications for this kind of malignancy. Despite the characterization of functions and mechanisms of certain lncRNAs in ESCC, our understanding of lncRNAs in this particular malignancy remains significantly restricted. The expression pattern and significance of numerous lncRNAs are yet to be fully elucidated. Consequently, the primary aim of this investigation was to identify the association of lncRNAs with different ESCC stages to clarify the mechanisms underlying ESCC progression, and to candidate

potential biomarkers or therapeutic targets for this malignancy. Hence, in order to ascertain lncRNAs associated with various ESCC stages, an analysis was conducted on RNA-seq profiles from different stages and their adjacent non-tumor samples, utilizing publicly available RNA sequencing data. The findings of this study unveil the expression patterns of lncRNAs in ESCC tissues, thereby presenting potential novel candidates for ESCC treatment and prognosis.

Materials and Methods

Data collection and processing

High-throughput gene expression data, including protein-coding and long noncoding genes, related to ESCC tissues and normal esophagus samples were extracted from The Cancer Genome Atlas (TCGA) data portal (<https://tcga-data.ncbi.nih.gov/tcga>). TCGA database is a cancer research initiative that endeavors to comprehend the intricacies of carcinogenesis and the progression of cancer cells. Its primary objective is identifying novel biomarkers and treatment protocols by collecting diverse cancer-related omics data. TCGA platform is accessible to the public and can be utilized without copyright restrictions (Tomczak, Czerwińska et al., 2015). The RNA-seq data in this platform have been generated by the Illumina HiSeq sequencing machine. In this study, unstranded HTSeq raw counts RNA-seq data, including 87 ESCC samples, specimens, consisting of 5 stage 1 51 stage 2 specimens, 27 stage 3 specimens, and 4 stage 4 specimens, and also three adjacent non-cancerous samples, were extracted from ESCA project in Genomic Data Commons (GDC) data portal (<https://portal.gdc.cancer.gov/>).

Identification of the aberrantly expressed lncRNAs and protein-coding genes in different ESCC stages

To compare the expression profiles of ESCC tissues with normal sample ones and identify differentially expressed genes (DEGs), the R software (version 4.2.3) was utilized. The DESeq2 Bioconductor package (Love et al., 2014) was employed for the differential expression analysis of individual lncRNAs and protein-coding genes. DEGs were determined based on the threshold values of $|\log_2 \text{fold change}| > 1$, $P\text{-value} < 0.05$, and false discovery rate (FDR) < 0.05 (Wang et al., 2021a). The volcano plots corresponding to each ESCC stage were generated using the “EnhancedVolcano” package of R software (Blighe et al., 2019).

Key coexpression WGCNA modules

The method of coexpression network generation, was used to screen candidate biomarkers and key players corresponding to each ESCC stage, based on network calculation using the R package WGCNA. WGCNA can be used to partition genes into several modules, containing a certain number of genes that are closely correlated in their expression patterns, which could be associated with a selected clinical phenotype (Langfelder and Horvath, 2008). In this study, the gene expression data matrix of ESCC samples, obtained by DESeq2, was utilized, and DEGs with P -values < 0.05 were filtered and selected as WGCNA input to generate gene coexpression networks using the “unsigned” method. This method can consider both positive and negative correlations into consideration, and is helpful for investigating noncoding RNAs, such as miRNAs and lncRNAs, along with protein-coding genes (Langfelder and Horvath, 2008). To further identify the modules associated with different ESCC stages, the relationship between modules eigengenes and each ESCC stage was calculated. Therefore, modules with high correlation coefficients related to different stages were identified and selected for subsequent analyses.

Gene set enrichment analysis

The Database for Annotation Visualization and Integrated Discovery (DAVID) (Huang et al., 2007) is an online resource that offers a comprehensive range of functional annotation tools to interpret the biological significance of genes. In order to identify the pathways that are most associated with each stage of ESCC, various module genes correlating to different stages of ESCC were submitted to DAVID (version 2021), for gene ontology (GO) biological process analysis with default parameters. The results were then visualized in a lollipop chart using “ggplot2” in R software (Wickham et al., 2016). $FDR < 0.05$ was considered as statistically significant.

Construction of protein-protein interactions

The STRING database (<https://string-db.org/>) is an online tool for exploring the interacting genes. Its primary objective is the construction and visualization of protein-protein interaction (PPI) networks belonging to various genes, which is achieved based on known and predicted interactions among proteins (Szklarczyk et al., 2015). Using the STRING tool (version 11.5), the PPI of each ESCC stage-correlating genes with a confidence score > 0.4 , was constructed, and visualized by Cytoscape <http://jcmr.um.ac.ir>

software (version 3.9.2) (Shannon et al., 2003). In addition, using the CytoHubba plug-in (version 0.1) (Chin et al., 2014) of Cytoscape software, the Maximal Clique Centrality (MCC) score for each gene was calculated, and the top 5 genes with the highest MCC score and score cut-off ≥ 2 in each network, were selected as the hub mRNAs for subsequent analyses.

Construction of lncRNA-mRNA interaction network

lncRNA-mRNA interaction between each module lncRNAs and identified hub genes was predicted using RIBlast, which is a fast and accurate RNA-RNA interaction prediction tool, using default parameters as described previously (Fukunaga and Hamada, 2017). Hub lncRNAs were selected as those that could interact with the hub genes in the same module and visualized in Sankey diagram using “ggplot2” in R software.

Investigating the prognostic value of hub lncRNAs

The prognostic significance of selected hub lncRNAs in terms of OS time was assessed using the Kaplan-Meier (KM) plotter tool developed by Lánczky and Györffy (Lánczky and Györffy, 2021). Expression levels of a specific lncRNA were compared to the median expression level of that lncRNA in all patients with ESCC. Based on this comparison, patients were divided into two groups: high-expression and low-expression. KM survival curves were then generated to visualize the survival outcomes of these two groups. A cut-off criterion of P -values less than 0.05 was applied to determine the statistical significance of the results.

Results

Differential expression analysis of lncRNAs and mRNAs

The expression level of each gene was calculated using DESeq2. Through the application of calculating criteria, we were able to identify aberrantly expressed lncRNAs and mRNAs in various stages of ESCC. The findings, presented in Table 1, suggest that these genes may have significant implications for the progression of ESCC. Figure 1 represents differentially expressed genes in each ESCC stage compared to esophagus normal samples.

Table 1. Number of differentially expressed genes corresponding to each stage of ESCC.

