Medicine®

OPEN

Identification of prognostic miRNA biomarkers for esophageal cancer based on The Cancer Genome Atlas and Gene Expression Omnibus

Jinru Xue, MS[a], Erna Jia, MS[b], Na Ren, MS[a], Hua Xin, MD[a],*

Abstract

MicroRNAs (miRNAs) in tumor and tumor-adjacent tissues can be effective diagnostic and prognostic markers to monitor tumor occurrence and progression. Despite improvements in the diagnosis and treatment of esophageal cancer (EC), the survival rate is <25%; consequently, more effective EC-specific prognostic biomarkers are urgently needed to design effective treatment regimens. In this study, we focused on identifying independent prognostic miRNA signatures in tumor and tumor-adjacent tissues in EC.

We screened candidate miRNAs using a genome-wide miRNA transcriptome dataset from The Cancer Genome Atlas (TCGA) database that included 82 patients with esophageal adenocarcinoma (EADC) and 83 patients with esophageal squamous cell carcinoma (ESCC). We validated potential prognostic miRNA markers using a microarray profiling dataset that included information of 32 patients with EADC and 44 patients with ESCC from the Gene Expression Omnibus database. TCGA dataset was additionally used to identify differentially expressed mRNAs (DEMs) between the tumor and tumor-adjacent tissues. Univariate and multivariate Cox analyses were performed to detect the relationship between miRNAs and the overall survival of patients with EC. Kaplan–Meier method was applied to assess the survival differences between groups with differential miRNA expression. Lastly, functional enrichment analysis was conducted using miRWalk 2.0 online database for annotation.

Although there was a considerable difference between the DEMs of EADC and ESCC, 73 DEMs were differentially expressed in both EADC and ESCC samples in TCGA dataset. Cox regression and Kaplan–Meier survival analyses showed that a higher expression of hsa-miR-186-5p and hsa-let-7d-5p was independently associated with a poor prognosis of EADC and ESCC, respectively. Furthermore, gene functional enrichment analysis revealed that the target genes of hsa-miR-186-5p and hsa-let-7d-5p participated in various cancer-related pathways, including the MAPK signaling pathway, proteoglycans in cancer, and AGE-RAGE signaling pathway.

Our results revealed that hsa-miR-186-5p and hsa-let-7d-5p could be used as independent prognostic biomarkers for EADC and ESCC, respectively.

Abbreviations: DEM = differentially expressed microRNAs, EADC = esophageal adenocarcinoma, EC = esophageal cancer, ESCC = squamous cell cancer, GEO = gene expression omnibus, HR = hazard ratio, microRNA = miRNA, OS = overall survival, TCGA = The Cancer Genome Atlas.

Keywords: biomarker, esophageal cancer, Gene Expression Omnibus, microRNA, prognosis, The Cancer Genome Atlas

1. Introduction

Esophageal cancer (EC) is a malignant cancer that originates from the epithelial tissues of the esophageal mucosa. EC is the eighth most common and sixth deadliest cancer worldwide.[1] The majority of cases of EC are diagnosed at an advanced stage and are usually untreatable by the time of diagnosis.[2] Despite constant improvements in the diagnosis and treatment of EC, the overall 5-year survival rate is still below 25% owing to its frequent early metastasis.[3,4] To design more optimal treatment regimens that improve the survival of patients with EC, effective EC-specific prognostic biomarkers are urgently needed.

MicroRNAs (miRNAs) are short, non-coding, single-stranded RNA that are approximately 19 to 24 nucleotides in length.[5,6] miRNAs can stably exist in various body fluids and tissues, and play roles in cell differentiation, proliferation, migration, invasion, and apoptosis.[7] Dysregulation of miRNAs is closely associated with the occurrence and development of tumors by regulating the expression of target genes.[6,8–11] Previous studies have indicated the predictive ability of miRNA markers for the prognosis of EC.[12–15] However, most of these studies focused on

the prognostic value of differentially expressed miRNAs (DEMs) between esophageal tumors and tumor-adjacent tissues. Recent studies have shown that there are significant differences between tumor-adjacent tissues and normal tissues, and the microenvironment of tumor-adjacent tissues is conducive to tumor invasion and metastasis.[16,17] Therefore, it is necessary to screen miRNA markers with prognostic values on a genome-wide scale.

In this study, we screened candidate miRNAs using a genome-wide miRNA transcriptome dataset from The Cancer Genome Atlas (TCGA) database and validated prognostic miRNA markers for EC using a microarray profile dataset from the Gene Expression Omnibus (GEO) database. Besides validating miRNA markers in tumor tissues, we also evaluated the prognostic value of miRNA markers in tumor-adjacent tissues. Additionally, functional annotation analysis was used to further investigate the biological processes of prognostic miRNAs in EC.

2. Methods

2.1. Data profiling and normalization

Genome-wide miRNA profiling data and corresponding clinical information were obtained from TCGA database (http://www.cancergenome.nih.gov; ESCA project) as the marker screening set, and included 182 EC and 13 adjacent noncancerous samples. We excluded samples with overall survival (OS) time ≤3 months. As a result, 82 esophageal adenocarcinoma (EADC) and 83 esophageal squamous cell carcinoma (ESCC) samples remained in the dataset (Fig. 1). The downloaded files had recorded raw read counts for 2072 miRNAs. Raw counts for each miRNA were normalized to trimmed mean of M-values (TMM) and then

subjected to log-transformation using the "edgeR" package in R software (version 4.0.1).

Microarray profiling data from GSE13937 and corresponding clinical information were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/) as a marker validation set, and included 32 EADC samples, 44 ESCC samples, and their corresponding 76 paired-adjacent noncancerous samples (Fig. 1). All samples in this dataset had an OS of >3 months. The normalized expression matrix of GSE13937 displayed 632 miRNAs.

As all data for this study were downloaded from TCGA and GEO databases, there was no requirement for an ethics committee approval.

