Identification of microRNAs as novel biomarkers for esophageal squamous cell carcinoma: a study based on The Cancer Genome Atlas (TCGA) and bioinformatics

Cheng-Yun Li¹, Wen-Wen Zhang¹, Ji-Lian Xiang², Xing-Hua Wang³, Jin Li¹, Jun-Ling Wang¹

¹Department of Toxicology, School of Public Health, Lanzhou University, Lanzhou, Gansu 730000, China; ²Department of Gastroenterology, Third People’s Hospital of Gansu Province, Lanzhou, Gansu 730000, China; ³Department of Gastrointestinal Surgery, Gansu Wuwei Tumor Hospital, Wuwei, Gansu 733000, China.

Abstract

Background: MicroRNAs (miRNAs) have played important roles in the regulation of gene expression in many cancers, but their roles in esophageal squamous cell carcinoma (ESCC) are still unclear. The aim of this study was to determine the potential ESCC-specific key miRNAs from a large sample dataset in The Cancer Genome Atlas (TCGA).

Methods: Integrative bioinformatics analysis was used to identify key ESCC-specific miRNAs related to the ESCC patients’ tumor histological grade and lymphatic metastasis from TCGA. Next, these key miRNA potential gene regulatory functions and relationships with ESCC patients’ clinical characteristics and overall survival were analyzed. Finally, three key miRNAs were selected randomly and quantification real-time polymerase chain reaction (qRT-PCR) was used to validate in 51 newly diagnosed ESCC patients’ tissues samples (collected from Nov. 2017 to Feb. 2019, in Wuwei, China) whether the bioinformatics analyses results were reliable and valid. Two-tailed Student’s t test, Pearson Chi-squared test and Kaplan-Meier survival analysis were used in this study.

Results: Thirty-five ESCC-specific miRNAs from TCGA database were investigated (fold-change > 2.0, P < 0.05), and 28 participated in the miRNAs-mRNAs co-expression network construction, while 17 were related with ESCC patients’ tumor histological grade, TNM stage, and lymphatic metastasis (P < 0.05). Meanwhile, six miRNAs (including miR-200b-3p, miR-31-5p, miR-135b-5p, miR-141-3p, miR-133b-5p, and miR-195-5p) were correlated with overall survival of ESCC patients (log-rank, P < 0.05). MiR-135b-5p, miR-15b-5p, and miR-195-5p were selected for verification of the expression levels in 51 ESCC patients’ tissue samples by using qRT-PCR. We found that the fold-changes between qRT-PCR and TCGA were completely consistent. The results also suggested that miR-135b-5p, miR-15b-5p, and miR-195-5p were significantly correlated with tumor differentiation degrees (P < 0.05), miR-195-5p was significantly correlated with tumor TNM stage (P < 0.05), and miR-133b-5p was significantly correlated with lymph-node metastasis (P < 0.05), MiR-135b-5p, miR-15b-5p, and miR-195-5p expression levels, ESCC patient clinical features association analysis results and the aforementioned TCGA bioinformatics analyses were similar.

Conclusion: This study identified key ESCC-related miRNAs. The key miRNAs are worthy of further investigation as potential novel biomarkers for diagnosis, classification, and prognosis of ESCC.

Keywords: Esophageal cancer; microRNA; Clinical features; Overall survival

Introduction

Esophageal squamous cell carcinoma (ESCC) is a common fatal malignant disease in humans, which causes more than 400,000 deaths per year.¹ In spite of the advances in multidisciplinary diagnosis and treatment, there are still important topics that need to be explored for the early diagnosis and prognosis biomarkers of ESCC.² The ESCC detection methods are based on the endoscopy examination, while this method was often affected by doctors’ individual experience and technology, with a high rate of missed diagnosis. Meanwhile, endoscopic diagnoses of ESCC cannot find the early stage of lymph node metastasis status in esophageal carcinoma.³ In addition, conventional non-invasive serological biomarkers, such as carbohydrate antigen 19-9 (CA19-9) and carcinoembryonic antigen (CEA), are lack of specificity in the early diagnosis of ESCC.⁴ Therefore, the significance of exploration of innovative and favorable biomarkers for ESCC diagnosis and prognosis should be emphasized.

Access this article online

Quick Response Code:
Website:
www.cmj.org
DOI:
10.1097/CM9.0000000000000427

Correspondence to: Dr. Jun-Ling Wang, Department of Toxicology, School of Public Health, Lanzhou University, Lanzhou, Gansu 730000, China
E-Mail: 13919995713@163.com

Received: 23-04-2019 Edited by: Yi Cui

2213
MicroRNAs (miRNAs, 18–22 nucleotides) are small, non-coding RNAs that can bind to the messenger RNAs and negatively regulate the gene expression. Studies have revealed that miRNAs function in many key biological roles in various disease progressions, including cancers. Therefore, further insights into the important roles of miRNAs in diseases, especially in cancers, have shown miRNAs as attractive tools and targets for novel diagnoses, prognoses, and therapeutic approaches. ESCC is a multistep-process disease, which involves multiple effects of gene-environmental factor interactions. The miRNAs play important regulatory roles in epigenetics and are closely related with many diseases progressions.