Stage	lncRNA		mRNA	
	Up-regulated (n)	Down-regulated (n)	Up-regulated (n)	Down-regulated (n)
Stage 1	20	12	87	59
Stage 2	588	144	987	944
Stage 3	615	200	952	1102
Stage 4	300	123	524	498

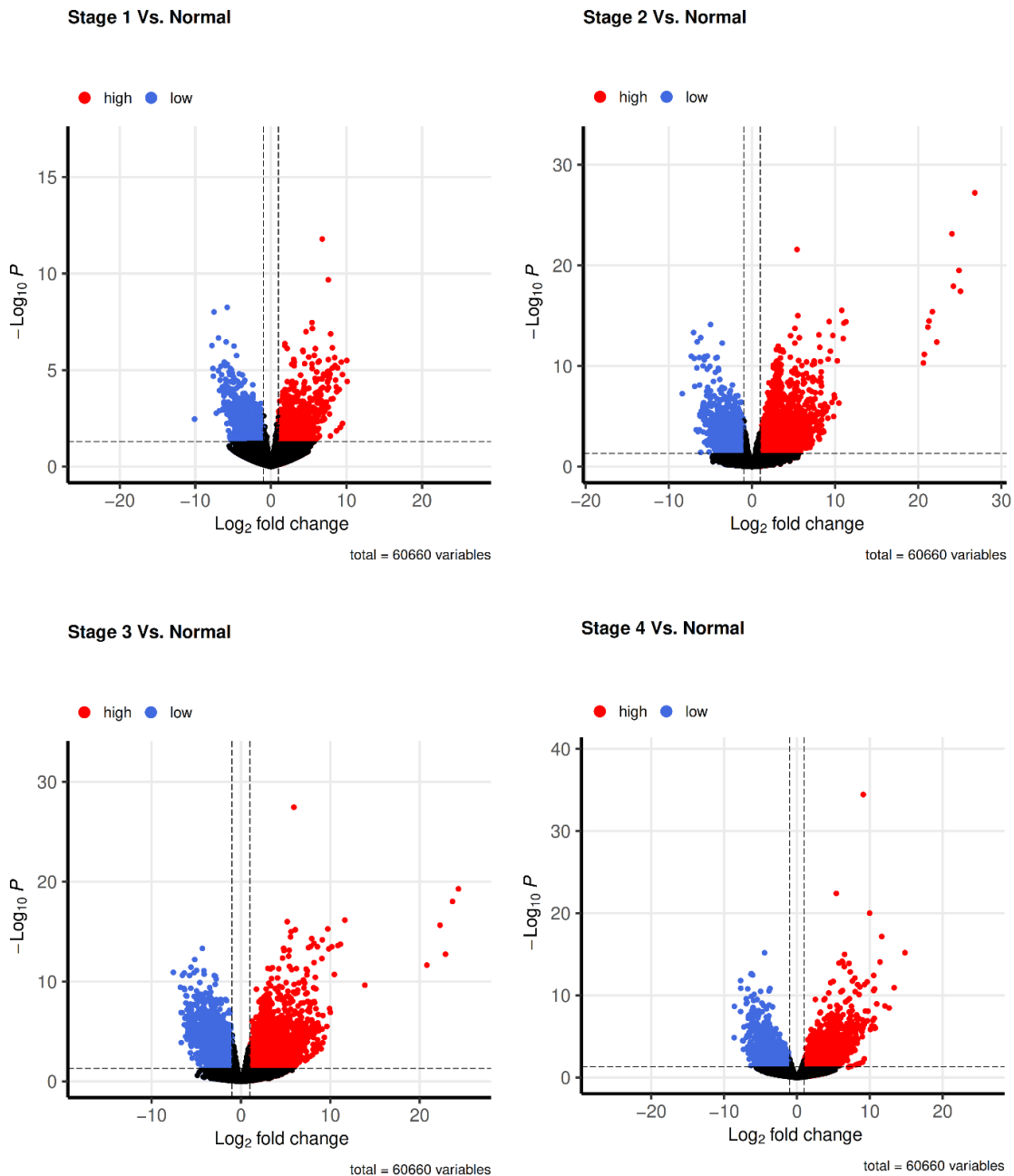


Figure 1. Volcano plots of differentially expressed lncRNAs in each ESCC stage. The horizontal line at P -value = 0.05

Constructing the coexpression modules of lncRNAs and mRNAs

After filtering for differential expression, 2794 genes with $FDR < 0.05$ were identified, consisting of 721 lncRNAs and 2073 mRNAs. These genes were then utilized as input data for establishing coexpression modules using WGCNA. The power value, which plays a crucial role in determining the average connectivity degree and independence of each coexpression module, was set to $\beta = 3$ to ensure high-scale independence and low mean connectivity (Figure 2A). The minimum module size was set as 30 with a height cut of 0.2, the gene adjacency was analyzed and, correlated modules were merged, and eventually, 13 modules were obtained in total. The dendrogram of all selected genes was clustered with the adjacency matrix, resulting in the identification of co-expression modules, as visually depicted in Figure 2B. The correlations of each module eigengene with ESCC stages were calculated using Pearson correlation analysis. We found that the brown module ($R=0.29$) was significantly associated with stage 1, the blue module ($R=0.25$) with stage 2, the yellow module ($R=0.21$) with stage 3, and the magenta module ($R=0.35$) with stage 4 phenotype

(Figure 3), and these modules were selected for subsequent analysis (Table 2).

Gene set enrichment analysis

In order to investigate the biological processes in which genes corresponding to each stage act, the protein-coding genes of selected modules were analyzed. The main aberrantly regulated biological processes related to ESCC stage 1 correspond to the retinoic acid biosynthesis process and cellular response to UV-A. ESCC stage 2 correlated genes mainly participate in spindle organization and mitotic chromosome segregation processes, which reflects the increased cell proliferation in this stage. ESCC stage 3 correlated genes mainly participate in ion transmembrane transportation, and stage 4 genes were mostly correlated to cytokines signaling, cardiomyopathy, and also cell adhesion molecules, which can be correlated to cancer metastasis, which is the main characteristic of this stage (Figure 4).