2.2. Screening of DEMs

Principal component analysis (PCA) showed that the variations among TCGA EC samples were mainly dependent on the race of the patient and histopathological type of the tumor (See Fig. S1, http://links.lww.com/MD/F711, supplemental comment, which illustrates the results of PCA for age, gender, race, smoking, alcohol, and histopathological subtypes of EC). Since the histopathological type of EC is highly related to race, differential expression analyses were performed separately for the EADC and ESCC samples.

DEMs between EADC and ESCC samples and their corresponding adjacent noncancerous samples were analyzed in TCGA miRNA counts data using the "edgeR" package in R. A DEM was defined as a miRNA with a false discovery rate < 0.05 and |log2fold change| (|log2FC|) > 1. Volcano plots and Venn diagrams were used to summarize the identified DEMs.
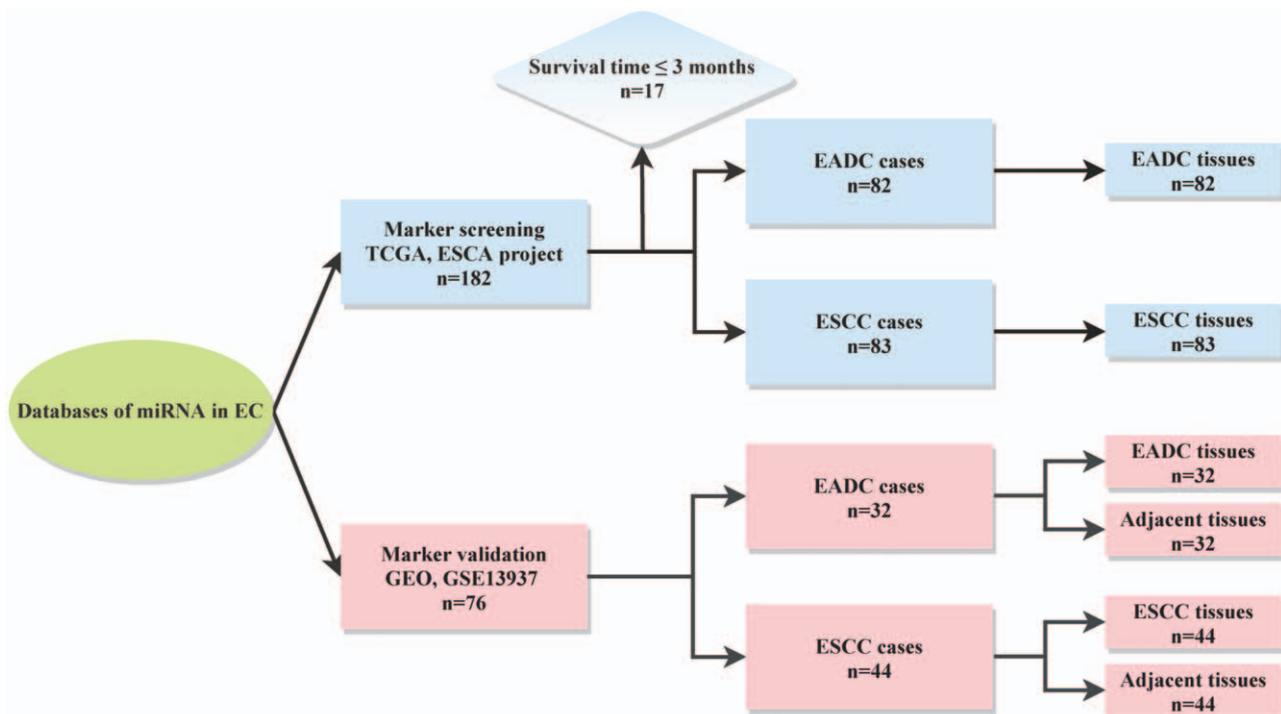
Figure 1. Flow diagram of the study population. EADC = esophageal adenocarcinoma, EC = esophageal cancer, ESCC = esophageal squamous cell carcinoma, GEO = Gene Expression Omnibus, miRNA = microRNA, TCGA = The Cancer Genome Atlas.

2.3. Cox regression analysis of miRNA prognostic markers

Cox proportional hazard regression was performed using the "survival" package in R to screen miRNA candidates on a genome-wide scale using TMM normalized data for determining the OS of patients with EADC and ESCC in TCGA dataset. miRNAs with more than 20% zero expression values were not included in the analyses. First, univariate Cox analyses were used to calculate the hazard ratio (HR) and *P* value of each miRNA. Then, miRNAs whose *P* value was less than .05 were used in the multivariate Cox analyses with adjustment for tumor–node–metastasis (TNM) stage and chemoradiation therapy. Thereafter, miRNAs with *p* value < 0.05 were labeled as "independent prognostic markers" for EADC and ESCC. GSE13937 was used to validate these independent prognostic miRNAs. miRNAs with more than 20% missing values in this dataset were removed from the Cox analyses; whereas, the remaining missing values were imputed with median values of each miRNA. Based on the median values of the expression levels of validated miRNAs, patients with EADC, and ESCC in GSE13937 were classified into high and low-risk groups, accordingly. The Kaplan–Meier method was used to assess the survival differences between these 2 groups.

2.4. Functional enrichment analyses

Target genes of the validated miRNAs were extracted from the miRWalk 2.0 online database, using a comparison of binding sites resulting from 12 existing miRNA-target prediction programs (miRWalk, Microt4, miRanda, mirbridge, miRDB, miRNAMap, Pictar2, PITA, RNA22, RNAhybrid, miRMap, and Targetscan). Target genes that overlapped in 3 or more databases were selected as target genes for the subsequent Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses performed using the "clusterProfiler" package in R.

2.5. Statistical analyses

All statistical analyses were conducted using R software (version 4.0.1; R Foundation, Vienna, Austria). We considered a *P* value <.05 as an indicator of a significant difference.