Further understanding of the ESCC tissue alterations and miRNA expression profiles may help us to identify better biomarkers for ESCC diagnosis, classification, and prognosis.

Further evidence suggested that aberrantly expressed miRNAs in ESCC tissues have a clinical influence on disease diagnosis and prognosis. Recently, miRNA-associated biomarkers have been reported for the diagnosis and prognoses of ESCC. However, the small sample sizes used for microarray detection often present a bias toward the identification of ESCC-associated miRNAs due to lack of RNA sequencing data, and therefore often generate errors. Thus far, The Cancer Genome Atlas (TCGA) database has provided a resource for large sample size genome sequencing data from 33 human cancer types including ESCC. In the present study, we obtained ESCC patients’ tissues samples including normal esophageal mucosa tissues 11 cases, pathologic Stage I 27 cases, Stage II 74 cases and Stage III-IV 60 cases. Before further analyses and processing, RNA-sequencing data normalization was performed in full compliance with the publication guidelines of TCGA; the approval of ethics committee was not required.

Methods

Ethical approval
Informed consent forms and clinical information were obtained for all patients by the investigator and from medical records, respectively. The procedures used in this study were approved by the Wuwei Tumor Hospital of Gansu (Wuwei, China) and conformed to the Declaration of Helsinki and current legislation.

Data collection
The ESCC patient tumor tissues and normal esophageal epithelial tissue miRNA sequencing data set and patient personal information and clinical pathological features are included from the TCGA database (supplementary Table S1, http://links.lww.com/CM9/A89). RNA sequencing data were obtained from Illumina HiSeq platforms (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga/using-tcga/types). ESCC patient exclusion criteria included: (1) the admitting diagnosis was not ESCC; (2) suffering from two or more types of cancers including ESCC; (3) clinical information or RNA sequencing data were deficient; (4) patients treated with radiation and chemotherapy before specimen collection, and (5) overall survival time more than 5 years. In the present study, we obtained ESCC patients’ tissues samples including normal esophageal mucosa tissues 11 cases, pathologic Stage I 27 cases, Stage II 74 cases and Stage III-IV 60 cases. Further analyses and processing, RNA-sequencing data normalization was performed in full compliance with the publication guidelines of TCGA; the approval of ethics committee was not required.

Study design and data preprocessing
Hierarchical cluster analyses were performed using R package heat maps. Then, according to the z-score values to deduce the expression levels of samples RNA sequencing data, we analyzed the gene pair Euclidean distances to detect gene clusters. Subsequently, the significantly different miRNA and mRNA expression profiles in ESCC were analyzed separately, and we determined the significantly different genes of the normal vs. cancerous tissues of ESCC patients. We compared the differentially expressed miRNAs and mRNA of four stage levels, including ESCC (G I stage) non-lymphatic metastases vs. non-tumor esophageal tissues, ESCC (G II stage) lymphatic metastases vs. non-tumor esophageal tissues, ESCC (G II stage) lymphatic metastases vs. non-tumor esophageal tissues, and ESCC (G III-IV stage) lymphatic metastases vs. non-tumor esophageal tissues, respectively. The intersections of the miRNAs and mRNAs were selected in the above four stage levels for further analyses using bioinformatics methods. Figure 1 is the flow diagram of this study design.

Co-expression network of miRNAs and mRNAs
For the current study, according to the findings that miRNAs can bind to the messenger RNAs and negatively regulate gene expression, a miRNA-mRNA network was built. Significantly differentially expressed intersecting miRNA and mRNA were used to build a co-expression network, which showed fold change >2.0 and P < 0.05. Target genes of miRNAs were predicted by miRanda (http://www.microrna.org/microrna/home.do) and Targetscan (http://www.targetscan.org/). The combined miRNA predicted target gene information and significantly differentially expressed intersecting miRNAs of TCGA ESCC patients were used to build the miRNA-mRNA network.

The miRNAs-mRNAs co-expression network was constructed using significantly differential gene expression profiles using R package software (The University of Auckland, Auckland, New Zealand). Pearson correlation matrices models were used to find the potential relevance of all pair-wise genes. Subsequently, based on the weighted adjacency matrix, we measured the network connectivity of a gene with other gene network generations, and built the miRNAs-mRNAs co-expression network by using Cytoscape 3.0 software (National Institute of General Medical Sciences, MD, USA).