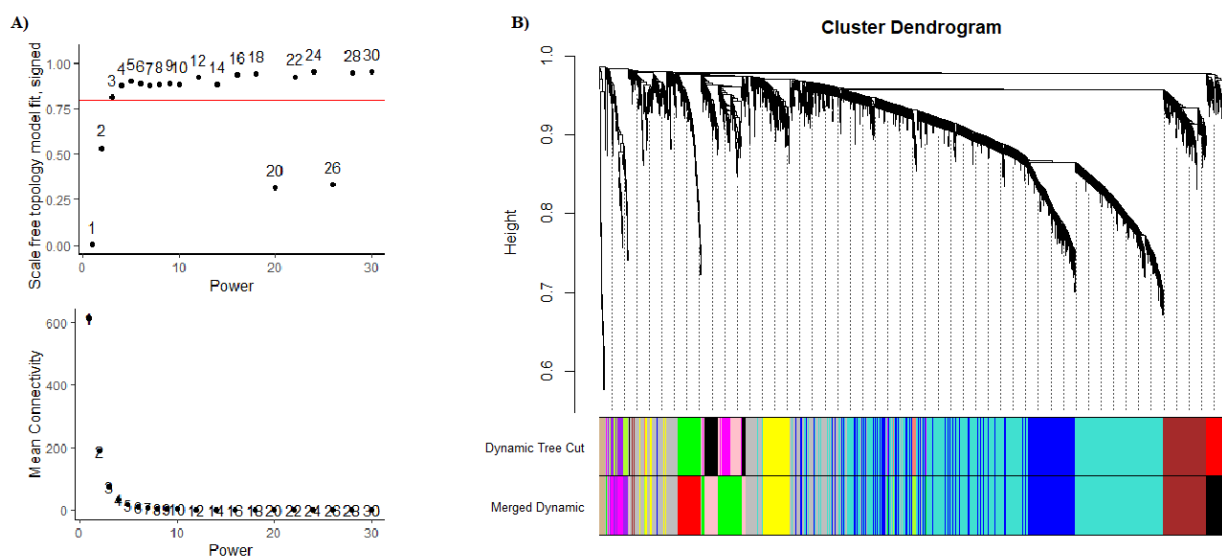


Figure 2. WGCNA analysis of lncRNAs and mRNAs. A) The network topology scale-free fitting index, obtained through the soft-threshold power analysis method, is provided as an indication. B) The identified co-expressed genes are visualized through a hierarchical clustering dendrogram. Each row is color-coded to represent a module consisting of highly connected genes.

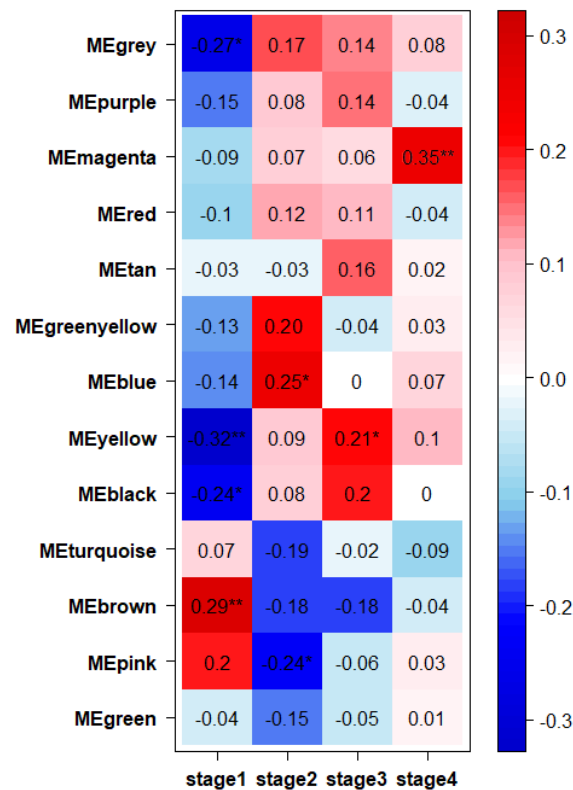


Figure 3. Heatmap of module-phenotype relationships depicting correlations between module eigengenes and ESCC stages using Pearson correlation analysis. Numbers in the table correspond to the correlation, and asterisks represent the significance of correlations (*: $P < 0.05$, **: $P < 0.01$). The degree of correlation is illustrated with the color legend.

Table 2. Most significant modules correlated to each ESCC stage and their specifications.

Module color	genes No.	lncRNAs No.	mRNAs No.	Correlated stage	Correlation	<i>P</i> -value
Brown	93	25	68	Stage 1	0.29	0.006
Blue	457	54	403	Stage 2	0.25	0.013
Yellow	133	97	36	Stage 3	0.21	0.020
Magenta	62	33	29	Stage 4	0.35	0.002

Construction of PPI networks and identification of hub mRNAs

The CytoHubba plug-in in Cytoscape was used to analyze the STRING data and construct the protein-protein interaction network formed by selected modules mRNAs. Using the MCC method, the hub mRNAs in each network corresponding to various ESCC stages were identified as RPS4Y1, KDM5D, USP9Y, DDX3Y and ZFY related to ESCC stage 1, BUB1B, BIRC5, MELK, DLGAP5 and NUF2 related to stage 2, GABRA3, GABRR1, GABRQ, HTR2C and BOLL related to stage 3 and MAGEA12, MAGEA3, CTAG2, CSAG2 and CT45A1 in the case of ESCC stage 4 (Figure 5).

These mRNAs had the least betweenness centrality ranks, suggesting they might be potential critical regulatory factors in their corresponding ESCC stages.