3. Results

3.1. Identification of DEMs

We compared the tumor tissues with their corresponding adjacent normal tissues in TCGA dataset (See Tables S1, http://links.lww.com/MD/F712 and S2, http://links.lww.com/MD/F713, supplemental comment, which illustrates the DEMs in TCGA EADC and TCGA ESCC samples), observed that a total of 124 (74 upregulated and 50 downregulated) and 155 miRNAs (84 upregulated and 71 downregulated) were differentially expressed in EADC and ESCC samples, respectively. Consistent with results from the PCA analysis (Fig. S1, http://links.lww.com/MD/F711), the Venn diagram showed a considerable difference between the DEMs of EADC and ESCC, and only 73 DEMs were differentially expressed in both the EADC and ESCC samples (Fig. 2).
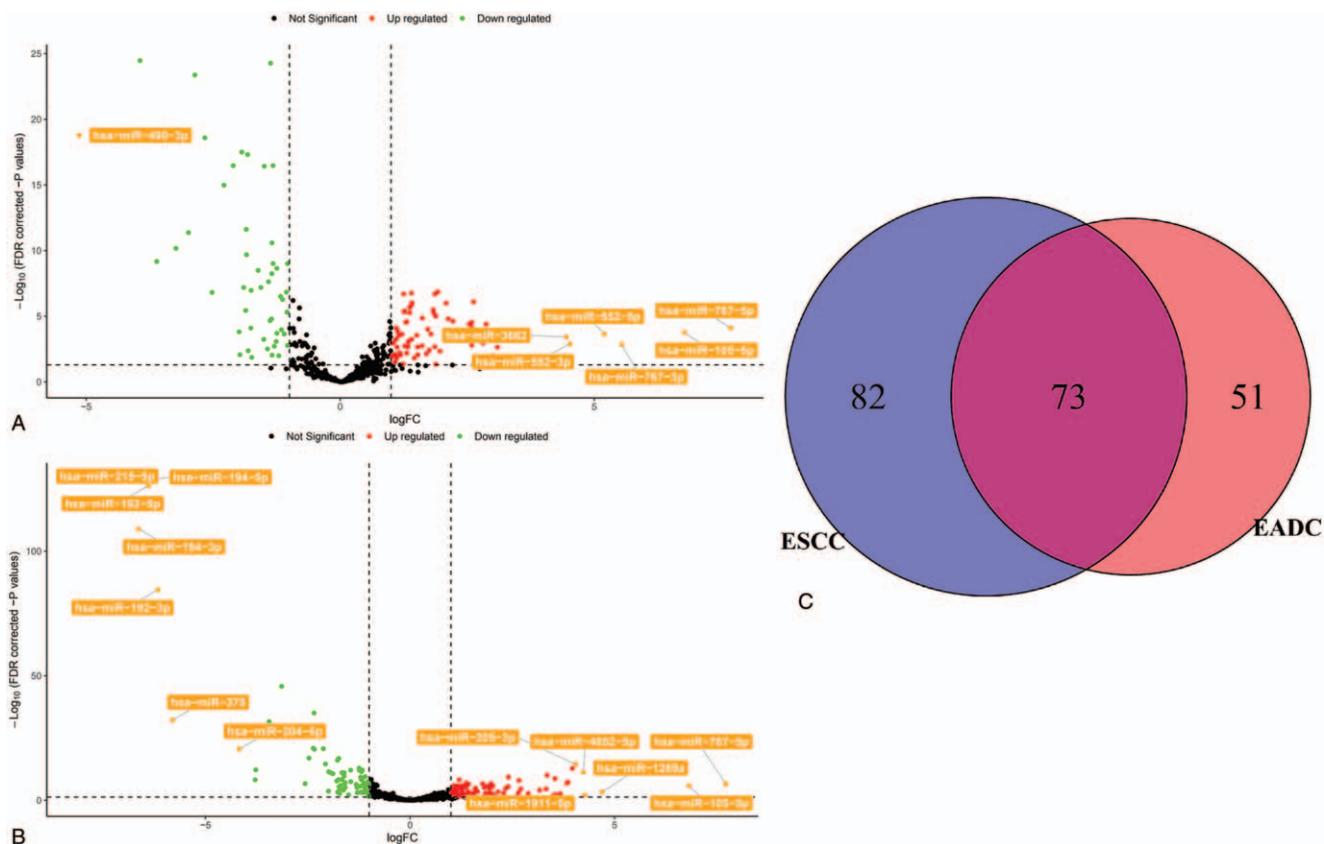
Figure 2. Differentially expressed microRNAs between the tumor and tumor-adjacent tissues in esophageal cancer samples from The Cancer Genome Atlas. (A) Volcano plot shows the differentially expressed miRNAs for esophageal adenocarcinoma (EADC). (B) Volcano plot shows the differentially expressed miRNAs for esophageal squamous cell carcinoma (ESCC). (C) Venn diagram shows the number of differentially expressed miRNAs of both EADC and ESCC. FDR = false discovery rate, logFC = log fold change.

3.2. Identification and validation of prognostic miRNAs

The characteristics of patients with EC are listed in Table 1. Results from the Cox regression analyses of TCGA dataset are listed in Tables S3, http://links.lww.com/MD/F714 and S4, http://links.lww.com/MD/F715 (supplemental comment, which illustrates the results of Cox regression analyses for patients with EADC and ESCC in TCGA dataset). Using the univariate Cox regression analyses, 63 and 53 miRNAs (all with $P < .05$) were screened out for predicting the OS of patients with EADC and ESCC, respectively. After adjusting for TNM stage and chemoradiation therapy in the multivariate Cox regression analyses, 37 and 32 miRNAs (all with $P < .05$) showed independent prognostic values for patients with EADC and ESCC, respectively. Twelve of the 37 miRNAs for EADC and 10 of the 32 miRNAs for ESCC were identified as DEMs in TCGA dataset.

Among the 37 miRNAs identified for EADC in TCGA-ESCA dataset, 9 were available in the GSE13937 dataset. Table 2 shows the Cox regression analyses of these 9 miRNAs in patients with EADC. In TCGA dataset, 3 miRNAs (hsa-miR-197-3p, hsa-miR-186-5p, and hsa-miR-191-5p) were identified as independent risk factors for predicting the OS (all with $P < .05$) and 6 miRNAs (hsa-miR-150-5p, hsa-miR-29c-3p, hsa-miR-29a-3p, hsa-miR-23b-3p, hsa-miR-27b-3p, and hsa-let-7b-5p) were identified as independent protective factors for OS (all with $P < .05$). In the GSE13937 dataset, multivariate Cox analyses with adjustment for TNM stage and chemoradiation therapy showed that hsa-miR-186-5p and hsa-miR-27b-3p were independent risk factors for predicting the OS (HR = 7.53, 95% CI = 1.47–38.58, $P = .015$; HR = 2.49, 95% CI = 1.04–5.98, $P = .041$, respectively).