Analysis of the association between key miRNAs and clinical status
According to the co-expression of miRNAs and mRNAs, the key miRNAs involved in co-expression networks were selected as our targets to determine if they positively correlated with
ESCC. Subsequently, we further analyzed the potential associations between co-expression network key miRNAs and TCGA ESCC patients’ clinical features, including sex, TNM stage, tumor grade, lymphatic metastases status, and pathological stage by using multiple linear regression analyses. Kaplan-Meier survival analyses
To analyze the correlation between the changes of the selected key miRNA expression levels and ESCC patients’ prognostic survival, correlations were determined by Gene Expression Profiling Interactive Analysis (GEPIA) (http://gepia.cancer-pku.cn/). Publicly available data and GEPIA tools were prepared to calculate the Kaplan-Meier survival analysis parameters using TCGA ESCC patients’ datasets. All the RNA sequencing data probe sets or records were averaged by the quantile normalized data. Finally, the survival distributions equality of different TCGA ESCC patients and the key miRNA expression level changes were analyzed using the Kaplan-Meier, log-rank, and the hazard ratio (HR) tests.

Preparation for human ESCC samples
Next, 51 ESCC patients’ tumor tissues samples and paired non-tumor esophagus tissues (aged 40–69 years) were collected from Gansu Wuwei Tumor Hospital, China, for RT-qPCR validation. All of these patients were diagnosed with ESCC according to their histopathology. Samples were collected and stored in RNAlater (Ambion, Foster City, CA, USA) at −80°C. Collection and use of ESCC patients’ tumor tissue samples were approved by the ethics committee of Gansu Wuwei Tumor Hospital.

Quantitative real time PCR (qRT-PCR) validation of bioinformatics data
Total RNA of ESCC patients’ tissues samples were isolated by TRIzol® reagent (Invitrogen, Carlsbad, CA, USA). Total RNA quantity was evaluated by a NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). A Reverse Transcription Kit (Promega, Madison, WI, USA) and GoTaq® qPCR Master Mix of Power SYBR® Green (Promega) were used to synthesized cDNA and for qRT-PCR detection. Finally, qRT-PCR reactions were performed using the Step One PlusTM PCR System (Applied Biosystems, Foster City, CA, USA).

Statistical analyses
Data were analyzed using the SPSS Statistics 21.0 (International Business Machines Corporation, Armonk, NY, USA) and expressed as mean ± standard deviation (SD). All analyses were performed three times and represent data from three individual experiments. Two-tailed Student’s t test was used to significance of differences between subgroups. Pearson Chi-squared test was used to find the association between the expression of miRNAs and ESCC patients’ clinicopathological characteristics. Kaplan-Meier survival analysis was used to investigate the correlation between the changes of miRNAs expression levels and patients’ prognostic overall survival times. Statistical significance was set at P < 0.05. The qRT-PCR results were calculated and evaluated by the 2−ΔΔCt method.

Results
Differentially expressed miRNAs screening in ESCC patients
We used ESCC patients’ RNA-seq datasets collected from 161 cancerous and 11 normal specimens for further analyses. There were 232 miRNAs that were significantly differentially expressed between ESCC patients’ tumor vs. normal esophageal tissues samples in the Level 3 ESCC RNA-Sequencing (RNA-Seq) data from TCGA database. Next, we integrated and analyzed the relevance between the 232 miRNA expression levels and ESCC patients’ differential clinical stage and pathologic grade. There were 103 miRNAs that showed significant differential changes between G1 stage (non-lymphatic metastasis) ESCC tumors and normal esophageal tissues samples. In addition, 119 miRNAs were significantly differentially changed between G1 stage (lymphatic metastasis) ESCC tumors and normal esophageal
tissues samples; 125 miRNAs were found to be significantly differentially changed between G II stage (lymphatic metastasis) ESCC tumors and normal esophageal tissue samples; and 150 miRNAs were found to be significantly differentially changed between G III-IV stage (lymphatic metastasis) ESCC tumors and normal esophageal tissues samples. Subsequently, we used Venn diagrams (http://bioinfogp.cnb.csic.es/tools/venny/index.html) to analyze the overlapping subclasses of miRNAs, and found there were 35 key intersecting miRNAs [Figure 2]. These 35 intersecting miRNAs included 25 miRNAs that were overexpressed and 10 miRNAs that were downregulated (supplementary Table S2, http://links.lww.com/CM9/A89).

**MiRNAs-mRNAs co-expression network construction and potential regulatory function analyses**

We chose the abovementioned 35 intersecting miRNAs from 232 ESCC associated miRNAs by using a Venn diagram overlapping subclass analysis to construct the miRNAs-miRNAs network. Next, to build the miRNA-mRNA co-expression network, we also predicted the mRNAs targeted by the miRNAs. According to these 35 intersecting miRNAs (supplementary Table S2, http://links.lww.com/CM9/A89), we predicted that these miRNAs targeted mRNAs by using the miRanda and Targetscan databases. Based on the bioinformatics analysis, we found that there were 376 intersecting mRNAs and 2235 significantly differentially expressed mRNAs in ESCC cancer tissues from TCGA database. Finally, based on the 35 intersecting miRNAs and 376 intersecting mRNAs, we constructed a miRNAs-mRNAs co-expression network. The relationships between miRNAs and mRNAs were analyzed by their negative regulation mechanisms and the co-expression network was drawn using Cytoscape 3.0. There were 28 targeted miRNAs that potentially regulated 72 mRNAs and were involved in the miRNA-mRNA co-expression network [Figure 3]. To further analyze these 72 negatively regulated mRNAs in the
co-expression network, we found that some played important roles in cancer development, such as E2F2, FGF2, COL4A3, ERBB4, FGFRI, RELN, and TCF3.

