

Identification of crucial miRNAs and genes in esophageal squamous cell carcinoma by miRNA-mRNA integrated analysis

Xiaowu Zhong, PhD^{a,b,c}, Guangcheng Huang, MD^a, Qiang Ma, MD^a, Hebin Liao, MD^b, Chang Liu, MD^c, Wenjie Pu, MD^c, Lei Xu, MD^b, Yan Cai, PhD^a, Xiaolan Guo, PhD^{a,b,c,*}

Abstract

Esophageal squamous cell carcinoma (ESCC) is a malignancy that severely threatens human health and carries a high incidence rate and a low 5-year survival rate. MicroRNAs (miRNAs) are commonly accepted as a key regulatory function in human cancer, but the potential regulatory mechanisms of miRNA-mRNA related to ESCC remain poorly understood.

The GSE55857, GSE43732, and GSE6188 miRNA microarray datasets and the gene expression microarray datasets GSE70409, GSE29001, and GSE20347 were downloaded from Gene Expression Omnibus databases. The differentially expressed miRNAs (DEMs) and differentially expressed genes (DEGs) were obtained using GEO2R. Gene ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis for DEGs were performed by Database for Annotation, Visualization and Integrated Discovery (DAVID). A protein-protein interaction (PPI) network and functional modules were established using the STRING database and were visualized by Cytoscape. Kaplan-Meier analysis was constructed based on The Cancer Genome Atlas (TCGA) database.

In total, 26 DEMs and 280 DEGs that consisted of 96 upregulated and 184 downregulated genes were screened out. A functional enrichment analysis showed that the DEGs were mainly enriched in the ECM-receptor interaction and cytochrome P450 metabolic pathways. In addition, *MMP9*, *PCNA*, *TOP2A*, *MMP1*, *AURKA*, *MCM2*, *IVL*, *CYP2E1*, *SPRR3*, *FOS*, *FLG*, *TGM1*, and *CYP2C9* were considered to be hub genes owing to high degrees in the PPI network. MiR-183-5p was with the highest connectivity target genes in hub genes. FOS was predicted to be a common target gene of the significant DEMs. Hsa-miR-9-3p, hsa-miR-34c-3p and FOS were related to patient prognosis and higher expression of the transcripts were associated with a poor OS in patients with ESCC.

Our study revealed the miRNA-mediated hub genes regulatory network as a model for predicting the molecular mechanism of ESCC. This may provide novel insights for unraveling the pathogenesis of ESCC.

Abbreviations: AURKA = aurora kinase A, BP = biological process, CC = cellular component, CYP2C9 = cytochrome P450 2C9, CYP2E1 = cytochrome P450 2E1, DAVID = database for annotation, visualization and integrated discovery, DEGs = differentially expressed genes, DEMs = differentially expressed miRNAs, EC = esophageal carcinoma, EMT = epithelial-mesenchymal transition, ESCC = esophageal squamous cell carcinoma, FDR = false discovery rate, FLG = filaggrin, GEO = gene expression omnibus, GO = gene ontology, IVL = involucrin, KEGG = Kyoto Encyclopedia of Genes and Genomes, MCM2 = minichromosome maintenance protein 2, MF = molecular function, MMPs = matrix metalloproteinases, OS = overall survival, PCNA = proliferating cell nuclear antigen, PDCD4 = programmed Cell Death 4, PDK1 = phosphoinositide dependent kinase 1, PPI = protein-protein interactions, SPRR3 = small proline-rich repeat protein 3, STRING = search tool for the retrieval of interacting genes/proteins, TGM1 = Transglutaminase-1, TOP2A = topoisomerase II α .

Keywords: bioinformatics, esophageal squamous cell carcinoma, miRNA, regulatory network

Editor: Surinder Kumar.

XZ, GH, and HL contributed equally to this work.

This study was a secondary retrospective study. No ethical approval was required.

The authors acknowledge financial support from the Applied Basic Research Program of Sichuan Province (Grant No. 2017JY0151), and Science and Technology Support Program of Nanchong (Grant Nos. NSMC20170436, NCKL201703, and 18SXHZ0514).

The authors have no conflicts of interest to disclose.

^a Department of Clinical Laboratory, Affiliated Hospital of North Sichuan Medical College, ^b Translational Medicine Research Center, ^c Department of Laboratory Medicine, North Sichuan Medical College, Nanchong, Sichuan, China.

* Correspondence: Xiaolan Guo, Department of Clinical Laboratory, Affiliated Hospital of North Sichuan Medical College, Nanchong 637000, Sichuan, China (e-mail: alan5200@hotmail.com).

Copyright © 2019 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medicine (2019) 98:27(e16269)

Received: 9 January 2019 / Received in final form: 14 May 2019 / Accepted: 8 June 2019

<http://dx.doi.org/10.1097/MD.00000000000016269>

1. Introduction

Esophageal carcinoma (EC) is the ninth most common cancer worldwide, with 572,034 new cases and 508,585 mortalities reported in 2018.^[1] The incidence and mortality of EC are rapidly growing worldwide in developing countries. Esophageal squamous cell carcinoma (ESCC) is the predominant histological type of EC, which is more prevalent in the developing world, with very high incidence areas found in East Asia, mainly in China.^[2,3] Epidemiological evidence suggests a different risk factor profile including cigarette smoking, alcohol consumption, highly salted meats, and so on.^[4,5] If earlier detection and aggressive medical and surgical treatment of EC could be achieved, then the five-year survival rates would exceed 80%.^[6,7] Unfortunately, because effective early diagnosis for EC remains elusive, EC often presents in an insidious and nonspecific manner, and the one-year survival rate remains lower than 15%.^[8–10] Therefore, the identification of a molecular mechanism and more efficient diagnostic methods for the early stages of EC can help to develop treatments for patients with ESCC.

During recent years, developments in molecular biology have provided gene expressions profiles for studying the pathogenesis of ESCC. Many scholars have carried out in-depth research on the pathogenesis of ESCC at the gene level, through which variety of significant genes were found. For instance, it was found that Programmed Cell Death 4 (PDCD4) is down-regulated in ESCC.^[11] Phosphoinositide dependent kinase 1 (PDK1) is reported to be upregulated in ESCC, which is associated with a poor prognosis.^[12] In addition, microRNAs (miRNAs) are small noncoding RNA molecules of 20 to 25 nucleotides in length that participate in posttranscriptional regulation of gene expression. It has previously been demonstrated that miRNAs play an important role in tumor initiation and progression.^[13] In ESCC, miR-375 is frequently down-regulated in cancer cells and functions as a tumor suppressor.^[14] MiR-183 might play an oncogenic role in suppressing apoptosis and promoting proliferation in ESCC by regulating the PDCD4 expression.^[15] It is known that miR-21 acts as an oncogenic miRNA in several types of cancer. The over-expression of miR-21 is related to advanced clinical stage, lymph node metastasis and a poor prognosis in ESCC.^[16,17] However, little is known regarding an integrated analysis of the miRNA-mRNA regulatory network in ESCC. Therefore, further study is imperative.

In this study, we analyzed differentially expressed miRNAs (DEMs) and genes (DEGs) between ESCC tumor tissues and normal tissue samples using bioinformatics methods. We performed functional and pathway enrichment analysis of DEGs, and established the miRNA-mRNA and protein-protein interactions (PPI) network to reveal regulatory mechanisms in ESCC. Here, we aimed to explore the main pathways and processes that

are associated with ESCC and provide cancer biomarkers for diagnosis and therapy.

2. Materials and methods

2.1. Microarray data collection

The GSE55857, GSE43732, and GSE6188 miRNA expression profile data and three gene expression profiles (GSE70409, GSE29001, and GSE20347) were acquired from Gene Expression Omnibus (GEO: <http://www.ncbi.nlm.nih.gov/geo/>).^[18] For these datasets, only normal ESCC and ESCC tissue samples were selected for further analysis. Every included dataset contained more than 10 normal ESCC tissues and ESCC samples. Basic information on the miRNA and mRNA expression microarrays of ESCC is shown in Table 1.

2.2. Identification of DEMs and DEGs

The differentially expressed miRNAs (DEMs) and differentially expressed genes (DEGs) were obtained from 2 groups of samples (ESCC vs normal ESCC tissues) in each GEO database. The raw data of miRNA and gene expression profiles were preprocessed by a way of GEO2R analysis (www.ncbi.nlm.nih.gov/geo/geo2r/). The adjusted *P* value < .01 and $|\log_{2}FC| > 2$ were set as the cut-off criteria of the DEM and DEG analyses to identify a statistically significant difference.