However, univariate Cox analyses showed marginal significance for hsa-miR-27b-3p ($P = .051$). Kaplan–Meier survival analysis indicated that higher expression of hsa-miR-186-5p in the GSE13937 EADC samples was negatively correlated with the survival rate of patients with EADC ($P = .045$, Fig. 3A). Among the 9 miRNAs in the EADC adjacent tissues of the GSE13937 dataset, only hsa-miR-150-5p was found to be a significant risk factor for OS in both univariate and multivariate Cox analyses ($P = .044$ and $P = .037$, respectively). Both hsa-miR-186-5p and hsa-miR-27b-3p were not significant in the Cox analyses in the EADC adjacent tissues (all showed $P > .05$).

Among the 32 miRNAs for ESCC in TCGA dataset, 10 were available in the GSE13937 dataset. Table 3 shows the Cox regression analyses of these 10 miRNAs in patients with ESCC. In TCGA dataset, 4 miRNAs (hsa-miR-126-3p, hsa-miR-340-3p, hsa-let-7d-5p, and hsa-miR-192-5p) were identified as independent risk factors for predicting the OS (all $P < .05$), and six miRNAs (hsa-miR-105-5p, hsa-miR-200b-3p, hsa-miR-429, hsa-miR-494-3p, hsa-miR-376b-3p, and hsa-miR-320a) were identified as independent protective factors for OS (all $P < .05$). In the GSE13937 dataset, multivariate Cox analyses with adjustment for TNM stage and chemoradiation therapy demonstrated that hsa-let-7d-5p was an independent risk factor for OS (HR = 2.65, 95% CI = 1.06–6.65, $P = .038$). Kaplan–Meier survival analysis indicated that higher expression of hsa-let-7d-5p in the GSE13937 ESCC samples was negatively correlated with the survival rate of patients with ESCC ($P = .007$, Fig. 3B). Among the 10 miRNAs in the ESCC adjacent tissues of the GSE13937 dataset, no miRNAs were associated with OS (all had $P > .05$).

3.3. Functional enrichment analyses of hsa-miR-186-5p and hsa-let-7d-5p

A total of 9686 and 6586 target genes were identified for hsa-miR-186-5p and hsa-let-7d-5p, respectively. Figure 4 shows the top 10 GO annotations and KEGG enrichment terms for hsa-miR-186-5p and hsa-let-7d-5p. In the functional enrichment analyses, the biological processes of hsa-miR-186-5p-targeted genes were enriched in regulation of cell morphogenesis, and hsa-let-7d-5p-targeted genes were enriched in regulation of cell morphogenesis, regulation of GTPase, and regulation of Ras protein signal transduction. The KEGG pathway results revealed that both the hsa-miR-186-5p- and hsa-let-7d-5p-targeted genes were significantly enriched in cancer-related pathways, including the MAPK signaling pathway, proteoglycans in cancer, and AGE-RAGE signaling pathway.

4. Discussion

Abnormal miRNA expression affects the molecular functions and biological processes of various tumors; consequently, numerous attempts have been made to improve the precision of EC prognosis prediction using miRNA biomarkers. However, most previous studies have focused on a specific histopathological subtype of EC (EADC or ESCC) and have not conducted external validation of the predicted miRNAs.[18–20] In addition, only a few studies have focused on identifying EC-specific prognostic miRNAs in tumor-adjacent tissues. Therefore, in this study, we comprehensively analyzed prognostic miRNAs of 2 main histopathological subtypes of EC (EADC and ESCC) on a genome-wide scale in TCGA-ESCA dataset, and validated these miRNAs in both tumor and tumor-adjacent tissues of EC in the

Table 1

Clinical characteristics of the study populations.

Characteristics	TCGA-ESCA (n = 165)	GSE13937 (n = 76)
Age		
<60	75 (45.5%)	/
≥60	90 (54.5%)	/
Sex		
Female	27 (16.4%)	/
Male	138 (83.6%)	/
Smoking history		
Yes	96 (58.2%)	58 (76.3%)
No	51 (30.9%)	14 (18.4%)
Unknown	18 (10.9%)	4 (5.3%)
Alcohol history		
Yes	116 (70.3%)	57 (75.0%)
No	48 (29.1%)	13 (17.1%)
Unknown	1 (0.6%)	6 (7.9%)
Histopathology		
EADC	82 (49.7%)	32 (42.1%)
ESCC	83 (50.3%)	44 (57.9%)
TNM stage		
0	0 (0.0%)	14 (18.4%)
I	17 (10.3%)	11 (14.5%)
II	76 (46.1%)	31 (40.8%)
III	52 (31.5%)	10 (13.2%)
IV	15 (9.1%)	10 (13.2%)
Unknown	5 (3.0%)	0 (0.0%)
Chemoradiation therapy		
Yes	75 (45.5%)	39 (51.3%)
No	77 (46.7%)	37 (48.7%)
Unknown	13 (7.9%)	0 (0.0%)

EADC = esophageal adenocarcinoma, ESCC = esophageal squamous cell cancer.

Table 2

Univariate and multivariate Cox regression analysis of the 9 miRNAs in patients with EADC.