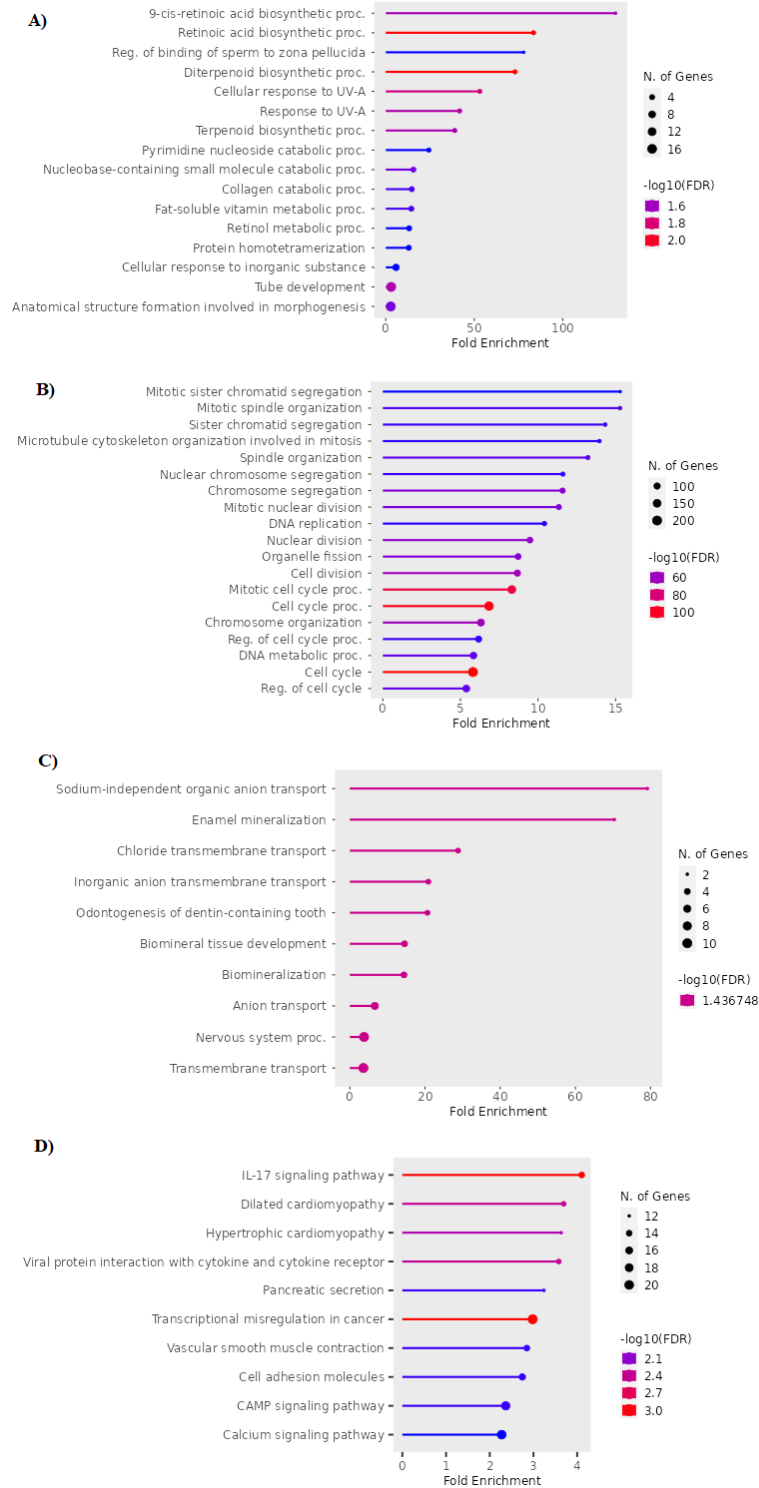


Figure 4. Gene ontology analysis of protein-coding genes related to each ESCC stage. Lollipop plots show top GO terms for biological processes of selected module genes corresponding to different ESCC stages: A) ESCC stage 1, B) ESCC stage 2, C) ESCC stage 3, and D) ESCC stage 4.

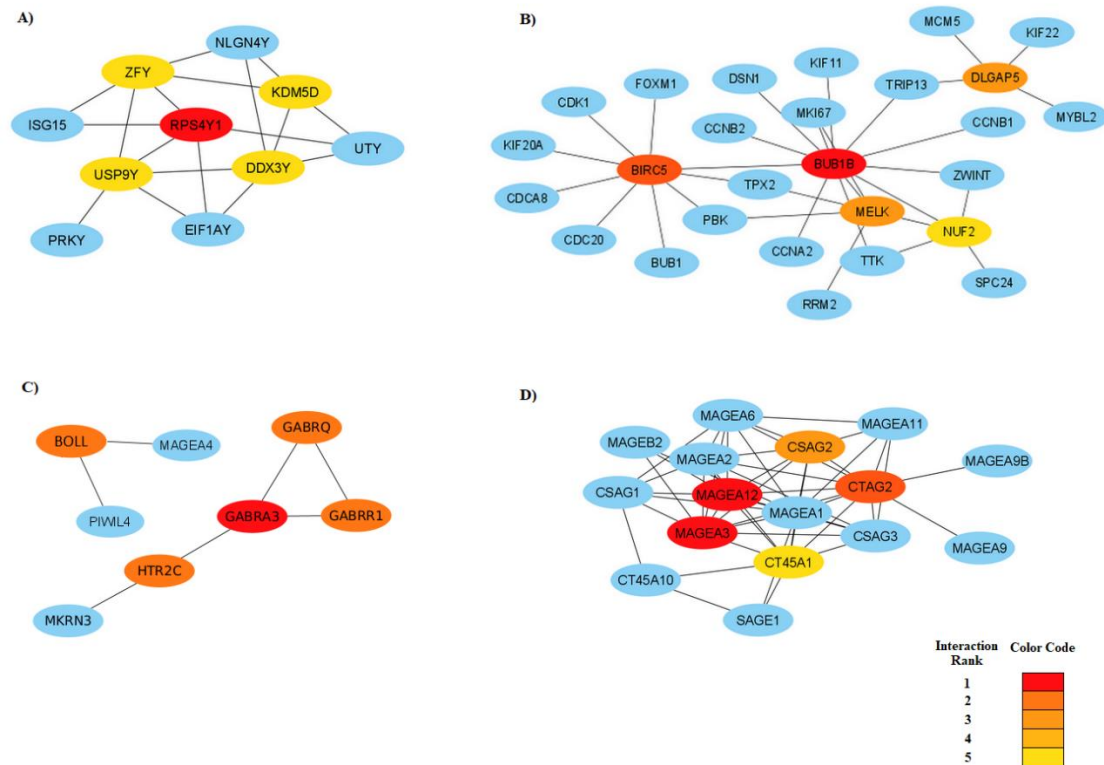


Figure 5. PPI subnetworks containing hub mRNAs corresponding to A) ESCC stage 1, B) ESCC stage 2, C) ESCC stage 3, and D) ESCC stage 4. Sky blue ellipses indicate first-stage nodes interacting with hub mRNAs. Black lines indicate the interactions between genes.

lncRNA-mRNA interaction network and identification of hub lncRNAs

Based on the identification of hub mRNAs, an investigation was conducted on the lncRNA-mRNA interaction network to identify hub lncRNAs. Consequently, some of the hub lncRNAs corresponding to each ESCC stage, were identified including TTTY14, LINC00486, LINC02154 and HOXC-AS1 related to stage 1, AC104088.1, AL161891.1, BACE1-AS, DUXAP8, LINC00342, LINC02457, MIR924HG, AC022405.1, AC093827.4, AC013391.2, TPRG1-AS1, AC021016.1 and AC092718.2 related to stage 2, AL355834.2, CASC20, LINC01707, LINC02196, PCAT1, AC007128.2, AC009264.1, AC010275.1, AL049828.2, AL139023.1, AL583808.1, CASC9, G2E3-AS1, LINC00491, LINC01807, LINC02428, AC026341.3 and LINC01981 related to stage 3, and LINC00540, LINC01193, LINC00470, TRPM2-AS, and AC109439.1 related to stage 4 (Figure 6).

Prognosis value of candidate hub lncRNAs

To identify hub lncRNAs with potential prognostic characteristics, univariate Cox regression analysis

was performed for identified hub lncRNAs corresponding to each ESCC stage. The results indicate that six lncRNAs comprising two stage 2 identified hub lncRNAs, including AC013391.2, AC104088.1, and four stage 3 identified hub lncRNAs, including AC026341.3, AL139023.1, AL583808.1 and LINC01707 were significantly associated with overall survival of ESCC patients (Figure 7). These lncRNAs showed HR values of more than 1, indicating that they might be considered potential prognostic factors for this type of malignancy. Cox analysis of the rest of the hub lncRNAs indicated *P*-values higher than 0.05, and thus, these lncRNAs were not significantly associated with patients' survival times.