To obtain further insight into the potential regulatory functions of the 28 miRNAs on the 72 mRNAs involved in miRNAs-mRNAs networks, the up- and down-regulated miRNAs were analyzed using the DAVID database. Gene Ontology (GO) analyses revealed upregulated miRNAs enriched in ‘regulation of cellular process, regulation of biological process and protein binding’ and downregulated miRNAs transcripts enriched in ‘single-organism cellular process, single-organism process, and response to stimulus’ [Figure 4A]. These significantly enriched GO functions provided a better understanding of the roles of miRNAs in regulatory functions. In addition, pathway analysis was performed for all 72 mRNAs, and the results suggested that the enriched pathways targeted by upregulated mRNAs transcripts were ‘microRNAs in cancer, human T-cell leukaemia virus type I (HTLV-I) infection, cell cycle and hepatitis B.’ The enriched pathways of downregulated transcripts were ‘insulin secretion, glutamatergic synapse, proteoglycans in cancer, and vascular smooth muscle contraction’ [Figure 4B].

Correlations between miRNA signatures and clinical characteristics

We selected the 28 miRNAs, which were involved in miRNA-mRNA co-expression networks, and further analyzed the correlation between these miRNAs and TCGA database of 161 ESCC patients’ clinical pathology features. The ESCC clinical/demographic features included patients’ sex, race, tumor grade, TNM stage, and lymphatic metastasis taken from TCGA database. Multiple linear regression analysis results suggested that there were 17 miRNAs differentially expressed with different clinical pathology features (P < 0.05). We showed that miR-32-5p, miR-330-5p, miR-135b-5p, miR-195-5p, and miR-145-5p were associated with ESCC patients’ tumor grade, miR-16-5p, miR-32-5p, miR-93-5p, miR-185-5p, miR-15b-5p, miR-135b-5p, miR133a-3p, miR-28-5p, and miR-195-5p were associated with TMN stages, and miR-32-5p, miR-93-5p, miR-15b-5p, miR-135b-5p, miR-330-5p, miR-106b-5p, miR-182-5p, miR-20b-5p, miR-141-3p, miR-28-5p, miR-195-5p, and miR-320a were associated with lymphatic metastasis. Meanwhile, miR-16-5p and miR-183-5p were found associated with racial/ethnic background [Table 1].

Prognostic analysis of miRNA expression levels and ESCC patient survival

A Kaplan-Meier survival analysis was used to investigate the relationships between the miRNAs-mRNAs co-expression network 28 key miRNAs and ESCC patients’ survival time. The 28 key miRNA expression levels and the ESCC patient overall survival information in

![Figure 4: Top 20 enrichments of Gene Ontology (GO) terms for differentially expressed miRNAs and pathways for differentially expressed intersection miRNAs. (A) Top 20 enrichments of GO terms for differentially expressed miRNAs (the bar plot shows the enrichment scores of the significant enrichment GO terms). (B) Top 20 enrichments of pathways for differentially expressed intersection miRNAs (the bar plot shows the enrichment scores of the significant enrichment pathways).](image-url)
TCGA database were calculated using univariate Cox proportional hazard regression model analyses. We found six miRNAs closely related with ESCC patient overall survival time (log-rank, \(P < 0.05\)). Among these six miRNAs, there were five miRNAs (miR-200b-3p, miR-31-5p, miR-15b-5p, miR-141-3p, and miR-135b-5p) that were negatively related with ESCC patient overall survival time (\(P < 0.05\)), and miR-195-5p was positively correlated with ESCC patient overall survival time (\(P < 0.05\)) [Figure 5].