Dataset	miRNA	Univariate analysis		Multivariate analysis	
		HR (95% CI)	p value	HR (95% CI)	P value
TCGA-ESCA, EADC	hsa-let-7b-5p	0.49 (0.33–0.73)	<.001[*]	0.53 (0.34–0.85)	.008[*]
	hsa-miR-150-5p	0.79 (0.63–1.00)	.050[*]	0.68 (0.52–0.91)	.009[*]
	hsa-miR-186-5p	2.39 (1.15–4.96)	.020[*]	3.10 (1.32–7.26)	.009[*]
	hsa-miR-191-5p	1.65 (1.09–2.50)	.018[*]	1.69 (1.08-2.65)	.022[*]
	hsa-miR-197-3p	2.82 (1.55–5.12)	.001[*]	2.73 (1.38–5.42)	.004[*]
	hsa-miR-23b-3p	0.52 (0.35–0.79)	.002[*]	0.59 (0.38–0.93)	.021[*]
	hsa-miR-27b-3p	0.51 (0.32–0.82)	.006[*]	0.56 (0.35–0.91)	.018[*]
	hsa-miR-29a-3p	0.57 (0.36–0.88)	.012[*]	0.55 (0.34–0.89)	.016[*]
	hsa-miR-29c-3p	0.62 (0.44–0.87)	.006[*]	0.64 (0.45–0.91)	.014[*]
GSE13937, EADC	hsa-let-7b-5p	1.66 (0.79–3.51)	.181	2.13 (0.85–5.33)	.107
	hsa-miR-150-5p	1.45 (0.86–2.45)	.163	1.49 (0.77–2.88)	.241
	hsa-miR-186-5p	4.13 (1.26–13.50)	.019[*]	7.53 (1.47–38.58)	.015[*]
	hsa-miR-191-5p	1.31 (0.67–2.55)	.425	1.25 (0.59–2.67)	.561
	hsa-miR-197-3p	0.53 (0.17–1.59)	.256	0.66 (0.19–2.32)	.515
	hsa-miR-23b-3p	1.09 (0.46–2.61)	.842	1.09 (0.34–3.52)	.883
	hsa-miR-27b-3p	1.74 (1.00–3.05)	.051	2.49 (1.04–5.98)	.041
	hsa-miR-29a-3p	1.44 (0.76–2.73)	.265	1.56 (0.73–3.35)	.256
	hsa-miR-29c-3p	1.26 (0.72–2.21)	.427	1.28 (0.65–2.52)	.477
GSE13937, EADC adjacent tissue	hsa-let-7b-5p	1.31 (0.70–2.47)	.399	1.60 (0.74–3.44)	.232
	hsa-miR-150-5p	2.01 (1.02–3.97)	.044[*]	2.50 (1.06–5.90)	.037[*]
	hsa-miR-186-5p	1.01 (0.66–1.54)	.981	0.99 (0.63–1.55)	.948
	hsa-miR-191-5p	1.67 (0.92–3.02)	.092	1.51 (0.83–2.76)	.180
	hsa-miR-197-3p	0.63 (0.29–1.38)	.250	0.70 (0.28–1.75)	.449
	hsa-miR-23b-3p	1.29 (0.64–2.60)	.468	1.61 (0.61–4.22)	.337
	hsa-miR-27b-3p	1.38 (0.89–2.14)	.152	1.64 (0.85–3.17)	.143
	hsa-miR-29a-3p	1.55 (0.95–2.51)	.078	1.61 (0.92–2.81)	.095
	hsa-miR-29c-3p	1.61 (1.01–2.57)	.046[*]	1.63 (0.94–2.83)	.081

CI = confidence interval, EADC = esophageal adenocarcinoma, HR = hazard ratio.

[*] represents $P < .05$.
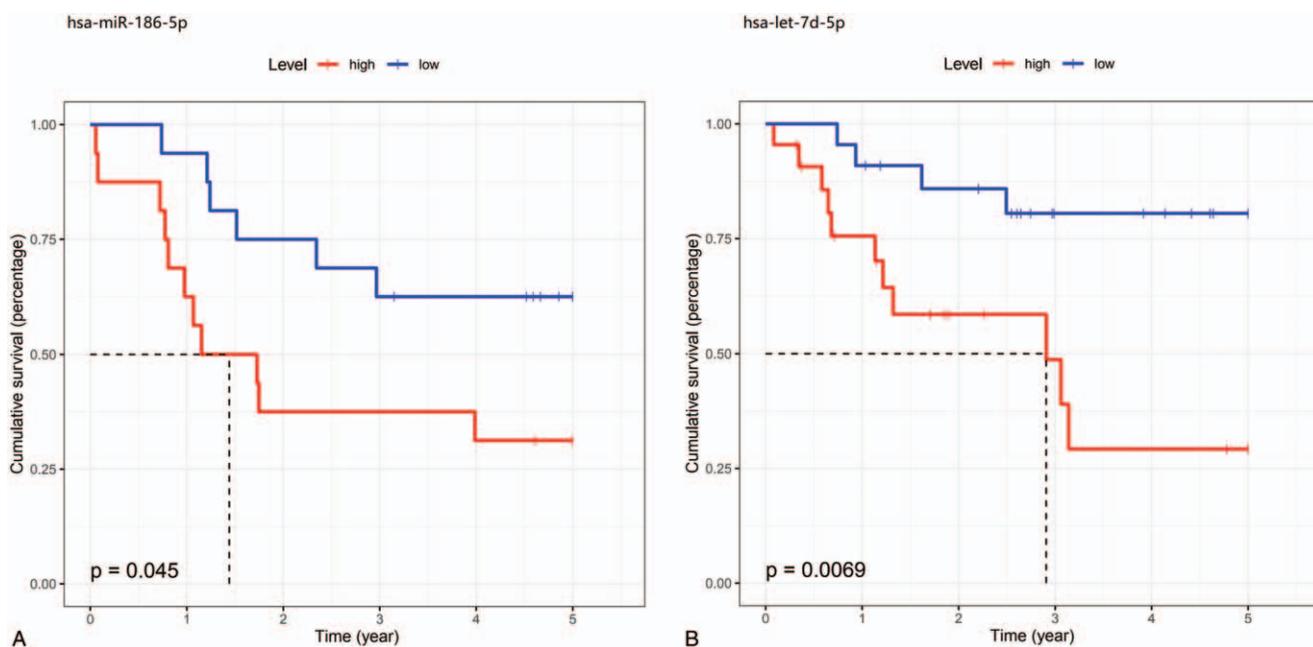
Figure 3. Kaplan–Meier curves show the differences in survival between the high-risk and low-risk groups of patients with esophageal cancer. (A) Kaplan–Meier curves for hsa-miR-186-5p in esophageal adenocarcinoma. (B) Kaplan–Meier curves for hsa-let-7d-5p in esophageal squamous cell carcinoma.