The miRWalk 2.0 database (<http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2/index.html>) is a comprehensive atlas that supplies a large collection of predicted and experimentally verified miRNA-target interactions.^[19] We submitted the significant DEMs to miRwalk 2.0 and selected the data of their target mRNAs, which were based on experimental literature reports. Then, we extracted by intersecting the target mRNAs of the significant DEMs and the overlapping genes of the DEGs from the three GEO datasets. Finally, these genes were defined as significant DEGs (Fig. 1).

2.3. Gene ontology and KEGG pathway analysis

The Database for Annotation, Visualization and Integrated Discovery (DAVID, <https://david-d.ncicrf.gov/>) is an online bioinformatics resource that provides functional and pathway enrichment analysis.^[20] Gene ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis for DEGs was performed by with the DAVID online tool. A false discovery rate (FDR) < 0.05 and gene count > 2 were used as the cut-off criteria.

2.4. Construction of regulatory network

Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, <https://string-db.org/>) is an online platform that is used

Table 1

Basic information for the miRNA and mRNA expression microarrays for ESCC.

Gene/miRNA	GEO ID	Platform	Case: Control	Sample type	Time	Authors
miRNA	GSE55857	GPL14613 [miRNA-2] Affymetrix Multispecies miRNA-2 Array	108:108	Esophageal squamous cell carcinoma	2017	Jang et al
miRNA	GSE43732	GPL16543 Agilent-038166 cbc_human_miR18.0	119:119	Esophageal squamous cell carcinoma	2014	Chen et al
miRNA	GSE6188	GPL4508 Tsinghua University mammalian 2K microRNA microarray	104:153	Esophageal squamous cell carcinoma	2008	Guo et al
Gene	GSE70409	GPL13287 Phalanx Human OneArray [Annotation HOA5 release 1.0]	17:17	Esophageal squamous cell carcinoma	2016	Chen et al
Gene	GSE29001	GPL571 [HG-U133A_2] Affymetrix Human Genome U133A 2.0 Array	21:24	Esophageal squamous cell carcinoma	2012	Yan et al
Gene	GSE20347	GPL571 [HG-U133A_2] Affymetrix Human Genome U133A 2.0 Array	17:17	Esophageal squamous cell carcinoma	2010	Hu et al

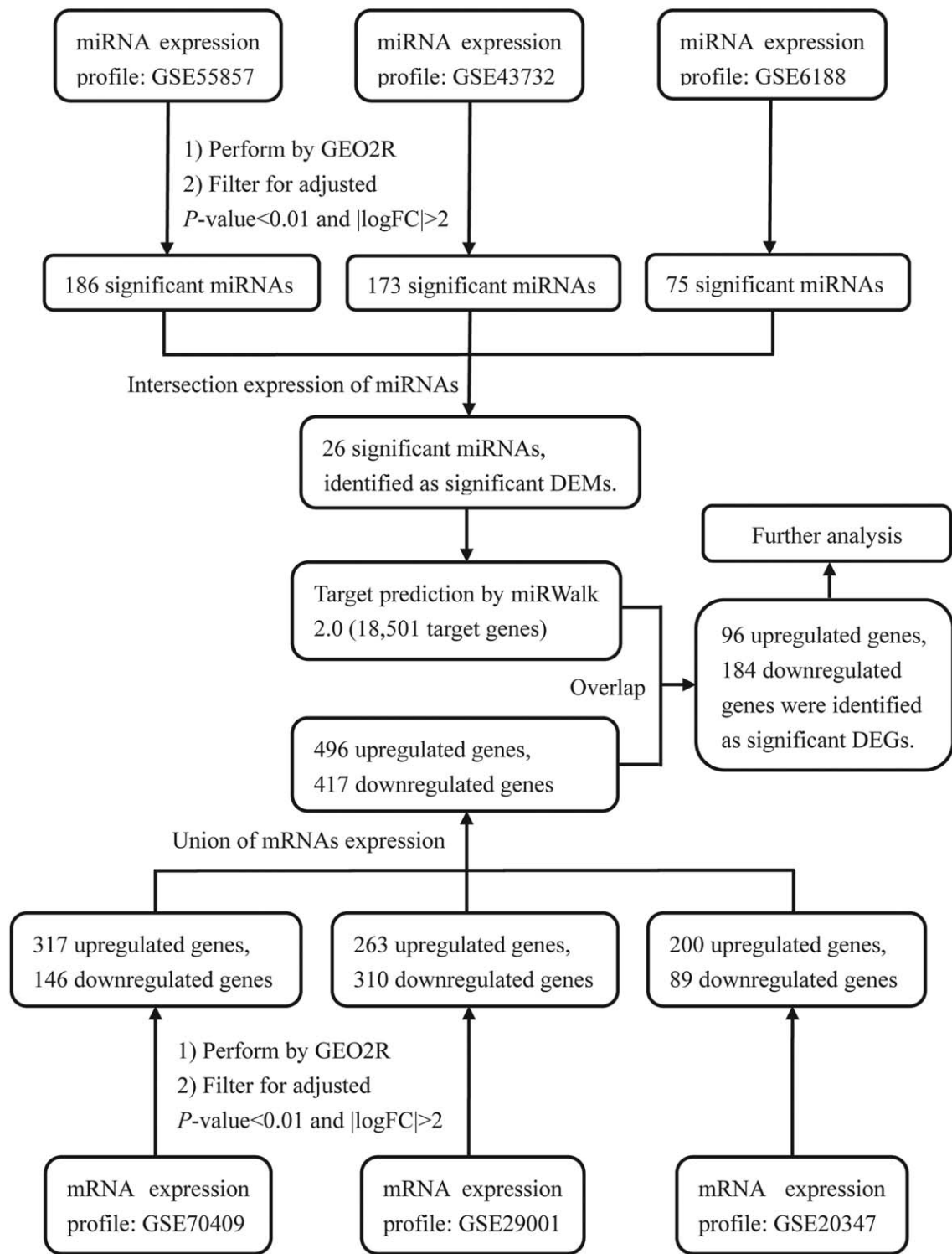


Figure 1. Flowchart of bioinformatics analysis.

to predict functional protein association networks.^[21] To construct protein-protein interactions (PPI) of DEGs, we submitted the DEGs to STRING, and only validated interactions that had a combined score >0.40, which was considered to be significant and the PPI was visualized.

Furthermore, the target genes of the significant DEMs were predicted by miRTarBase ([http://mirtarbase.mbc.nctu.](http://mirtarbase.mbc.nctu.edu.tw/)

<http://www.targetscan.org/>).^[23] When the miRNAs shared a common target mRNA with the hub genes of DEGs, they might exist in a similar regulatory pathway. Finally, the miRNA-mRNA regulatory network depicted interactions between miRNAs and their potential targets in ESCC were visualized by using Cytoscape 3.7.0.^[24]

2.5. Survival analysis

A Kaplan–Meier analysis was performed using OncoLnc (<http://www.oncolnc.org/>),^[25] in which boxplots were employed to visualize the expression level of all prognosis-related genes between ESCC and normal tissues, that was based on data from the TCGA database (<https://cancergenome.nih.gov/>). Data from 144 ESCC patients were found in the TCGA database. Among them, 117 patients were male, and 27 were female. Among the 144 available clinical outcome events, 85 patients were living and 59 were deceased when the follow-up period ended. The *P* value < .05 was considered to be statistically significant.

3. Results

3.1. Identification of DEMs and DEGs

A total of 186, 173, and 75 DEMs were identified from the GSE55857, GSE43732, and GSE6188 datasets, respectively. As shown in Figure 2A, 26 DEMs were screened out in all 3 datasets, so that they could be identified as significant DEMs, which were hsa-miR-18b-5p, hsa-miR-503-5p, hsa-miR-9-3p, hsa-miR-

133a-3p, hsa-miR-34b-5p, hsa-miR-34c-5p, hsa-miR-339-3p, hsa-miR-375, hsa-miR-133b, hsa-miR-203a, hsa-miR-150-5p, hsa-miR-493-3p, hsa-miR-18a-3p, hsa-miR-130b-3p, hsa-miR-431-5p, hsa-miR-7-5p, hsa-miR-382-5p, hsa-miR-182-3p, hsa-miR-424-5p, hsa-miR-146b-5p, hsa-miR-196a-3p, hsa-miR-183-5p, hsa-miR-1-3p, hsa-miR-196b-3p, and hsa-miR-497-5p.

Compared with the normal samples, a total of 496 upregulated DEGs and 417 downregulated DEGs were identified in the ESCC samples (Fig. 2B and C). We obtained 18,501 target genes of the 26 significant DEMs from the miRWalk 2.0 database. The intersection number of these target genes and the upregulated DEGs that appeared in the three datasets was 96, while the intersection number of these candidate genes and the down-regulated DEGs which appeared in three datasets was 184. Therefore, the 96 upregulated and 184 downregulated genes were identified as the final sets of significant DEGs.