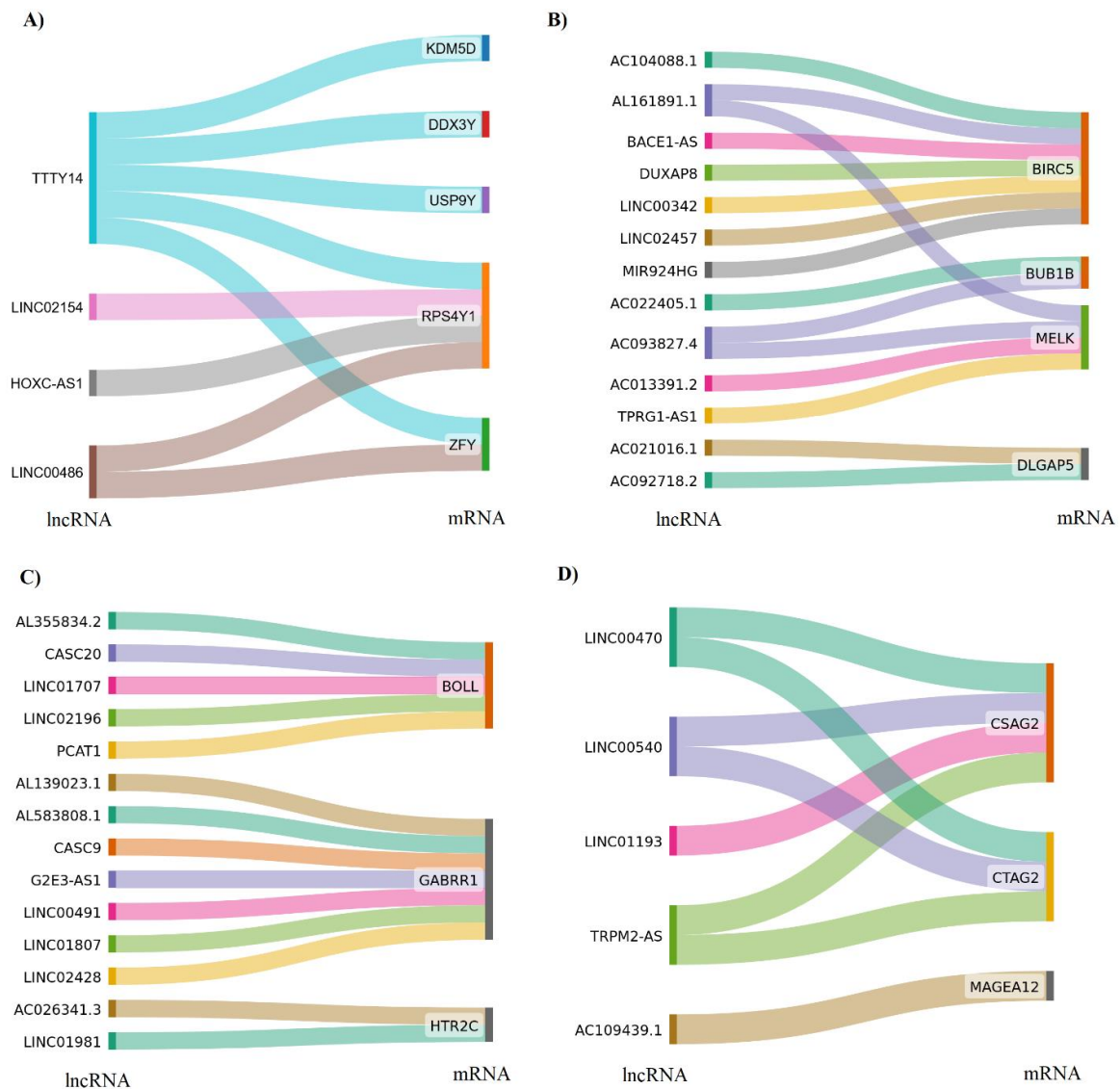


Figure 6. Identification of hub lncRNAs based on hub mRNAs through lncRNA-mRNA interaction networks. lncRNA-mRNA interaction in A) ESCC stage 1, B) ESCC stage 2, C) ESCC stage 3, and D) ESCC stage 4.

Discussion

ESCC is one of the most prevalent malignancies in both men and women worldwide. Despite its poor prognosis and high mortality rates, the underlying mechanisms of ESCC are not well comprehended. Therefore, it is imperative to conduct meticulous investigations into the precise involvement of lncRNAs within the genomic contexts at different stages of this cancer. Such investigations are crucial in order to gain a comprehensive understanding of this particular malignancy. Notably, lncRNAs have demonstrated their potential as valuable biomarkers for various types of cancers, including ovarian cancer, colorectal cancer, thyroid cancer, and others

(Feng et al., 2023; Liao et al., 2023; Yuan et al., 2022). Moreover, lncRNAs have been implicated in ESCC as crucial players in the proliferation, invasion, metastasis, and angiogenesis (Luo et al., 2022; Ren et al., 2020; Zhang et al., 2022a).