Table 3

Univariate and multivariate Cox regression analysis of the 10 miRNAs in patients with ESCC.

Dataset	miRNA	Univariate analysis		Multivariate analysis	
		HR (95% CI)	*p* value	HR (95% CI)	*P* value
TCGA-ESCA, ESCC	hsa-let-7d-5p	2.24 (1.04–4.82)	.040*	2.56 (1.05–6.24)	.039*
	hsa-miR-105-5p	0.85 (0.74–0.98)	.025*	0.82 (0.69–0.97)	.022*
	hsa-miR-126-3p	3.01 (1.56–5.78)	.001*	3.04 (1.42–6.54)	.004*
	hsa-miR-192-5p	1.84 (1.07–3.17)	.028*	2.28 (1.14–4.54)	.019*
	hsa-miR-200b-3p	0.73 (0.56–0.97)	.028*	0.68 (0.52–0.91)	.008*
	hsa-miR-320a	0.43 (0.22–0.86)	.017*	0.36 (0.16–0.81)	.013*
	hsa-miR-340-3p	2.36 (1.12–4.99)	.024*	3.11 (1.28–7.51)	.012*
	hsa-miR-376b-3p	0.51 (0.28–0.91)	.023*	0.48 (0.24–0.96)	.038*
	hsa-miR-429	0.72 (0.54–0.96)	.023*	0.67 (0.50–0.89)	.006*
	hsa-miR-494-3p	0.52 (0.30–0.90)	.019*	0.47 (0.25–0.89)	.020*
GSE13937, ESCC	hsa-let-7d-5p	2.26 (1.11–4.62)	.025*	2.65 (1.06–6.65)	.038*
	hsa-miR-105-5p	1.8 (0.71–4.55)	.214	1.38 (0.49–3.86)	.543
	hsa-miR-126-3p	1.42 (0.81–2.5)	.222	1.63 (0.79–3.33)	.184
	hsa-miR-192-5p	1.74 (0.71–4.26)	.228	1.42 (0.53–3.81)	.482
	hsa-miR-200b-3p	1.27 (0.85–1.89)	.238	1.5 (0.88–2.54)	.135
	hsa-miR-320a	3.06 (1.25–7.53)	.015*	2.37 (0.86–6.51)	.094
	hsa-miR-340-3p	1.13 (0.7–1.83)	.621	0.94 (0.56–1.57)	.815
	hsa-miR-376b-3p	1.59 (0.79–3.21)	.196	1.84 (0.83–4.07)	.130
	hsa-miR-429	2.24 (0.87–5.74)	.094	1.74 (0.57–5.3)	.331
	hsa-miR-494-3p	1.01 (0.72–1.42)	.939	0.92 (0.61–1.38)	.675
GSE13937, ESCC adjacent tissue	hsa-let-7d-5p	0.69 (0.28–1.74)	.432	0.98 (0.33–2.9)	.972
	hsa-miR-105-5p	1.66 (0.85–3.24)	.139	1.51 (0.68–3.35)	.314
	hsa-miR-126-3p	0.94 (0.47–1.88)	.863	1.12 (0.45–2.78)	.803
	hsa-miR-192-5p	0.94 (0.63–1.4)	.761	0.9 (0.58–1.41)	.648
	hsa-miR-200b-3p	0.87 (0.6–1.24)	.435	0.69 (0.42–1.14)	.148
	hsa-miR-320a	2.24 (0.82–6.14)	.115	1.45 (0.52–4.01)	.474
	hsa-miR-340-3p	1.69 (0.92–3.12)	.091	1.47 (0.82–2.64)	.201
	hsa-miR-376b-3p	1.31 (0.71–2.42)	.390	1.52 (0.75–3.08)	.246
	hsa-miR-429	1.57 (0.65–3.77)	.315	1.49 (0.57–3.91)	.417
	hsa-miR-494-3p	0.98 (0.66–1.44)	.903	0.82 (0.56–1.22)	.336

CI = confidence interval, ESCC = esophageal squamous cell cancer, HR = hazard ratio.
* represents $P < .05$.

GSE13937 dataset. To eliminate the effects of potential confounders, we performed multivariate Cox analyses by adjusting for TNM stage and chemoradiation therapy. We found that hsa-miR-186-5p and hsa-let-7d-5p may serve as independent prognostic markers for EADC and ESCC, respectively. Moreover, hsa-miR-186-5p and hsa-let-7d-5p could classify the patients with EADC and ESCC from the GSE13937 dataset, into 2 risk groups with significantly different OS rates.

Previous studies have shown that the differential expression of hsa-miR-186-5p is correlated with survival in patients with various tumors, including acute myeloid leukemia, non-small cell lung cancer, neuroblastoma, gastrointestinal stromal tumors, pancreatic ductal adenocarcinoma, and EC.[21] Zhao et al[22] found that in EC cases with the same TNM stage, those with poor prognosis exhibited a lower level of ESCC hsa-miR-186-5p expression than those with good prognosis. In contrast, our study showed that elevated hsa-miR-186-5p expression in tumor tissues was associated with a poor prognosis of EADC. These contrasting results may be explained by the difference in the histopathological subtypes of EC, wherein hsa-miR-186-5p may display an opposite prognostic effect in EADC and ESCC patients. To date, only a few studies have distinguished the prognostic effect of hsa-miR-186-5p in different histopathological subtypes of EC. In addition, EADC is more common in populations from Europe and the United States; whereas, ESCC is more common in Asian populations.[23] Thus, racial differences

may also lead to an opposite prognostic effect of hsa-miR-186-5p in patients with EADC and ESCC.