3.2. Functional annotation analysis

GO ontology contains three terms: molecular function (MF), cellular component (CC) and biological process (BP). In MF

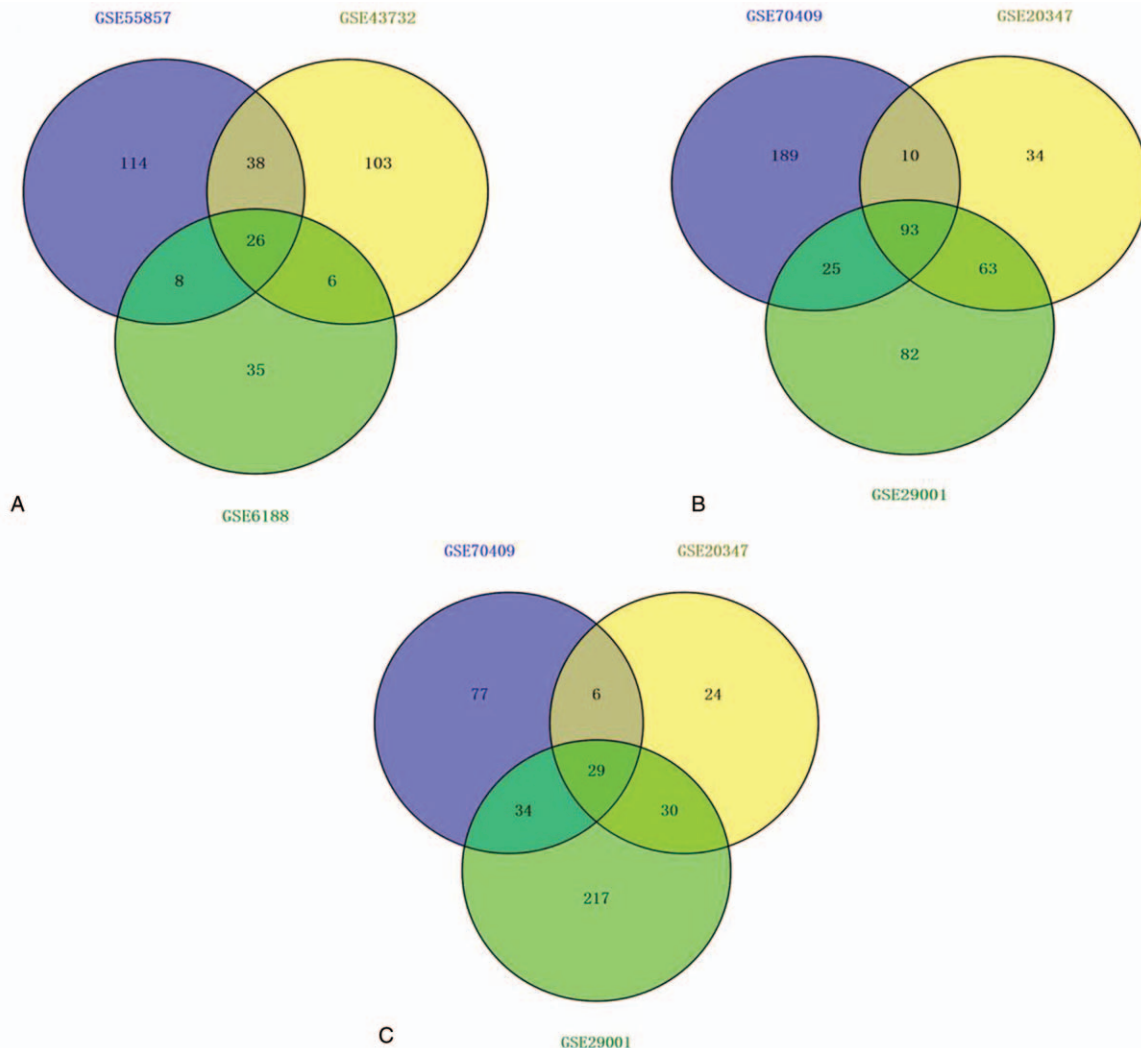


Figure 2. Identification of differentially expressed miRNAs (DEMs) and genes (DEGs). (A) Identification of DEMs; (B) Identification of upregulated DEGs; (C) Identification of downregulated DEGs.

Table 2**Significantly enriched GO biological process terms and KEGG pathways.**

Category	Term	Description	Gene counts	FDR	
Upregulated					
	GO:0005201	MF	Extracellular matrix structural constituent	7	2.43E-03
	GO:0048407	MF	Platelet-derived growth factor binding	4	3.36E-02
	GO:0004222	MF	Metalloendopeptidase activity	7	5.00E-02
	GO:0005576	CC	Extracellular region	33	6.98E-09
	GO:0005578	CC	Proteinaceous extracellular matrix	15	9.88E-08
	GO:0031012	CC	Extracellular matrix	14	4.07E-06
	GO:0005581	CC	Collagen trimer	9	2.64E-05
	GO:0005615	CC	Extracellular space	24	2.31E-04
	GO:0005788	CC	Endoplasmic reticulum lumen	10	7.09E-04
	GO:0030574	BP	Collagen catabolic process	14	2.64E-14
	GO:0030198	BP	Extracellular matrix organization	14	7.71E-08
	GO:0001501	BP	Skeletal system development	12	3.36E-07
	GO:0022617	BP	Extracellular matrix disassembly	9	1.28E-05
	GO:0030199	BP	Collagen fibril organization	7	1.00E-04
	hsa04512	KEGG	ECM-receptor interaction	8	1.57E-03
	hsa05146	KEGG	Amoebiasis	8	5.94E-03
	hsa04510	KEGG	Focal adhesion	10	7.37E-03
	hsa04974	KEGG	Protein digestion and absorption	7	2.47E-02
Downregulated					
	GO:0005506	MF	Iron ion binding	11	4.01E-03
	GO:0004497	MF	Monoxygenase activity	7	3.10E-02
	GO:0070062	CC	Extracellular exosome	75	3.26E-14
	GO:0001533	CC	Cornified envelope	8	3.55E-04
	GO:0005615	CC	Extracellular space	31	2.13E-02
	GO:0030216	BP	Keratinocyte differentiation	11	7.71E-06
	GO:0031424	BP	Keratinization	8	8.14E-04
	GO:0018149	BP	Peptide cross-linking	8	1.09E-03
	GO:0030855	BP	Epithelial cell differentiation	8	1.10E-02
	GO:0008544	BP	Epidermis development	8	3.98E-02
	hsa00982	KEGG	Drug metabolism - cytochrome P450	7	5.00E-02

BP = biological process, CC = cellular component, FDR = false discovery rate, GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes, MF = molecular function.

ontology, the most significant GO terms for upregulated significant DEGs were the extracellular matrix structural constituent, platelet-derived growth factor binding and metalloendopeptidase activity. Whereas the terms for the downregulated significant DEGs, the terms were iron ion binding and monoxygenase activity. In CC ontology, the extracellular region, the proteinaceous extracellular matrix, the extracellular matrix, the collagen trimer, the extracellular space and the endoplasmic reticulum lumen were significantly enriched GO terms for upregulated genes, while the extracellular exosome, the cornified envelope and the extracellular space were significantly enriched GO terms for downregulated genes. In BP ontology, the upregulated genes were mainly enriched in collagen catabolic process, extracellular matrix organization, skeletal system development, extracellular matrix disassembly and collagen fibril organization. The downregulated genes were mainly enriched in keratinocyte differentiation, keratinization, peptide cross-linking, epithelial cell differentiation and epidermis development.

Moreover, 4 KEGG pathways were overrepresented in the upregulated genes, including ECM-receptor interaction, amoebiasis, focal adhesion and protein digestion and absorption, and only 1 pathway (drug metabolism-cytochrome P450) was identified in the downregulated genes. The detailed results are presented in Table 2.

3.3. PPI network

The PPI network of DEGs was established on the basis of STRING. The 103 nodes and 316 edges in total constituted the PPI network of significantly upregulated DEGs (Fig. 3). The network that significantly downregulated DEGs was composed of 188 nodes and 138 edges (Fig. 4).

The edge information plays an important role in the identification of hub genes in a PPI network. We used the parameter “degree” to calculate the edge counts of every single gene in a PPI network. The top 5% degree genes are shown in Table 3, which were assessed as hub genes. Thirteen genes in the PPI network were selected as hub genes in ESCC. These hub genes might play crucial roles in ESCC.