**Verification of selected miRNAs expression levels in ESCC tissues by qRT-PCR**

Based on the above bioinformatics analyses, three miRNAs (miR-135b-5p, miR-15b-5p, and miR-195-5p) were randomly selected. The three key miRNA expression levels in newly diagnosed 51 ESCC patients’ tumor samples and their self-paired normal esophagus epithelial tissues samples were compared to previous TCGA bioinformatics analysis results for reliability and validity. The expression levels of miR-135b-5p and miR-15b-5p were consistently upregulated in 51 ESCC patient tumor tissue samples compared with normal esophageal epithelial tissues (\(P < 0.05\)); however, miR-195-5p expression levels were consistently downregulated (\(P < 0.05\)) [Table 2]. Subsequently, we also compared the increased or decreased fold-changes between these three miRNAs verified by qRT-PCR and TCGA ESCC RNA fold-changes. The fold-changes measured by qRT-PCR of the 51 collected ESCC patients’ tissue samples and TCGA ESCC RNA sequencing results were consistent (supplementary Table S2, http://links.lww.com/CM9/A89 and Table 2).
Table 2: Relative expression of miRNAs in 51 pairs of ESCC tumor and non-tumor esophagus tissues (mean ± SD).

| Gene symbol | miRNA expression, ΔCt | Tumor tissues | Adjacent non-tumor tissues | 2^−ΔCt Change in expression in ESCC | t | P |
|-------------|----------------------|---------------|---------------------------|------------------------------------|---|---|
| miR-135b-5p |                      | 14.017 ± 2.854 | 15.551 ± 2.990            | ▲ 14.850 ± 7.21                      | 3.026 | 0.000 |
| miR-15b-5p  |                      | 15.412 ± 1.393 | 17.299 ± 2.088            | ▲ 11.199 ± 6.35                      | 2.998 | 0.017 |
| miR-195-5p  |                      | 17.124 ± 2.145 | 15.108 ± 2.377            | △ −8.329 ± 4.88                      | 2.331 | 0.029 |

ΔCt = Ct target gene – Ct U6; ΔΔCt = ΔCt tumor tissues – ΔCt Adjacent non-tumor tissues; ▲ represent higher expressed in tumor tissues; △ represent lower expressed in tumor tissues; ESCC: esophageal squamous cell carcinoma.

Table 3: Association between the expression of miR-135b-5p, miR-15b-5p, miR-195-5p and clinicopathological characteristics of ESCC patients.

| Variables         | Cases, n (%) | miR-135b-5p Low, n High, n | χ² | P | miR-15b-5p Low, n High, n | χ² | P | miR-195-5p Low, n High, n | χ² | P |
|-------------------|--------------|-----------------------------|----|---|---------------------------|----|---|---------------------------|----|---|
| Gender            |              | 0.000 1.000                 | 3.975 | 0.046^* | 0.942 0.332 |
| Male              | 31 (61)      | 7 24                        | 6 25 | 24 | 7                         |
| Female            | 20 (39)      | 4 16                        | 10 10 | 13 | 7                         |
| Age               |              | 0.362 0.547                 | 1.377 0.241 | 3.557 0.059 |
| ≤50 years         | 17 (30)      | 5 12                        | 3 14 | 9 | 8                         |
| >50 years         | 34 (70)      | 6 28                        | 13 21 | 28 | 6                         |
| Degree of differentiation | 6.642 0.010^* | 4.438 0.034^* | 5.460 0.019^* |
| Poorly            | 33 (66)      | 3 30                        | 7 26 | 28 | 5                         |
| TNM stage         |              | 0.269 0.604                 | 1.634 0.201 | 6.297 0.012^* |
| I/II              | 22 (43)      | 6 16                        | 9 13 | 12 | 10                        |
| III/IV            | 29 (57)      | 5 24                        | 7 22 | 25 | 4                         |
| Lymph-node status |              | 3.877 0.049^*               | 0.000 1.000 | 0.278 0.398 |
| No metastasis     | 7 (14)       | 4 3                         | 2 5  | 4 | 3                         |
| Metastasis        | 44 (86)      | 7 37                        | 14 30 | 33 | 11                        |

^* P < 0.05. ESCC: esophageal squamous cell carcinoma. High: represent high expressed; Low: represent low expressed; TNM: tumor node metastasis.

Relationships between the three selected miRNAs and ESCC clinical pathological features

We found that miR-135b-5p, miR-15b-5p, and miR-195-5p expression levels were significantly related with the 51 ESCC patients tumor differentiation degrees (P < 0.05) [Table 3]; miR-195-5p was significantly related with tumor TNM stage (P < 0.05), and miR-135b-5p was significantly related with lymph-node metastasis (P < 0.05). The differential expression of miR-15b-5p was significantly associated with the 51 ESCC patients’ sex (P < 0.05). These three miRNAs from ESCC patients’ correlated with the above TCGA bioinformatics comprehensive analysis [Table 1] and were in good agreement, suggesting that TCGA bioinformatics analysis was reliable.

Discussion

ESCC is the main type of esophageal cancer and is a common cause of cancer-related death worldwide.[11,12] ESCC development and progression is a complex process, which includes a variety of obvious changes in gene expression levels and subsequent physiological changes. MiRNAs are small, non-coding regulatory RNAs that inhibit target miRNAs.[13] Recent advances in miRNA expression profiling have suggested new directions to study the pathological changes of ESCC, and also indicate that miRNAs have great potential as novel biomarkers in ESCC diagnosis and prognosis.[14,15] Recent studies mainly concentrated on small sample size RNA sequencing or microarray data detection and, therefore, may not accurately reflect the dysregulated miRNAs that link to tumorigenesis and progression of ESCC.[16,17] Therefore, the ESCC related significant differences of miRNAs should be identified in a large sample size to enhance accuracy and reliability of predictive value.