Let-7 is one of the earliest discovered miRNA families in *Homo sapiens*. Studies have shown that let-7 plays an important role in ESCC by regulating oncogenes associated with cell proliferation and differentiation directly or indirectly to affect the invasion and metastasis of ESCC.[24–26] miRNAs in the let-7 family usually act as tumor suppressor genes, and lower expression of these miRNAs is positively correlated with the prognosis of patients with cancer.[27] Ling et al[28] revealed that lower expression of let-7a expression is correlated with higher TNM stages and recurrence in patients with ESCC. In our study, hsa-let-7d-5p was an independent prognostic factor for ESCC. However, here, an elevated hsa-let-7d-5p level implied poorer prognosis in both TCGA-ESCA and GSE13937 datasets. In addition, hsa-let-7d-5p was not identified as a DEM between the ESCC and corresponding tumor-adjacent tissues. Therefore, future studies should focus on explicating the role of hsa-let-7d in the development of ESCC before its clinical application.

This study revealed that some miRNAs in the EC-adjacent tissues were useful for prognosis prediction. For instance, in the GSE13937 dataset, hsa-miR-150-5p in the EADC-adjacent tissues was significantly associated with EADC prognosis in both univariate and multivariate analyses, which suggested that the tumor-adjacent tissues may be linked to EC. Recently, a study[29] showed that normal tissues adjacent to tumors can be
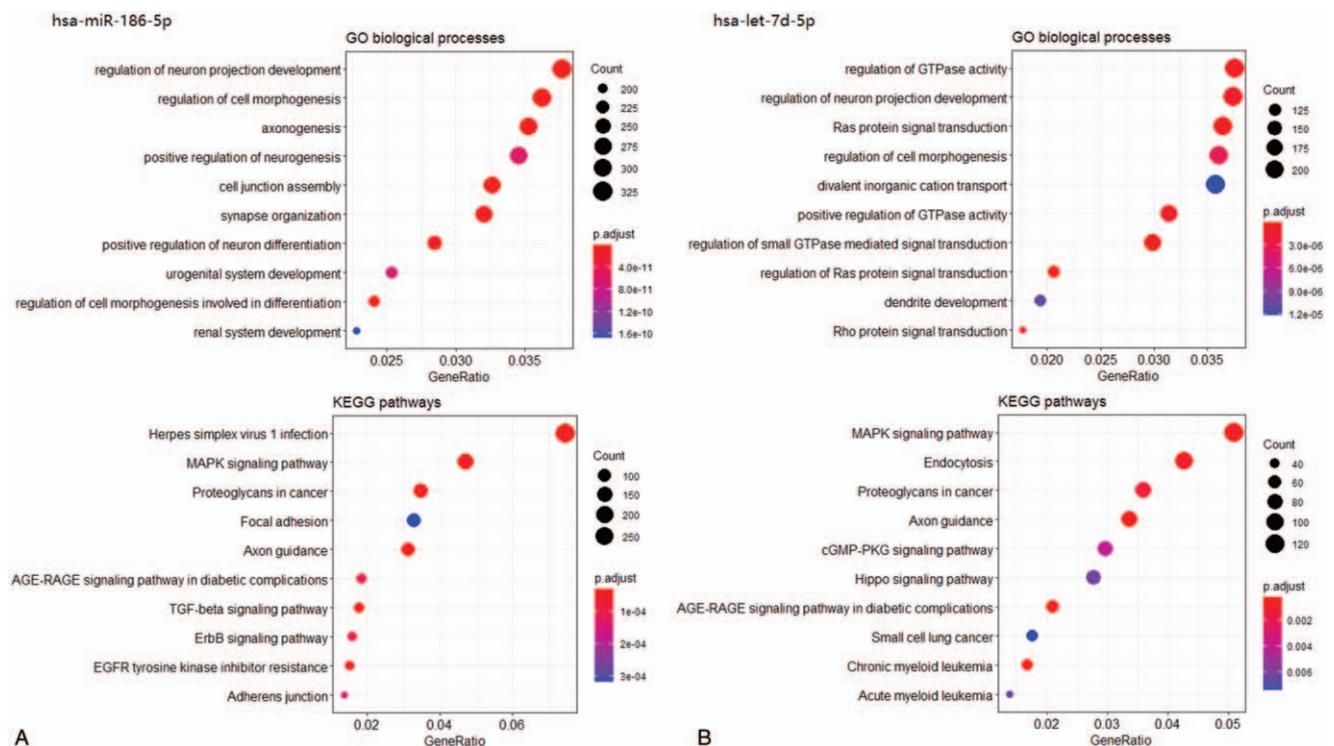
Figure 4. Functional annotation analysis of hsa-miR-186-5p- and hsa-let-7d-5p-targeted genes. (A) Top 10 enriched gene ontology (GO) biological processes and top 10 enriched KEGG pathways of hsa-miR-186-5p-targeted genes. (B) Top 10 enriched GO biological processes and top 10 enriched KEGG pathways of hsa-let-7d-5p-targeted genes.

characterized by peculiar gene expression profiles and biological pathways that are different from both the tumor and "real" normal tissues. Rebeca et al[30] have used a large sample size of normal mucosa (as a reference for tumor-adjacent mucosa) and tumor samples to compare gene expression profiles, and identified many differentially expressed genes, which could be grouped into 3 alteration patterns: "tumor-like," "trend," and "adjacent-specific". A comprehensive analysis of transcriptomes of tumor-adjacent tissues has indicated that these tissues have a unique intermediate status, and approximately 63.8% of differentially expressed genes between tumor and normal tissues are identical to the differentially expressed genes between the tumor and tumor-adjacent tissues. In this study, we also showed that tumor-adjacent tissues are enriched in inflammatory response-related genes (e.g., tumor necrosis factor-α- signaling) and several cancer-related signatures (e.g., KRAS signaling), and can spread proinflammatory signals of a tumor to the surroundings to induce tumor development.[16,17] Therefore, the status of tumor-adjacent tissues could be a transitional stage between tumor and non-adjacent normal tissues, and biomarkers in the tumor-adjacent tissues may be potential tumor prognosis predictive markers as well.