3.4. miRNA-mRNA interaction network

To further understand the regulatory relationship between 26 significant DEMs and hub genes, the miRNA-mRNA regulation network was established and is shown in Figure 5. There were no hub genes that could be the target genes of hsa-miR-339-3p, hsa-miR-375 or hsa-miR-196b-3p. MiR-183-5p was with the highest connectivity target genes, targeting 7 hub genes (upregulated: MMP9, TOP2A, AURKA, and MCM2; downregulated: IVL, FOS, and CYP2C9). Five hub genes could be target genes of hsa-miR-9-3p and hsa-miR-130b-3p. Moreover, several genes were predicted to be common targets of different miRNAs. For

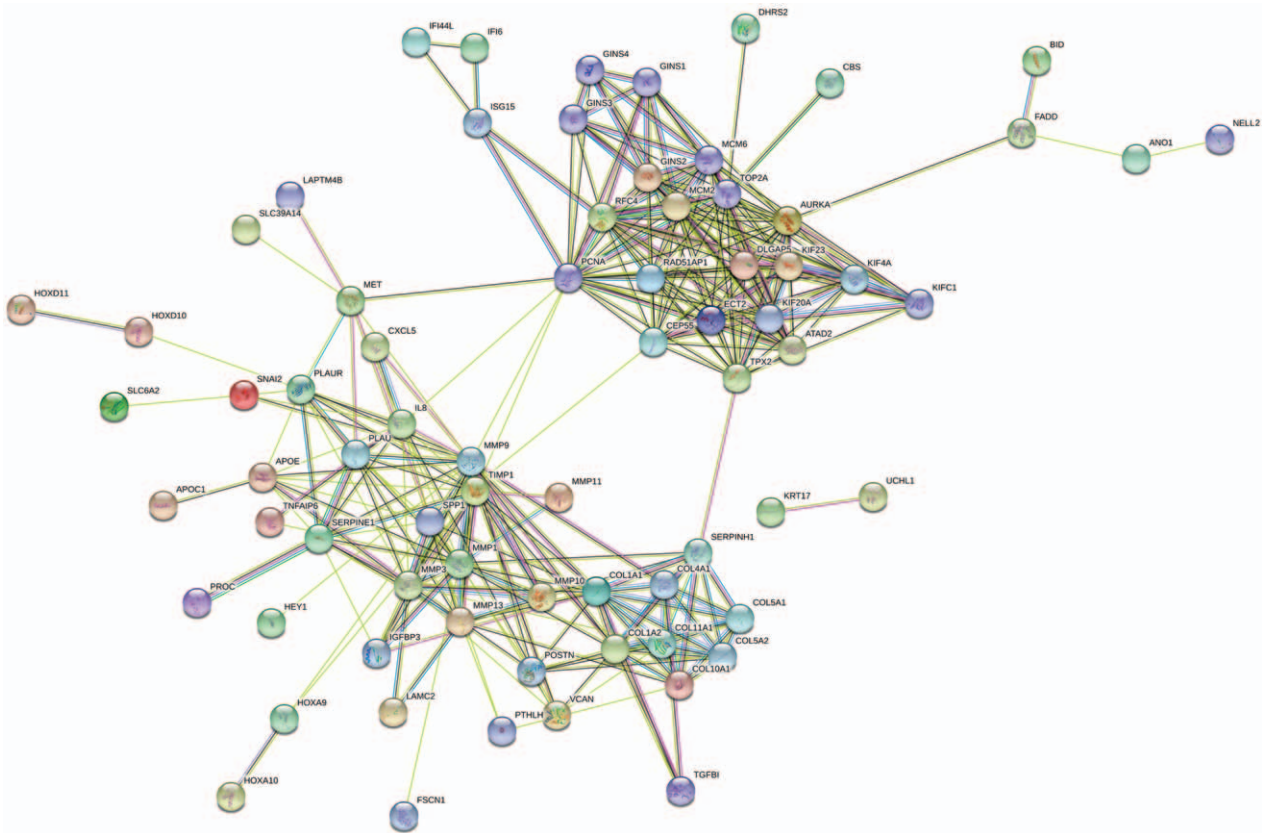


Figure 3. PPI networks of significantly upregulated DEGs. Colored nodes: query proteins and first shell of interactors, white nodes: second shell of interactors; Blue-green line: known interactions from curated databases, purple line: known interactions from experimentally determined, green line: predicted interactions form gene neighborhood, red line: predicted interactions form gene fusions, dark blue: predicted interactions form gene co-occurrence, yellow line: interactions form textmining, black line: interactions form co-expression, light blue: interactions form protein homology.

instance, FOS was predicted as a common target of hsa-miR-133a-3p, hsa-miR-34b-5p, hsa-miR-34c-5p, hsa-miR-133b, hsa-miR-150-5p, hsa-miR-130b-3p, hsa-miR-431-5p, hsa-miR-382-5p, hsa-miR-424-5p, hsa-miR-196a-3p, hsa-miR-183-5p, and hsa-miR-497-5p.

3.5. Stage and survival analysis of miRNA/mRNA in ESCC

We investigated the miRNAs expression level of 26 significant DEMs in the TCGA dataset, and found that 23 miRNAs were consistent with the results from GEO databases. However, the expression levels of hsa-miR-203a, hsa-miR-182-3p, and hsa-miR-1-3p were not significantly different between the ESCC samples and normal ESCC tissues. Using the TCGA ESCA dataset, we assessed the relationship of miRNAs with the clinical TNM stages. Detailed results of the significant DEMs are presented in Table 4.

Based on the TCGA survival data, survival analysis using the Kaplan–Meier method and the log-rank test was conducted. The results indicated that hsa-miR-9-3p, hsa-miR-34c-3p and FOS were related to overall survival (OS) in ESCC patients (Fig. 6). Higher expression of the transcript was associated with the poor OS in patients with ESCC.

4. Discussion

ESCC is one of the most lethal cancers and is a public health issue of great concern worldwide.^[26] Despite many advances in

diagnosis and treatment, there is still no effective treatment for ESCC, and survival remains very poor.^[27] Thus, the molecular mechanisms that are involved in the progress of ESCC have not been clarified. Therefore, it is crucial to study the mechanism and identify the molecular targets for diagnosis and treatment. In the present study, we identified 26 significant DEMs, 96 upregulated DEGs, and 184 downregulated DEGs. The results of the functional enrichment analysis indicated that significant DEGs were related to the ECM-receptor interaction and cytochrome P450 pathways in ESCC. Thirteen genes in the PPI network were selected as hub genes in ESCC. Importantly, hsa-miR-9-3p, hsa-miR-34c-3p and FOS were validated and were found to be correlated with tumor stages and survival, which meant they could not only regulate cellular process but could also be of valuable in clinical practice.

As was suggested by the functional enrichment analysis of significant DEGs, we found that the DEGs were mainly enriched in biological processes including, the ECM-receptor interaction and the cytochrome P450 metabolic pathways, and so on. ECM-receptor interaction was a chief contributor of cancer progression.^[28] Cytochrome P450 levels showed significant differences between ESCC patients and healthy subjects and may contribute to the development of ESCC.^[29] The genetic polymorphism RsaI/PstI in cytochrome P450 2E1 (CYP2E1) was a risk factor for EC. It demonstrated that CYP2E1 was a genetic determinant in the development of ESCC.^[30] CYP2C9 expression was relevant for high Ki-67 labeling indices in EC. Moreover, inhibition of