In this study, we analyzed the dysregulated miRNAs in ESCC from TCGA database. According to the ESCC patient tissue sample miRNA sequencing data sets in TCGA, we further explored the significant differences of miRNAs related to patients’ clinical features and overall survival time by using integrated bioinformatics analysis. Three miRNAs were selected randomly and we detected their expression levels in 51 newly collected ESCC patients’ tumor samples and their self-paired normal esophageal epithelial tissue samples by qRT-PCR. We further analyzed the correlations between the three miRNAs and ESCC patients’ clinical pathological characteristics to confirm the accuracy and reliability of the bioinformatics analysis results.
According to the integrated bioinformatics overlapping subclasses analysis results we found 35 dysregulated miRNAs in ESCC tumor tissues and significant correlation between ESCC histological grades and lymphatic metastasis status. Among these 35 dysregulated miRNAs, some have been detected that showed significant differences expressed in esophageal cancer, including upregulated miR-205, which may act as an oncogene in ESCC.\(^{[18]}\) Alder et al\(^{[19]}\) reported that miR-31 and miR-21 were dysregulated in ESCC and their expression levels were associated with zinc deficiency promoted esophageal cancer. Gu et al\(^{[20]}\) investigated the miRNA diagnostic values in ESCC pathogenesis and prognosis, and found that in ESCC upregulation of miR-93, miR-106, and miR-205 were closely related with ESCC pathogenesis and prognosis. In addition, upregulated miR-20b-3p, miR-93-5p, miR-16-5p, miR-330-5p, miR-106b-5p, and miR-484 were reported in esophageal cancer and were involved in regulating the progression of this disease.\(^{[21-23]}\) The downregulation of miR-133a expression showed changes from Barrett esophagus to esophageal cancer.\(^{[24]}\) Derouet et al reported that miR-145 expression affected the progression of esophageal cancer by enhancing cell invasion and apoptosis.\(^{[25]}\) MiR-143-3p, miR-125a-5p, and miR-101-3p levels also were reported to be significantly different in esophageal cancer tissues and cell lines.\(^{[20,26-28]}\) For the remaining 16 dysregulated miRNAs, including 10 upregulated and six downregulated miRNAs, their functions in the development and progression of ESCC have not been reported.

Based on the theory that mRNA transcripts bind by single-stranded miRNAs,\(^{[29]}\) we constructed a miRNAs-miRNAs network according to their negative regulatory relationships. Among these ESCC related 35 intersecting miRNAs, there were 28 key miRNAs involved in the miRNAs-miRNAs co-expression network. The miRNAs-miRNAs co-expression network revealed potential regulatory relationships between miRNAs and mRNAs in ESCC. In the miRNAs-miRNAs co-expression network, we also found some of these 28 key miRNAs were reported as potential diagnostic and prognostic biomarkers in ESCC, including miR-101-3p, miR-125a-5p, miR-16-5p, and miR-20b-5p.\(^{[22,30-32]}\) To further analyze these negatively regulated 72 mRNAs in the co-expression network, we showed that some played crucial roles in the development and progression of ESCC, such as E2F2, FGF2, and TCF3.\(^{[31-33]}\) Subsequently, we analyzed these 28 key miRNAs in the miRNAs-miRNAs co-expression network indirectly by GO enrichment analyses and they were involved in cell signaling pathway regulation. GO analysis results suggested that the enriched GOs mainly targeted cell molecules and biological processes. The pathway analysis also revealed that some pathways were related with cancer, such as the PI3K-AKT signaling pathway and MAPK signaling pathway.\(^{[34,37]}\) Ishibashi et al also reported that the overexpressed miR-141-3p targeted PHLPP-2 gene expression and affected the activity of the PI3K/AKT pathway in ESCC.\(^{[38]}\) Therefore, our bioinformatics analysis results suggested that these 28 key miRNAs may play crucial roles in the progression of ESCC and cancer pathways.

To investigate the correlation between the 28 key miRNA expression levels and the 161 ESCC patients’ clinical pathology features, we identified 17 miRNAs that were related with 161 ESCC patients’ clinical pathology features, including miR-32-3p, miR-16-5p, miR-32-5p, and miR-183-5p. These miRNAs mainly related with ESCC patients’ tumor grade, TNM stages, and lymphatic metastasis status. Among these 17 miRNAs, miR-16-5p, miR-330-5p, miR-145-5p, miR-93-5p, miR-93-5p, miR133a-3p, miR-28-5p, miR-15b-5p, miR-20b-5p, miR-141-3p, and miR-320a were previously reported to be important indicators in lymphatic metastasis, invasion, and as diagnostic biomarkers of ESCC.\(^{[22,32,39-43]}\) However, the functions of other miRNAs have not been correlated with ESCC clinical pathology features. We further analyzed the relationships between the 28 key miRNAs and ESCC patients’ overall survival time. There were six miRNAs that correlated with ESCC patients’ overall survival in TCGA. Among these six miRNAs, only miR-15b-5p has been reported with a correlation to ESCC survival.\(^{[22]}\) and other miRNAs have not been studied. Our bioinformatics analyses revealed the potential of using miRNAs as biomarkers for diagnosis, classification, and prognosis of ESCC.