Our study has several limitations. First, although the identified DEMs were associated with EC prognosis, no additional in vitro experiments were performed to validate these findings. Second, the expression of miRNA in the GEO dataset was obtained using miRNA microarray rather than next-generation sequencing. Because of platform differences, the amount of miRNAs available in the GEO dataset was insufficient, which meant that the verification of all prognostic miRNAs identified in TCGA dataset was not possible. Therefore, future studies are needed to identify and verify more miRNAs for EC prognosis.

5. Conclusion

In the 2 independent miRNA datasets acquired from TCGA-ESCA and GSE13937, hsa-miR-186-5p and hsa-let-7d-5p were identified as independent prognostic miRNAs for EADC and ESCC, respectively. Higher expression levels of these miRNAs indicated a poorer OS. While our findings may be clinically valuable in predicting the prognosis of EC, further studies with larger sample sizes are needed to verify these findings.

Author contributions

Conceptualization: Hua Xin.
Data curation: Jinru Xue.
Formal analysis: Jinru Xue, Erna Jia.
Methodology: Hua Xin.
Software: Erna Jia, Na Ren.
Supervision: Hua Xin.
Writing – original draft: Jinru Xue.
Writing – review & editing: Jinru Xue, Erna Jia, Na Ren.

References

[1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018;68:7–30.
[2] Nieman DR, Peters JH. Treatment strategies for esophageal cancer. Gastroenterol Clin N Am 2013;42:187–97.
[3] Enzinger PC, Mayer RJ. Esophageal cancer. N Engl J Med 2003; 349:2241–52.

[4] Hamai Y, Hihara J, Emi M, et al. Results of neoadjuvant chemoradiotherapy with docetaxel and 5-fluorouracil followed by esophagectomy to treat locally advanced esophageal cancer. Ann Thorac Surg 2015;99:1887–93.

[5] Weber JA, Baxter DH, Zhang S, et al. The microRNA spectrum in 12 body fluids. Clin Chem 2010;56:1733–41.

[6] Kong YW, Ferland-McCollough D, Jackson TJ. Bushell M. microRNAs in cancer management. Lancet Oncol 2012;13:e249–58.

[7] O'Brien J, Hayder H, Zayed Y, et al. Overview of microRNA biogenesis, mechanisms of actions, and circulation. Front Endocrinol 2018; 9:402.

[8] Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281–97.

[9] Yang H, Gu J, Wang KK, et al. MicroRNA expression signatures in Barrett's esophagus and esophageal adenocarcinoma. Clin Cancer Res 2009;15:5744–52.

[10] Calin GA, Croce CM. MicroRNA signatures in human cancers. Nature reviews. Cancer 2006;6:857–66.

[11] He B, Yin B, Wang B, et al. MicroRNAs in esophageal cancer. Molecul Med Rep 2012;6:459–65.

[12] Mathé EA, Nguyen GH, Bowman ED, et al. MicroRNA expression in squamous cell carcinoma and adenocarcinoma of the esophagus: associations with survival. Clin Cancer Res 2009;15:6192–200.

[13] Hong L, Han Y, Zhang H, et al. Prognosis-related microRNAs in esophageal cancer. Expert Opin Biol Therapy 2014;14:483–9.

[14] Qi B, Yao WJ, Zhao BS, et al. Involvement of microRNA-198 overexpression in the poor prognosis of esophageal cancer. Asian Pacific J Cancer Prevent 2013;14:5073–6.

[15] Feber A, Xi L, Luketich JD, et al. MicroRNA expression profiles of esophageal cancer. J Thorac Cardiovasc Surg 2008;135:255–60. discussion 260.

[16] Aran D, Camarda R, Odegaard J, et al. Comprehensive analysis of normal adjacent to tumor transcriptomes. Nat Commun 2017;8:1077.

[17] Singh R, Mishra MK, Aggarwal H. Inflammation, immunity, and cancer. Mediators Inflamm 2017;2017:6027305.

[18] Matsushima K, Isomoto H, Kohno S, et al. MicroRNAs and esophageal squamous cell carcinoma. Digestion 2010;82:138–44.

[19] Yu J, Zhu M, Lv M, et al. Characterization of a five-microRNA signature as a prognostic biomarker for esophageal squamous cell carcinoma. Scientific Rep 2019;9:19847.

[20] Fu W, Pang L, Chen Y, et al. The microRNAs as prognostic biomarkers for survival in esophageal cancer: a meta-analysis. Scientific World J 2014;2014:523979.

[21] Wang Z, Sha HH, Li HJ. Functions and mechanisms of miR-186 in human cancer. Biomed Pharmacotherap 2019;119:109428.

[22] Zhao BS, Liu SG, Wang TY, et al. Screening of microRNA in patients with esophageal cancer at same tumor node metastasis stage with different prognoses. Asian Pac J Precent 2013;14:139–43.

[23] 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 2018;68:394–424.

[24] Su JL, Chen PS, Johansson G, et al. Function and regulation of let-7 family microRNAs. MicroRNA 2012;1:34–9.

[25] Liu Q, Lv GD, Qin X, et al. Role of microRNA let-7 and effect to HMGA2 in esophageal squamous cell carcinoma. Mol Biol Rep 2012;39:1239–46.

[26] Büssing I, Slack FJ, Grosshans H. let-7 microRNAs in development, stem cells and cancer. Trends Mol Med 2008;14:400–9.

[27] Chirshev E, Oberg KC, Ioffe YJ, et al. Let-7 as biomarker, prognostic indicator, and therapy for precision medicine in cancer. Clin Transl Med 2019;8:24.

[28] Ling R, Zhou Y, Zhou L, et al. Lin28/microRNA-let-7a promotes metastasis under circumstances of hyperactive Wnt signaling in esophageal squamous cell carcinoma. Mol Med Rep 2018;17:5265–71.

[29] Russi S, Calice G, Ruggieri V, et al. Gastric normal adjacent mucosa versus healthy and cancer tissues: distinctive transcriptomic profiles and biological features. Cancers 2019;11:1248.

[30] Sanz-Pamplona R, Berenguer A, Cordero D, et al. Aberrant gene expression in mucosa adjacent to tumor reveals a molecular crosstalk in colon cancer. Mol Cancer 2014;13:46.