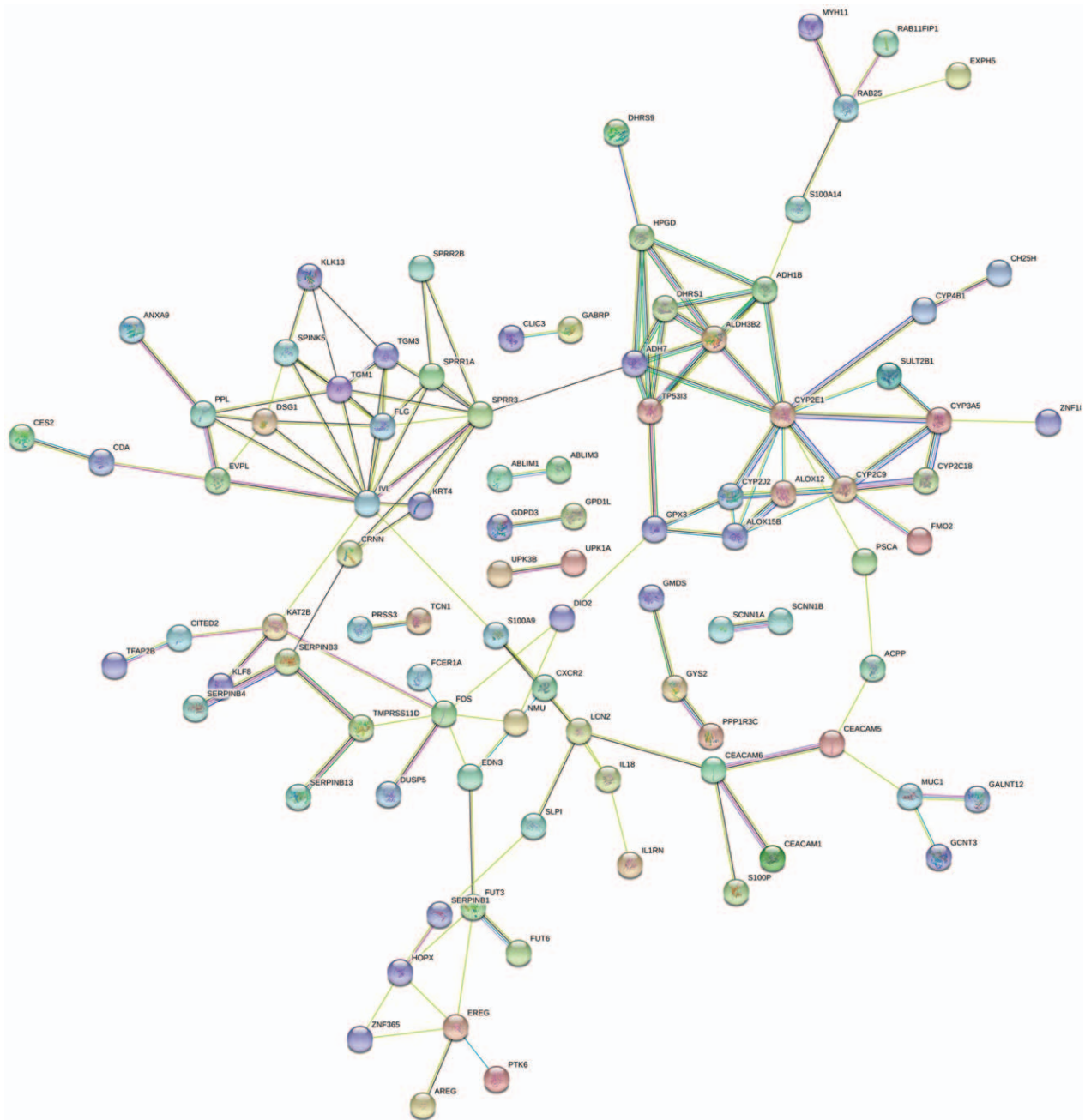


Figure 4. PPI networks of significantly downregulated DEGs. Colored nodes: query proteins and first shell of interactors, white nodes: second shell of interactors; Blue-green line: known interactions from curated databases, purple line: known interactions from experimentally determined, green line: predicted interactions from gene neighborhood, red line: predicted interactions from gene fusions, dark blue: predicted interactions from gene co-occurrence, yellow line: interactions from text mining, black line: interactions from co-expression, light blue: interactions from protein homology.

cytochrome P450 2C9 (CYP2C9) could affect the tumor cell proliferation in early EC development.^[31]

A growing body of evidence suggests that miRNAs have significant roles in human tumorigenesis, tumor progression and metastasis.^[32] MiRNAs are unique in their ability to regulate mRNA, which is more than 60% of protein-coding genes.^[33] The aberrant expression of miRNAs in ESCC causes destruction to the miRNA-regulated mRNA networks and can function as tumor suppressors or oncogenes.^[34] Therefore, the identification of the miRNA-mRNA regulatory network is important for

further research concerning ESCC. Compared with normal samples, a microarray analysis identified 26 significant DEMs in our results. MiR-18a/b,^[35,36] miR-503,^[37] miR-9,^[38] miR-133a/b,^[39,40] miR-34b/c,^[41,42] miR-375,^[43] miR-203a,^[44] miR-150,^[45] miR-130b,^[46] miR-7,^[47] miR-382,^[48] miR-424,^[49] miR-146b,^[50] miR-196a,^[51] miR-183,^[15,52] and miR-1^[53] were reported in ESCC studies, and mostly involved in ESCC pathogenesis. MiR-18a overexpression was positively correlated with the stage that promoted the expression Cyclin D1 regulating PTEN-PI3K-AKT-mTOR signaling axis in ESCC cells. MiR-150

Table 3
Top 5% hub genes in the PPI networks.

Ensembl gene ID	Gene symbol	Full gene name	Degree
Upregulated			
ENSP00000361405	MMP9	Matrix metalloproteinase 9	22
ENSP00000368438	PCNA	Proliferating cell nuclear antigen	22
ENSP00000411532	TOP2A	Topoisomerase (DNA) II alpha	19
ENSP00000322788	MMP1	Matrix metalloproteinase 1	19
ENSP00000216911	AURKA	Aurora kinase A	19
ENSP00000265056	MCM2	Minichromosome maintenance complex component 2	19
Downregulated			
ENSP00000357753	IVL	Involucrin	12
ENSP00000301258	CYP2E1	Cytochrome P450 family 2 subfamily E member 1	11
ENSP00000448220	SPRR3	Small proline rich protein 3	9
ENSP00000306245	FOS	Fos proto-oncogene	7
ENSP00000357789	FLG	Filaggrin	7
ENSP00000470555	TGM1	Transglutaminase 1	7
ENSP00000260682	CYP2C9	Cytochrome P450 family 2 subfamily C member 9	7

plays an oncogenic role in ESCC. MiR-150 promotes ESCC cell migration and invasion by directly targeting ZEB1, SPOCK1 and Gli1.^[45,54,55] MiR-133a, miR-133b and miR-375 are known diagnostic and prognostic markers that are associated with tumor suppressor miRNAs. MiR-133a, miR-133b and miR-375 are frequently downregulated in ESCC and are closely related to advanced clinical stage, tumor metastasis and poor prognosis.

MiR-133a can suppress the migration and invasion of ESCC cells by targeting Sox4 and the EMT process. MiR-133b can regulate metastases of ESCC by affecting the MAPK/ERK and PI3K/AKT signaling pathways by targeting EGFR. MiR-375 is involved in the development and progression of ESCC by repressing metadherin expression.^[39,40,43] In this study, higher expression levels of hsa-miR-9-3p and hsa-miR-34c-3p were found to be

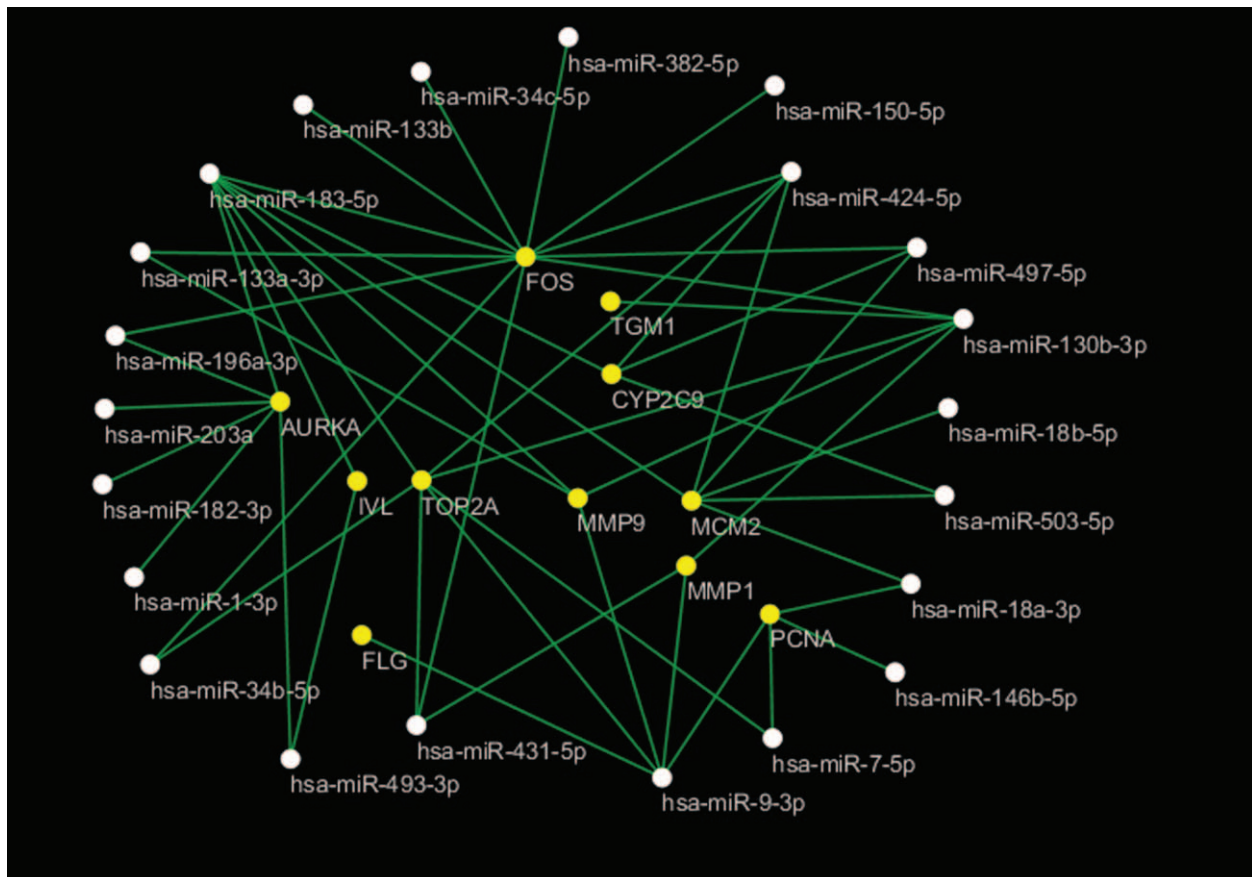


Figure 5. The miRNA-mRNA regulatory network. White nodes, miRNA; Green nodes, mRNA.