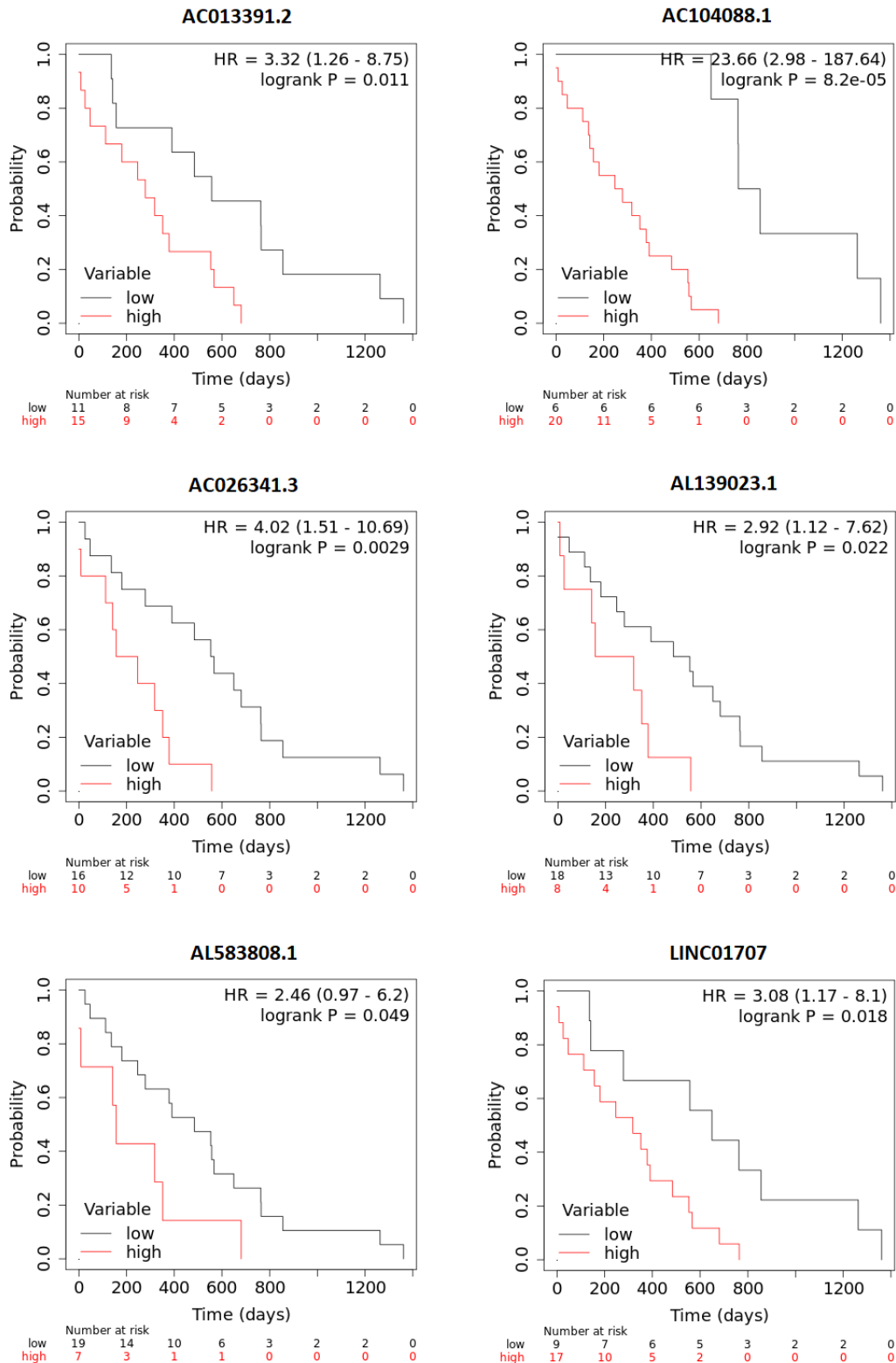


Figure 7. Kaplan–Meier survival curves of identified hub lncRNAs associated with overall survival of ESCC patients.

There are pieces of evidences, indicating the association between the dysregulation of certain

lncRNAs and the prognosis of patients with different types of cancers, such as hepatocellular carcinoma

(Wang et al., 2022), breast cancer (Li et al., 2021) and colorectal cancer (Chen et al., 2021). Moreover, transcriptome analyses have previously been conducted on patients with ESCC, and numerous research groups have documented the abnormal expression of several lncRNAs in ESCC samples, some of which can be used as ESCC prognostic biomarkers (Wang et al., 2021b; Zhang et al., 2020; Zhao et al., 2022). There are also several studies investigating the role of lncRNAs in the progression of ESCC (Chen et al., 2020; Jia et al., 2021; Tang et al., 2021). Moreover, Yu et al., identified a nine-lncRNA signature, including AC098973, AL133493, RP11-51M24, RP11-317N8, RP11-834C11, RP11-69C17, LINC00471, LINC01193 and RP1-124C based on random forest algorithm, which was able to predict the survival time of the patients (Yu et al., 2019). Zhang et al., also identified an eight-lncRNA signature by random survival forest algorithm and multivariate Cox regression analysis based on the GSE53625 dataset, which consisted of ADAMTS9-AS1, DLX6-AS1, LINC00470, LINC00520, LINC01497, LINC01749, MAMDC2-AS1, and SSTR5-AS1 (Zhang et al., 2022b). However, the exact roles of lncRNAs in different stages of this type of malignancy still remain to be elucidated. In the present study, we analyzed the differences in the mRNA and lncRNA expression profiles of normal and ESCC samples at four different stages and investigated the interactions and biological processes that are mainly associated with genes corresponding to each stage. For this purpose, the expression profile of lncRNAs along with mRNAs in four clinical stages of ESCC were investigated. Considering that WGCNA has excellent clustering efficiency for genomic materials with the correlated expression pattern, this methodology was recruited along with the investigation of lncRNA-mRNA interactions to discover more reliable stage-related hub lncRNAs. The purpose of finding hub lncRNAs, which can regulate the biological behavior of different ESCC stages, is to provide new ideas for understanding the molecular mechanisms involved in the progression of this cancer, and to candidate new targets for the treatment of ESCC. We analyzed ESCC stage-related genes, constructed protein-protein interaction networks, and consequently discovered five hub mRNAs related to each stage. lncRNAs interacting with each stage hub mRNAs were also identified as stage-related hub lncRNAs. Gene set enrichment analysis on stage-correlated modules was also performed to discover the main aberrantly regulated biological processes

corresponding to any stage during ESCC progression. Finally, prognostic risk factors were evaluated for predicting the OS time of ESCC patients. Five hub mRNAs were identified by integrating protein-protein interactions for any ESCC stage including RPS4Y1, KDM5D, USP9Y, DDX3Y and ZFY related to ESCC stage 1, BUB1B, BIRC5, MELK, DLGAP5 and NUF2 related to stage 2, GABRA3, GABRR1, GABRQ, HTR2C and BOLL related to stage 3, and MAGEA12, MAGEA3 CTAG2, CSAG2 and CT45A1 in the case of ESCC stage 4. Hub lncRNAs were also determined as lncRNAs interacting with identified hub mRNAs, which among them six lncRNAs were significantly associated with ESCC patients OS times, consisting of AC013391.2, AC104088.1, AC026341.3, AL139023.1, AL583808.1 and LINC01707. AC104088.1 is shown to have correlations with tongue squamous cell carcinoma and has been suggested as a prognostic marker for this malignancy (Hu and Messadi, 2023). AL139023.1 was also reported as the risk factor for patients' survival in lung adenocarcinoma and was correlated with poor prognosis in this type of malignancy (Geng et al., 2021). The correlation of AL583808.1 to non-small-cell lung cancer prognosis is also reported (Song et al., 2021). To our knowledge, the lncRNAs identified in this study, including AC013391.2, AC104088.1, AC026341.3, AL139023.1, AL583808.1, and LINC01707 have not been functionally annotated. Nevertheless, in the current investigation, the putative roles of these lncRNAs in ESCC were inferred by examining their interactions with mRNAs. The identification of genes associated with the expression pattern of lncRNAs, which are implicated in diverse biological processes, including spindle organization, mitotic chromosome segregation, and transmembrane transportation, implies that these lncRNAs may contribute to the progression of esophageal cancer through the modulation of these cellular mechanisms. Besides, this study accommodates some limitations which should be mentioned. First, the number of considered samples for normal esophagus squamous cells was limited, which was due to limitations in available RNA-seq samples in TCGA-ESCA project. Second, due to utilizing bulk RNA sequencing data, intratumor molecular heterogeneity might cause sampling bias. Third, the results were obtained based on computational analysis, thus, experimental investigations should be conducted to validate the introduced biomarkers, and also to examine their exact impact on ESCC progression in subsequent studies.

In the present study, a six-lncRNA signature associated with ESCC progression was identified. Notably, these lncRNAs were correlated to the overall survival time of patients with ESCC. However, their exact prognostic value should be confirmed using experimental examinations in further prospective studies in order to confirm their application in clinical settings.

Conflict of interest

The authors declare that they have no conflicts of interest to disclose.

Acknowledgments

This study was supported by Ferdowsi University of Mashhad, under grant number 50861.