Finally, miR-135b-5p, miR-15b-5p, and miR-195-5p were selected and validated using TCGA bioinformatics analysis results from 51 newly collected ESCC patients’ tissue samples. The correlation analyses between these three miRNA expression levels and the ESCC patients’ clinical pathology features also were investigated. The results suggested that the ESCC patients’ clinical pathology correlated with TCGA bioinformatics analyses results.

In summary, the present study successfully identified the correlation of ESCC-specific miRNAs from a large-scale sample in TCGA database by using integrated bioinformatics analyses. We investigated the ESCC related specific miRNAs in different clinical pathology features and patients’ overall survival time. Finally, we validate the reliability and validity of TCGA database ESCC patients’ RNA sequencing integrated bioinformatics results by qRT-PCR. We found that miRNAs (miR-16-5p, miR-183-5p, miR-32-5p, miR-330-5p, miR-135b-5p, miR-195-5p, miR-145-5p, miR-93-5p, miR-185-5p, miR133a-3p, miR-28-5p, miR-15b-5p, miR-106b-5p, miR-182-5p, miR-20b-5p, miR-141-3p, and miR-320a) and (miR-200b-3p, miR-31-5p, miR-15b-5p, miR-141-3p, miR-135b-5p and miR-195-5p) were related with ESCC patients’ clinical pathology features and patients’ overall survival time, respectively. Taken together, above TCGA database ESCC patients’ clinical pathology features and overall survival time related miRNAs are worthy of further study about their application functions as biomarkers indicators in diagnosis, clinical pathology classification, and prognosis estimation of ESCC patients.

Acknowledgements

We thank Dong-Lin Cheng who supported the bioinformatics analysis technical assistance in the project.
This study was supported by grants from the Fundamental Research Funds for the Central Universities (No. IZUJKBY-2018-13), the National Natural Science Foundation of China (No. 81803188), and the Gansu Province Science and Technology Project (No. 1606RJYA270).

Conflicts of interest
None.

References
1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide: IARC Cancer Base No. 11. Int J Cancer 2018;144:564-76. doi: 10.1002/ijc.31387.
2. Wang B, Liang T, Li J. Long noncoding RNA LINC01296 is associated with poor prognosis in ESCC and promotes ESCC cell proliferation, migration and invasion. Eur Rev Med Pharmacol Sci 2018;22:4524-4531. doi: 10.26355/eurrev_201807_15507.
3. Wang JW, Guan CT, Wang LL, Chang LY, Hao CQ, Li BY et al. Natural history analysis of 101 severe dysplasia and esophageal cancer cases by endoscopy. Gastrointest Res Pract 2017;2017:9612854. doi: 10.1155/2017/9612854.
4. Scarpa M, Noaro G, Saadeh L, Cavallin F, Cagol M, Alferi R, et al. Esophageal cancer management: preoperative CA19-9 and CEA serum levels may identify occult advanced adenocarcinoma. World J Surg 2015;39:424–432. doi: 10.1007/s00268-014-2835-1.
5. Rupaimoo R, Slack FJ. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. Nat Rev Drug Discov 2017;16:203–222. doi: 10.1038/nrd.2016.246.
6. Bracken CP, Scott HS, Goodall GJ. A network-biology perspective of microRNA function and dysfunction in cancer. Nat Rev Genet 2016;17:179–193. doi: 10.1038/nrg.2016.134.
7. He B, Yin B, Wang B, Xia Z, Chen C, Tang J. MicroRNAs in esophageal cancer (review). Mol Med Rep 2012;6:459–465. doi: 10.3892/mmr.2012.975.
8. Cai X, Yang X, Jin C, Li L, Cui Q, Guo Y, et al. Identification and verification of differentially expressed microRNAs and their target genes for the diagnosis of esophageal cancer. Oncol Lett 2018;16:3642–3650. doi: 10.3892/ol.2018.9066.
9. Wan J, Wu W, Che Y, Kang N, Zhang R. Insights into the potential use of microRNAs as a novel class of biomarkers in esophageal cancer. Dis Esophagus 2016;29:412–420. doi: 10.1111/dote.12338.
10. Mao S, Li Y, Lu Z, Che Y, Sun S, Huang J, et al. Survival-associated splicing signatures in esophageal carcinoma. Cancer Res 2018;78:385–391. doi: 10.1158/0008-5472.CAN-17-2818.
11. Zheng X, Song X, Shao Y, Xu B, Hu W, Zhou Q, et al. Prognostic role of tumor-infiltrating lymphocytes in esophageal cancer: a meta-analysis. Cell Physiol Biochem 2018;45:720–732. doi: 10.1159/00048164.
12. Xu D, Li G, Li H, Jia F. Comparison of IMRT versus 3-D-CRT in the treatment of esophageal cancer: a systematic review and meta-analysis. J Gastrointest Surg 2015;19:1208–1215. doi: 10.1007/s11605-015-2829-9.
13. Cai X, Yang X, Jin C, Li L, Cui Q, Guo Y, et al. MicroRNA function in respiratory diseases: mini-review. Front Cell Physiol 2018;9:596. doi: 10.3389/fcell.2018.00596.
14. Wang X, Li M, Zhang Z, Han S, Tang X, Ge Y, et al. Silencing of long noncoding RNA MALAT1 by miR-101 and miR-217 inhibits proliferation, migration and invasion of esophageal squamous cell carcinoma cells. J Biol Chem 2015;290:3925–3935. doi: 10.1074/jbc.M114.596866.
15. Mabty S, Plank M, Tay HL, Collison A, Foster PS. Targeting microRNA function in respiratory diseases: an update of review. Front Physiol 2016;7:21. doi: 10.3389/fphys.2016.00021.
16. Gong J, Chu Y, Xu M, Ho J, Lv L. Esophageal squamous cell carcinoma cell proliferation induced by exposure to low concentration of cigarette smoke extract is mediated via targeting miR-101-3p/COX-2 pathway. Oncol Rep 2016;35:463–471. doi: 10.3829/or.2015.4379.
17. Fassan M, Pirizzi P, Redaelli S, Balistreri M, Guzzardo V, Zagone V, et al. The HER2-miR125a5p/miR125b loop in gastric and oesophageal adenocarcinomas. Front Biochem 2013;4:1804–1810. doi: 10.1016/j.fibio.2013.01.023.
36. Shi H, Xu J, Zhao R, Wu H, Gu L, Chen Y. FGF2 regulates proliferation, migration, and invasion of ECA109 cells through PI3K/Akt signalling pathway in vitro. Cell Biol Int 2016;40:524–533. doi: 10.1002/cbin.10588.