Table 4

Results of the significant DEMs expression validation and the relationship with ESCC clinical stages in the TCGA dataset.

miRNA	Higher expression	P value (t test)	Clinical stages	P value (chi-square)
hsa-miR-18b-5p	Tumor	2.07E-04	Pathologic T Status/Clinical M Status	5.63E-04
hsa-miR-503-5p	Tumor	6.30E-03	Pathologic M Status	3.26E-02
hsa-miR-9-3p	Tumor	3.29E-02	Pathologic T/M Status	-
hsa-miR-133a-3p	Normal	1.09E-03	-	1.06E-02
hsa-miR-34b-5p	Tumor	5.41E-03	Clinical M Status	1.35E-02
hsa-miR-34c-5p	Tumor	9.45E-05	Pathologic T Status	2.83E-04
hsa-miR-339-3p	Tumor	3.87E-02	Pathologic N Status	-
hsa-miR-375	Normal	4.01E-03	-	5.90E-03
hsa-miR-133b	Normal	8.94E-03	Pathologic T/ N status	4.28E-02
hsa-miR-203a	Normal	2.48E-10	Pathologic N Status/ Clinical M Status	1.87E-09
hsa-miR-150-5p	Tumor	4.51E-02	Pathologic T Status/Clinical N/M Status	-
hsa-miR-493-3p	Tumor	2.36E-02	Pathologic T Status	-
hsa-miR-18a-3p	Tumor	2.42E-04	-	4.19E-03
hsa-miR-130b-3p	Tumor	1.01E-02	-	4.69E-02
hsa-miR-431-5p	Tumor	4.19E-02	Pathologic N Status	-
hsa-miR-7-5p	Tumor	2.93E-02	Pathologic T Stage	-
hsa-miR-382-5p	Tumor	6.24E-06	Pathologic T Status	2.32E-05
hsa-miR-182-3p	Tumor	1.44E-07	Clinical T/N/M Status	7.82E-07
hsa-miR-424-5p	Tumor	2.22E-03	Pathologic T/N/M Stage	1.52E-02
hsa-miR-146b-5p	Tumor	1.07E-03	-	1.05E-02
hsa-miR-196a-3p	Tumor	2.27E-06	Clinical T/N/M Status	2.96E-04
hsa-miR-183-5p	Tumor	2.77E-10	Pathologic T/N/M Stage	2.04E-09
hsa-miR-1-3p	Normal	1.02E-05	-	1.41E-04
hsa-miR-196b-3p	Tumor	3.16E-04	Pathologic T Status	4.83E-03
hsa-miR-497-5p	Normal	7.36E-05	Clinical T/N/M Status	6.93E-04

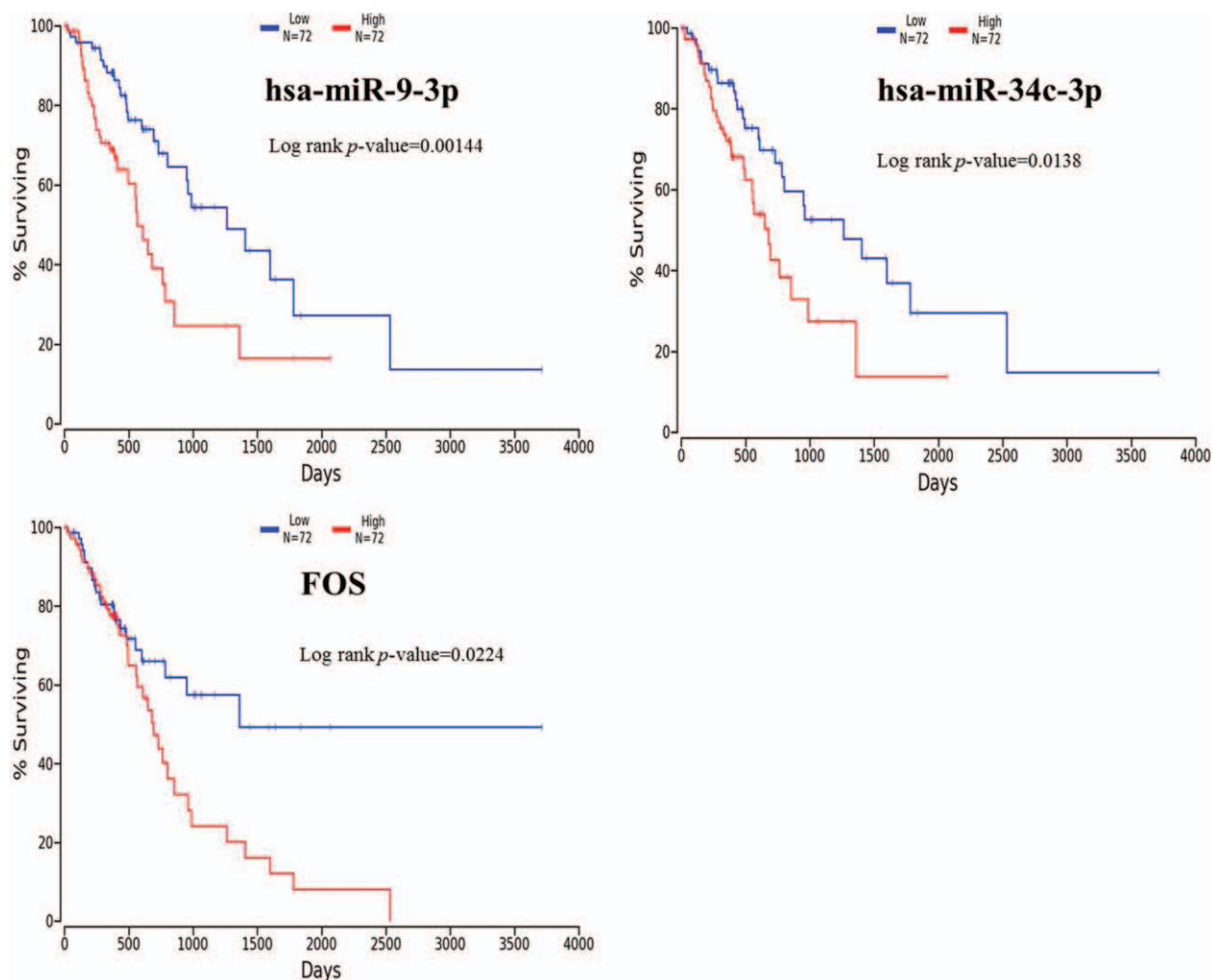


Figure 6. Prognostic values of hsa-miR-9-3p, hsa-miR-34c-3p and FOS for overall survival in ESCC patients. ESCC patients were divided into low- and high-expression groups. N represents the number of patients in each group.

associated with poor OS in patients with ESCC. It was reported that miR-9 induces epithelial-mesenchymal transition (EMT) in ESCC, which is a key event in tumor metastasis. High plasma miR-9 concentrations are significantly correlated with poor tumor differentiation.^[38,56] MiR-34b and miR-34c are located on intron 1 and exon 2 of the common primary transcript. MiR-34b and miR-34c expression levels were significantly higher in ESCC in the corresponding normal samples. Inhibiting the expression of miR-34b or miR-34c in ESCC cells may suppress cell growth in vitro. However, it may inactivate the p53-miR-34 pathway. Taken together, these significant DEMs provide potential biomarkers and molecular mechanisms for the diagnosis and therapy of ESCC.