References

- Aravalli, R. N., Cressman, E. N. & Steer, C. J. (2013). Cellular and molecular mechanisms of hepatocellular carcinoma: an update. *Archives of Toxicology* 87: 227-247.
- Blighe, K., Rana, S. & Lewis, M. (2019). EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. R package version 1.2.0. R package version 1.16.
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A. & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* 68: 394-424.
- Chen, J., Song, Y., Li, M., Zhang, Y., Lin, T., Sun, J., Wang, D., Liu, Y., Guo, J. & Yu, W. (2021). Comprehensive analysis of ceRNA networks reveals prognostic lncRNAs related to immune infiltration in colorectal cancer. *BMC Cancer* 21: 1-17.
- Chen, N., Zhang, G., Fu, J. & Wu, Q. (2020). Identification of key modules and hub genes involved in esophageal squamous cell carcinoma tumorigenesis using WCGNA. *Cancer Control* 27: 1073274820978817.
- Cherubini, C., Filippi, S. & Loppini, A. (2018). Systems Biology Modeling of Nonlinear Cancer Dynamics. *Systems Biology*: 203-213.
- Chin, C. H., Chen, S. H., Wu, H.H., Ho, C.W., Ko, M.T. & Lin, C.Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Systems Biology* 8: 1-7.
- Cui, H., Zhang, Y., Zhang, Q., Chen, W., Zhao, H. & Liang, J. (2017). A comprehensive genome-wide analysis of long noncoding RNA expression profile in hepatocellular carcinoma. *Cancer Medicine* 6: 2932-2941.
- DiSiena, M., Perelman, A., Birk, J. & Rezaizadeh, H. (2021). Esophageal cancer: An updated review. *Southern Medical Journal* 114: 161-168.
- Feng, J.L., Zheng, W.J., Xu, L., Zhou, Q.Y. & Chen, J. (2023). Identification of potential lncRNAs as papillary thyroid carcinoma biomarkers based on integrated bioinformatics analysis using TCGA and RNA sequencing data. *Scientific Reports* 13: 4350.
- Feng, Q., Zhang, H., Yao, D., Chen, W.D. & Wang, Y.D. (2019). Emerging role of non-coding RNAs in esophageal squamous cell carcinoma. *International Journal of Molecular Sciences* 21: 258.
- Fukunaga, T. & Hamada, M. (2017). RIBlast: an ultrafast RNA-RNA interaction prediction system based on a seed-and-extension approach. *Bioinformatics* 33: 2666-2674.
- Gao, G.D., Liu, X.Y., Lin, Y., Liu, H.F. & Zhang, G.J. (2018). lncRNA CASC9 promotes tumorigenesis by affecting EMT and predicts poor prognosis in esophageal squamous cell cancer. *European Review for Medical & Pharmacological Sciences* 22:422-429.
- Geng, W., Lv, Z., Fan, J., Xu, J., Mao, K., Yin, Z., Qing, W. & Jin, Y. (2021). Identification of the prognostic significance of somatic mutation-derived lncRNA signatures of genomic instability in lung adenocarcinoma. *Frontiers in Cell and Developmental Biology* 9: 657667.
- Hu, D. & Messadi, D. V. (2023). Immune-Related Long Non-Coding RNA Signatures for Tongue Squamous Cell Carcinoma. *Current Oncology* 30: 4817-4832.
- Huang, D. W., Sherman, B. T., Tan, Q., Collins, J. R., Alvord, W. G., Roayaei, J., Stephens, R., Baseler, M. W., Lane, H. C. & Lempicki, R. A. (2007). The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biology* 8: 1-16.
- Jia, J., Li, H., Chu, J., Sheng, J., Wang, C., Jia, Z., Meng, W., Yin, H., Wan, J. & He, F. (2021). lncRNA FAM83A-AS1 promotes ESCC progression by regulating miR-214/CDC25B axis. *Journal of Cancer* 12:1200-1211.

- Kadian, L., Arora, M., Prasad, C., Pramanik, R. & Chauhan, S. (2022). Signaling pathways and their potential therapeutic utility in esophageal squamous cell carcinoma. *Clinical and Translational Oncology* 24: 1014-1032.
- Lánczky, A. & Györfy, B. (2021). Web-based survival analysis tool tailored for medical research (KMplot): development and implementation. *Journal of Medical Internet Research* 23: 27633.
- Langfelder, P. & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9: 1-13.
- Li, G., Luo, J., Liang, C., Xiao, Q., Ding, P. & Zhang, Y. (2019). Prediction of LncRNA-disease associations based on network consistency projection. *Ieee Access* 7:58849-58856.
- Li, X., Jin, F. & Li, Y. (2021). A novel autophagy-related lncRNA prognostic risk model for breast cancer. *Journal of Cellular and Molecular Medicine* 25: 4-14.
- Li, X., Xiao, X., Chang, R. & Zhang, C. (2020). Comprehensive bioinformatics analysis identifies lncRNA HCG22 as a migration inhibitor in esophageal squamous cell carcinoma. *Journal of Cellular Biochemistry* 121: 468-481.
- Liau, X. L., Salvamani, S., Gunasekaran, B., Chellappan, D. K., Rhodes, A., Ulaganathan, V. & Tiong, Y. L. (2023). CCAT 1-A Pivotal Oncogenic Long Non-Coding RNA in Colorectal Cancer. *British Journal of Biomedical Science* 80: 11103.
- Lin, C., Wang, Y., Zhang, S., Yu, L., Guo, C. & Xu, H. (2017). Transcriptional and posttranscriptional regulation of HOXA13 by lncRNA HOTTIP facilitates tumorigenesis and metastasis in esophageal squamous carcinoma cells. *Oncogene* 36:5392-5406.
- Liu, J.Q., Deng, M., Xue, N.N., Li, T.X., Guo, Y.X., Gao, L., Zhao, D. & Fan, R.-T. (2020). lncRNA KLF3-AS1 suppresses cell migration and invasion in ESCC by impairing miR-185-5p-targeted KLF3 inhibition. *Molecular Therapy-Nucleic Acids* 20: 231-241.
- Love, M. I., Huber, W. & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15:1-21.
- Luo, X.J., He, M.M., Liu, J., Zheng, J.B., Wu, Q.N., Chen, Y.X., Meng, Q., Luo, K.J., Chen, D.L. & Xu, R.H. (2022). LncRNA TMPO-AS1 promotes esophageal squamous cell carcinoma progression by forming biomolecular condensates with FUS and p300 to regulate TMPO transcription. *Experimental & Molecular Medicine* 54: 834-847.
- Mambetsariev, I., Fricke, J., Gruber, S. B., Tan, T., Babikian, R., Kim, P., Vishnubhotla, P., Chen, J., Kulkarni, P. & Salgia, R. (2023). Clinical Network Systems Biology: Traversing the Cancer Multiverse. *Journal of Clinical Medicine* 12: 4535.
- McCabe, E. M. & Rasmussen, T. P. (2021). lncRNA involvement in cancer stem cell function and epithelial-mesenchymal transitions. *Seminars in Cancer Biology*, Vol. 75, 38-48: Elsevier.
- Pinu, F. R., Beale, D. J., Paten, A. M., Kouremenos, K., Swarup, S., Schirra, H. J. & Wishart, D. (2019). Systems biology and multi-omics integration: viewpoints from the metabolomics research community. *Metabolites* 9:76.
- Ren, L., Fang, X., Shrestha, S. M., Ji, Q., Ye, H., Liang, Y., Liu, Y., Feng, Y., Dong, J. & Shi, R. (2022). LncRNA SNHG16 promotes development of oesophageal squamous cell carcinoma by interacting with EIF4A3 and modulating RhoU mRNA stability. *Cellular & Molecular Biology Letters* 27: 89.
- Ren, P., Zhang, H., Chang, L., Hong, X.D. & Xing, L. (2020). LncRNA NR2F1-AS1 promotes proliferation and metastasis of ESCC cells via regulating EMT. *European Review for Medical & Pharmacological Sciences* 24:3686-3693.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* 13: 2498-2504.
- Sheikh, M., Roshandel, G., McCormack, V. & Malekzadeh, R. (2023). Current status and future prospects for esophageal cancer. *Cancers* 15:765.
- Shi, X., Sun, M., Liu, H., Yao, Y. & Song, Y. (2013). Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Letters* 339:159-166.
- Shin, J. W. & Chung, Y.-H. (2013). Molecular targeted therapy for hepatocellular carcinoma: current and future. *World Journal of Gastroenterology: WJG* 19: 6144.
- Siegel, R. L., Miller, K. D., Wagle, N. S. & Jemal, A. (2023). Cancer statistics, 2023. *Ca Cancer J Clin* 73:17-48.