37. Han S, Lee CW, Trevino JG, Hughes SJ, Sarosi GJ. Autocrine extrapancreatic trypsin 3 secretion promotes cell proliferation and survival in esophageal adenocarcinoma. PLoS One 2013;8:e76667. doi: 10.1371/journal.pone.0076667.

38. Ishibashi O, Akagi I, Ogawa Y, Inui T. MiR-141-3p is upregulated in esophageal squamous cell carcinoma and targets pleckstrin homology domain leucine-rich repeat protein phosphatase-2, a negative regulator of the PI3K/AKT pathway. Biochem Biophys Res Commun 2018;501:507–513. doi: 10.1016/j.bbrc.2018.05.025.

39. Bibby BA, Reynolds JV, Maher SG. MicroRNA-330-5p as a putative modulator of neoadjuvant chemoradiotherapy sensitivity in oesophageal adenocarcinoma. PLoS One 2015;10:e134180. doi: 10.1371/journal.pone.0134180.

40. Mei LL, Wang WJ, Qu YT, Xie XF, Bai J. Shi ZZ. miR-145-5p suppresses tumor cell migration, invasion and epithelial to mesenchymal transition by regulating the Sp1/NF-kappaB signaling pathway in esophageal squamous cell carcinoma. Int J Mol Sci 2017;18. doi: 10.3390/ijms18091833.

41. Liu MX, Liao J, Xie M, Gao ZK, Wang XH, Zhang Y, et al. miR-93-5p transferred by exosomes promotes the proliferation of esophageal cancer cells via intercellular communication by targeting PTEN. Biomed Environ Sci 2018;31:171–185. doi: 10.3967/bes2018.023.

42. Zaidi AH, Saldin LT, Kelly LA, Bergal L, Londono R, Kosovec JE, et al. MicroRNA signature characterizes primary tumors that metastasize in an esophageal adenocarcinoma rat model. PLoS One 2015;10:e0122375. doi: 10.1371/journal.pone.0122375.

43. Chen L, Jin Y, Wang L, Sun F, Yang X, Shi M, et al. Identification of reference genes and miRNAs for qRT-PCR in human esophageal squamous cell carcinoma. Med Oncol 2017;34:2. doi: 10.1007/s12032-016-0860-7.

How to cite this article: Li CY, Zhang WW, Xiang JL, Wang XH, Li J, Wang JL. Identification of microRNAs as novel biomarkers for esophageal squamous cell carcinoma: A study based on The Cancer Genome Atlas (TCGA) and bioinformatics. Chin Med J 2019;132:2213–2222. doi: 10.1097/CMI.0000000000000427