Global alterations of miRNA and mRNA expression are involved in the regulatory mechanisms of the development and progression of ESCC.^[57] Through PPI analyses, we found that the 13 mRNAs that were assessed as hub genes might play crucial roles in ESCC. Matrix metalloproteinases (MMPs) are strongly expressed in the cytoplasm of cancer cells, especially in the invasive margin, and are weakly expressed in stromal cells. MMP9 expression is positively associated with poor tumor cell differentiation.^[58] MMP1 promotes tumor growth and metastasis when the PI3K/AKT pathway is inhibited by LY294002.^[59] Proliferating Cell Nuclear Antigen (PCNA) expression can serve as an independent prognostic factor of EC.^[60] The c-FOS which belongs to FOS protein family that is involved in the transcriptional regulation of ID1, responds to etoposide in ESCC cells.^[61] Topoisomerase II α (TOP2A) is associated with active cell proliferation of mammalian cells, and have reported expression levels and prognostic value in ESCC patients.^[62] AURKA can directly interacts with β -catenin enhances tumor cell invasion and metastatic in vitro and in vivo, and upregulates MMP-2 expression through activating AKT/NF- κ B pathway in ESCC cells.^[63,64] Mini-chromosome maintenance protein 2 (MCM2) expression is useful marker than Ki-67 in predicting tumor aggressiveness and prognostic value in ESCC.^[65] Involucrin (IVL) IVL and filaggrin (FLG) are major components of the cornified envelope and are considered to be appropriate markers for terminal differentiation.^[66,67] Small proline-rich repeat protein 3 (SPRR3) was downregulated and involved in radiosensitivity in ESCC patients.^[68] Transglutaminase-1 (TGM1) can suppress the Wnt signaling pathway and involve in development of gastric cancer.^[69] By constructing a miRNA-mRNA network, miR-183 was shown to have the highest connectivity to target genes and to target, MMP9, TOP2A, AURKA, MCM2, IVL, FOS, and CYP2C9. MiR-183 might play an oncogenic role by suppressing apoptosis and promoting proliferation in ESCC by regulating PDCD4 expression.^[15,52] Five hub genes could be the target genes of hsa-miR-9-3p and hsa-miR-130b-3p. Moreover, the FOS gene was predicted to be a common target of different miRNAs. FOS and miR-34c had higher expression levels, which were found to be associated with poor OS in patients with ESCC. Interestingly, FOS was a target gene of hsa-miR-34c-3p determined by both miRTarBase and TargetScan. It revealed that miR-34c might be involved in the regulation of FOS. Further experimental studies are required to test our results. We identified that the hub genes that miRNAs may interact with might play important roles in ESCC. In addition, the identification of the miRNA-mRNA regulation network provided new insights into ESCC oncogenesis.

5. Conclusion

In this study, 26 DEMs and 280 DEGs that consisted of 96 upregulated and 184 downregulated genes were identified. Hsa-miR-9-3p, hsa-miR-34c-3p and FOS were found to be related to OS in ESCC patients. The higher expression levels of the transcript were associated with poor OS in patients with ESCC. The present study provides potential biomarkers and molecular mechanisms for the diagnosis and therapy of ESCC and provides valuable clues for further research.

Author contributions

Data curation: Guangcheng Huang.

Formal analysis: Guangcheng Huang, Wenjie Pu.

Funding acquisition: Xiaowu Zhong.

Methodology: Hebin Liao, Chang Liu.

Project administration: Xiaowu Zhong, Xiaolan Guo.

Resources: Lei Xu.

Software: Qiang Ma, Chang Liu, Yan Cai.

Supervision: Yan Cai, Xiaolan Guo.

Writing – original draft: Xiaowu Zhong.

Writing – review & editing: Xiaolan Guo.

References

- Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- Shibata A, Matsuda T, Ajiki W, et al. Trend in incidence of adenocarcinoma of the esophagus in Japan, 1993–2001. *Jpn J Clin Oncol* 2008;38:464–8.
- Lin Y, Totsuka Y, Shan B, et al. Esophageal cancer in high-risk areas of China: Research progress and challenges. *Ann Epidemiol* 2017; 27:215–21.
- Domper Arnal MJ, Ferrández Arenas Á, Lanás Arbeloa Á. Esophageal cancer: risk factors, screening and endoscopic treatment in Western and Eastern countries. *World J Gastroenterol* 2015;21:7933–43.
- Andreou A, Pesthy S, Struecker B, et al. Incidence and risk factors of symptomatic hiatal hernia following resection for gastric and esophageal cancer. *Anticancer Res* 2017;37:7031–6.
- Lv J, Cao XF, Zhu B, et al. Effect of neoadjuvant chemoradiotherapy on prognosis and surgery for esophageal carcinoma. *World J Gastroenterol* 2009;15:4962–8.
- Ho HJ, Chen HS, Hung WH, et al. Survival impact of total resected lymph nodes in esophageal cancer patients with and without neoadjuvant chemoradiation. *Ann Surg Oncol* 2018;25:3820–32.
- Portale G, Hagen JA, Peters JH, et al. Modern 5-year survival of resectable esophageal adenocarcinoma: single institution experience with 263 patients. *J Am Coll Surg* 2006;202:588–96.
- Chen JW, Xie JD, Ling YH, et al. The prognostic effect of perineural invasion in esophageal squamous cell carcinoma. *BMC Cancer* 2014;14:313.
- Saddoughi SA, Reinersman JM, Zhukov YO, et al. Survival after surgical resection of stage iv esophageal cancer. *Ann Thorac Surg* 2017;103:261–6.
- Wu Y, Hu L, Liang Y, et al. Up-regulation of lncRNA CASC9 promotes esophageal squamous cell carcinoma growth by negatively regulating PDCD4 expression through EZH2. *Mol Cancer* 2017;16:150.
- Yang Z, Wu Z, Liu T, et al. Upregulation of PDK1 associates with poor prognosis in esophageal squamous cell carcinoma with facilitating tumorigenicity in vitro. *Med Oncol* 2014;31:337.
- Tie J, Fan D. Big roles of microRNAs in tumorigenesis and tumor development. *Histol Histopathol* 2011;26:1353–61.
- Kinoshita T, Hanazawa T, Nohata N, et al. The functional significance of microRNA-375 in human squamous cell carcinoma: aberrant expression and effects on cancer pathways. *J Hum Genet* 2012;57:556–63.
- Yang M, Liu R, Li X, et al. miRNA-183 suppresses apoptosis and promotes proliferation in esophageal cancer by targeting PDCD4. *Mol Cells* 2014;37:873–80.