- Song, J., Zhang, S., Sun, Y., Gu, J., Ye, Z., Sun, X. & Tang, Q. (2021). A radioresponse-related lncRNA biomarker signature for risk classification and prognosis prediction in non-small-cell lung cancer. *Journal of Oncology* 2021:4338838.
- Sun, M. & Kraus, W. L. (2015). From discovery to function: the expanding roles of long noncoding RNAs in physiology and disease. *Endocrine Reviews* 36:25-64.
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A. & Tsafou, K. P. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research* 43: 447-452.
- Tan, Y. T., Lin, J. F., Li, T., Li, J. J., Xu, R. H. & Ju, H. Q. (2021). LncRNA-mediated posttranslational modifications and reprogramming of energy metabolism in cancer. *Cancer Communications* 41: 109-120.
- Tang, J., Xu, H., Liu, Q., Zheng, J., Pan, C., Li, Z., Wen, W., Wang, J., Zhu, Q. & Wang, Z. (2021). LncRNA LOC146880 promotes esophageal squamous cell carcinoma progression via miR-328-5p/FSCN1/MAPK axis. *Aging (Albany NY)* 13:14198.
- Wang, B.r., Chu, D.x., Cheng, M.y., Jin, Y., Luo, H.g. & Li, N. (2022). Progress of HOTAIR-microRNA in hepatocellular carcinoma. *Hereditary Cancer in Clinical Practice* 20: 4.
- Wang, C., Song, C., Liu, Q., Zhang, R., Fu, R., Wang, H., Yin, D., Song, W., Zhang, H. & Dou, K. (2021a). Gene expression analysis suggests immunological changes of peripheral blood monocytes in the progression of patients with coronary artery disease. *Frontiers in Genetics* 12: 641117.
- Wang, P., Chen, Y., Zheng, Y., Fu, Y. & Ding, Z. (2021b). Identification of epithelial-mesenchymal transition-(EMT-) related LncRNA for prognostic prediction and risk stratification in esophageal squamous cell carcinoma. *Disease Markers* 2021: 1-18.
- Wickham, H., Chang, W. & Wickham, M. H. (2016). Package 'ggplot2'. Create Elegant Data Visualisations Using the Grammar of Graphics. Version 2:1-189.
- Wu, Q., Zhang, H., Yang, D., Min, Q., Wang, Y., Zhang, W. & Zhan, Q. (2022). The m6A-induced lncRNA CASC8 promotes proliferation and chemoresistance via upregulation of hnRNPL in esophageal squamous cell carcinoma. *International Journal of Biological Sciences* 18: 4824-4836.
- Yang, M., Lu, H., Liu, J., Wu, S., Kim, P. & Zhou, X. (2022). lncRNAfunc: a knowledgebase of lncRNA function in human cancer. *Nucleic Acids Research* 50: 1295-1306.
- Yu, J., Wu, X., Huang, K., Zhu, M., Zhang, X., Zhang, Y., Chen, S., Xu, X. & Zhang, Q. (2019). Bioinformatics identification of lncRNA biomarkers associated with the progression of esophageal squamous cell carcinoma. *Molecular Medicine Reports* 19: 5309-5320.
- Yu, S. J. (2016). A concise review of updated guidelines regarding the management of hepatocellular carcinoma around the world: 2010-2016. *Clinical and Molecular Hepatology* 22: 7.
- Yuan, D., Guo, T., Zhu, D., Ge, H., Zhao, Y., Huang, A., Wang, X., Cao, X., He, C. & Qian, H. (2022). Exosomal lncRNA ATB derived from ovarian cancer cells promotes angiogenesis via regulating miR-204-3p/TGFβR2 axis. *Cancer Management and Research*: 327-337.
- Zhang, H., Pan, E., Zhang, Y., Zhao, C., Liu, Q., Pu, Y. & Yin, L. (2022a). LncRNA RPL34-AS1 suppresses the proliferation, migration and invasion of esophageal squamous cell carcinoma via targeting miR-575/ACAA2 axis. *BMC Cancer* 22: 1017.
- Zhang, J., Ling, X., Fang, C. & Ma, J. (2022b). Identification and validation of an eight-lncRNA signature that predicts prognosis in patients with esophageal squamous cell carcinoma. *Cellular & Molecular Biology Letters* 27: 1-18.
- Zhang, L., Li, P., Liu, E., Xing, C., Zhu, D., Zhang, J., Wang, W. & Jiang, G. (2020). Prognostic value of a five-lncRNA signature in esophageal squamous cell carcinoma. *Cancer Cell International* 20: 1-10.
- Zhao, F., Dong, Z., Li, Y., Liu, S., Guo, P., Zhang, D. & Li, S. (2022). Comprehensive analysis of molecular clusters and prognostic signature based on m7G-related LncRNAs in esophageal squamous cell carcinoma. *Frontiers in Oncology* 12:893186.
- Zheng, Z.J., Li, Y.S., Zhu, J.D., Zou, H.Y., Fang, W.K., Cui, Y.Y. & Xie, J.J. (2022). Construction of the Six-lncRNA prognosis signature as a novel biomarker in esophageal squamous cell carcinoma. *Frontiers in Genetics* 13:839589.

Open Access Statement:

This is an open access article distributed under the Creative Commons Attribution License (CC-BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.