- [16] Akagi I, Miyashita M, Ishibashi O, et al. Relationship between altered expression levels of MIR21, MIR143, MIR145, and MIR205 and clinicopathologic features of esophageal squamous cell carcinoma. *Dis Esophagus* 2011;24:523–30.
- [17] Hezova R, Kovarikova A, Srovnal J, et al. Diagnostic and prognostic potential of miR-21, miR-29c, miR-148 and miR-203 in adenocarcinoma and squamous cell carcinoma of esophagus. *Diagn Pathol* 2015;10:42.
- [18] Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res* 2013;41:D991–5.
- [19] Dweep H, Sticht C, Pandey P, et al. miRWalk—database: prediction of possible miRNA binding sites by “walking” the genes of three genomes. *J Biomed Inform* 2011;44:839–47.
- [20] Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protoc* 2009;4:44–57.
- [21] Szklarczyk D, Morris JH, Cook H, et al. The STRING database in 2017: quality-control protein-protein association networks, made broadly accessible. *Nucleic Acids Res* 2017;45(D1):D362–8.
- [22] Chou CH, Shrestha S, Yang CD, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res* 2018;46(D1):D296–302.
- [23] Agarwal V, Bell GW, Nam JW, et al. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 2015;4:e05005.
- [24] 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.
- [25] Anaya J. OncoRank: A pan-cancer method of combining survival correlations and its application to mRNAs, miRNAs, and lncRNAs. *PeerJ Preprints* 2016;4:e2574v1.
- [26] Talukdar FR, di Pietro M, Secrier M, et al. Molecular landscape of esophageal cancer: implications for early detection and personalized therapy. *Ann N Y Acad Sci* 2018;1434:342–59.
- [27] Zhang Y. Epidemiology of esophageal cancer. *World J Gastroenterol* 2013;19:5598–606.
- [28] Liu D, LYN, a key gene from bioinformatics analysis, contributes to development and progression of esophageal adenocarcinoma. *Med Sci Monit Basic Res* 2015;21:253–61.
- [29] Bergheim I, Wolfgarten E, Bollschweiler E, et al. cytochrome P450 levels are altered in patients with esophageal squamous-cell carcinoma. *World J Gastroenterol* 2007;13:997–1002.
- [30] Leng WD, Zeng XT, Chen YJ, et al. Cytochrome P450 2E1 RsaI/PstI polymorphism and risk of esophageal cancer: a meta-analysis of 17 case-control studies. *Exp Ther Med* 2012;4:938–48.
- [31] Schmelzle M, Dizdar L, Matthaehi H, et al. Esophageal cancer proliferation is mediated by cytochrome P450 2C9 (CYP2C9). *Prostaglandins Other Lipid Mediat* 2011;94:25–33.
- [32] Farazi TA, Spitzer JI, Morozov P, et al. miRNAs in human cancer. *J Pathol* 2011;223:102–15.
- [33] Friedman RC, Farh KK, Burge CB, et al. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009;19:92–105.
- [34] Mei LL, Qiu YT, Zhang B, et al. MicroRNAs in esophageal squamous cell carcinoma: potential biomarkers and therapeutic targets. *Cancer Biomark* 2017;19:1–9.
- [35] Zhang W, Lei C, Fan J, et al. miR-18a promotes cell proliferation of esophageal squamous cell carcinoma cells by increasing cyclin D1 via regulating PTEN-PI3K-AKT-mTOR signaling axis. *Biochem Biophys Res Commun* 2016;477:144–9.
- [36] Su H, Jin X, Zhang X, et al. Identification of microRNAs involved in the radioresistance of esophageal cancer cells. *Cell Biol Int* 2014;38:318–25.
- [37] Jiang L, Zhao Z, Zheng L, et al. Downregulation of miR-503 Promotes ESCC Cell Proliferation, Migration, and Invasion by Targeting Cyclin D1. *Genom Proteom Bioinform* 2017;15:208–17.
- [38] Song Y, Li J, Zhu Y, et al. MicroRNA-9 promotes tumor metastasis via repressing E-cadherin in esophageal squamous cell carcinoma. *Oncotarget* 2014;5:11669–80.
- [39] Li S, Qin X, Li Y, et al. MiR-133a suppresses the migration and invasion of esophageal cancer cells by targeting the EMT regulator SOX4. *Am J Transl Res* 2015;7:1390–403.
- [40] Zeng W, Zhu JF, Liu JY, et al. miR-133b inhibits cell proliferation, migration and invasion of esophageal squamous cell carcinoma by targeting EGFR. *Biomed Pharmacother* 2018;111:476–84.
- [41] Harata K, Ishiguro H, Kuwabara Y, et al. MicroRNA-34b has an oncogenic role in esophageal squamous cell carcinoma. *Oncol Lett* 2010;1:685–9.
- [42] Sreedharan L, Mayne GC, Watson DI, et al. MicroRNA profile in neosquamous esophageal mucosa following ablation of Barrett’s esophagus. *World J Gastroenterol* 2017;23:5508–18.
- [43] Hu C, Lv L, Peng J, et al. MicroRNA-375 suppresses esophageal cancer cell growth and invasion by repressing metadherin expression. *Oncol Lett* 2017;13:4769–75.
- [44] Liu Y, Dong Z, Liang J, et al. Methylation-mediated repression of potential tumor suppressor miR-203a and miR-203b contributes to esophageal squamous cell carcinoma development. *Tumour Biol* 2016;37:5621–32.
- [45] Yokobori T, Suzuki S, Tanaka N, et al. MiR-150 is associated with poor prognosis in esophageal squamous cell carcinoma via targeting the EMT inducer ZEB1. *Cancer Sci* 2013;104:48–54.
- [46] Yu T, Cao R, Li S, et al. MiR-130b plays an oncogenic role by repressing PTEN expression in esophageal squamous cell carcinoma cells. *BMC Cancer* 2015;15:29.
- [47] Dong W, Li B, Wang J, et al. Diagnostic and predictive significance of serum microRNA-7 in esophageal squamous cell carcinoma. *Oncol Rep* 2016;35:1449–56.
- [48] Qi B, Lu JG, Yao WJ, et al. Downregulation of microRNA-382 is associated with poor outcome of esophageal squamous cell carcinoma. *World J Gastroenterol* 2015;21:6884–91.
- [49] Wang F, Wang J, Yang X, et al. MiR-424-5p participates in esophageal squamous cell carcinoma invasion and metastasis via SMAD7 pathway mediated EMT. *Diagn Pathol* 2016;11:88.
- [50] Li J, Shan F, Xiong G, et al. Transcriptional regulation of miR-146b by C/EBP (LAP2) in esophageal cancer cells. *Biochem Biophys Res Commun* 2014;446:267–71.
- [51] Ma Y, Wang B, Guo Y, et al. Inhibition of miR-196a affects esophageal cancer cell growth in vitro. *Biomed Pharmacother* 2016;84:22–7.
- [52] Ren LH, Chen WX, Li S, et al. MicroRNA-183 promotes proliferation and invasion in oesophageal squamous cell carcinoma by targeting programmed cell death 4. *Br J Cancer* 2014;111:2003–13.
- [53] Liu W, Li M, Chen X, et al. MicroRNA-1 suppresses proliferation, migration and invasion by targeting Notch2 in esophageal squamous cell carcinoma. *Sci Rep* 2018;8:5183.
- [54] Osako Y, Seki N, Koshizuka K, et al. Regulation of SPOCK1 by dual strands of pre-miR-150 inhibit cancer cell migration and invasion in esophageal squamous cell carcinoma. *J Hum Genet* 2017;62:935–44.
- [55] Sun J, Wang D, Li X, et al. Targeting of miR-150 on Gli1 gene to inhibit proliferation and cell cycle of esophageal carcinoma EC9706. *Cancer Biomark* 2017;21:203–10.
- [56] Cui Y, Xue Y, Dong S, et al. Plasma microRNA-9 as a diagnostic and prognostic biomarker in patients with esophageal squamous cell carcinoma. *J Int Med Res* 2017;45:1310–7.
- [57] Liu M, Wang C, Gao L, et al. A nomogram to predict long-time survival for patients with M1 diseases of esophageal cancer. *J Cancer* 2018;9:3986–90.
- [58] Gu ZD, Li JY, Li M, et al. Matrix metalloproteinases expression correlates with survival in patients with esophageal squamous cell carcinoma. *Am J Gastroenterol* 2005;100:1835–43.
- [59] Liu M, Hu Y, Zhang MF, et al. MMP1 promotes tumor growth and metastasis in esophageal squamous cell carcinoma. *Cancer Lett* 2016;377:97–104.
- [60] Zhu SC, Li R, Wang YX, et al. Impact of simultaneous assay, the PCNA, cyclinD1, and DNA content with specimens before and after preoperative radiotherapy on prognosis of esophageal cancer-possible incorporation into clinical TNM staging system. *World J Gastroenterol* 2005;11:3823–9.
- [61] Zhao Y, Luo A, Li S, et al. Inhibitor of Differentiation/DNA Binding 1 (ID1) Inhibits Etoposide-induced Apoptosis in a c-Jun/c-Fos-dependent Manner. *J Biol Chem* 2016;291:6831–42.
- [62] Hanagiri T, Ono K, Kuwata T, et al. Evaluation of topoisomerase I/topoisomerase II alpha status in esophageal cancer. *J UOEH* 2011;3:205–16.
- [63] Jin S, Wang X, Tong T, et al. Aurora-A enhances malignant development of esophageal squamous cell carcinoma (ESCC) by phosphorylating (-catenin). *Mol Oncol* 2015;9:249–59.
- [64] Wang X, Li X, Li C, et al. Aurora-A modulates MMP-2 expression via AKT/NF- κ B pathway in esophageal squamous cell carcinoma cells. *Acta Biochim Biophys Sin* 2016;48:520–7.
- [65] Kato H, Miyazaki T, Fukai Y, et al. A new proliferation marker, minichromosome maintenance protein 2, is associated with tumor aggressiveness in esophageal squamous cell carcinoma. *J Surg Oncol* 2003;84:24–30.

- [66] Rice RH, Green H. Presence in human epidermal cells of a soluble protein precursor of the cross-linked envelope: activation of the cross-linking by calcium ions. *Cell* 1979;18:681–94.
- [67] Sandilands A, Sutherland C, Irvine AD, et al. Filaggrin in the frontline: role in skin barrier function and disease. *J Cell Sci* 2009;122:1285–94.
- [68] Luo A, Chen H, Ding F, et al. Small proline-rich repeat protein 3 enhances the sensitivity of esophageal cancer cells in response to DNA damage-induced apoptosis. *Mol Oncol* 2013;7:955–67.
- [69] Huang H, Chen Z, Ni X. Tissue transglutaminase-1 promotes stemness and chemoresistance in gastric cancer cells by regulating Wnt/ β -catenin signaling. *Exp Biol Med (Maywood)* 2017;242:194–202.