**Combined miRNA and DNA methylation as a diagnosis panel for esophageal squamous cell carcinoma**

**Abstract**

**Background：**

Although there are more biomarkers in the diagnosis of esophageal squamous cell carcinoma (ESCC), but because of the lack of sufficient sensitivity and specificity, still can't meet the clinical requirements. Therefore, it is immensely important to discover new biomarkers or combine several individually biomarkers to improve diagnostic effect for ESCC.

**Methods：**

MiRNA-seq data obtained from the Cancer Genome Atlas (TCGA) cohort to select differentially expressed miRNAs as our candidate miRNAs. Then quantitative real-time RT-PCR (qRT-PCR) was carried out to compare the candidate miRNA levels in 93 pairs of ESCC tissues and adjacent non-cancerous tissues,which further screened for target miRNAs. MicroRNA expression levels were further analyzed in regards to clinico-pathological features of esophageal cancer patients, and the miRNAs with prediction potential were identifed.

**Results：**

Nine miRNAs selected from the TCGA were as candidated markers for further validation. The qPCR results showed miR-30e-3p, miR-21-3p and miR-21-5p were significantly differentially expressed between 93 paired of ESCC and adjacent tissues performing paired t test. 表达量与临床信息的结果。ROC curve showed a strong separation between the ESCC and paracancerous groups, with an AUC of 0.80.

**Conclusions：**

We conclude that the miR-30e-3p, miR-21-3p and miR-21-5p can form a diagnosis panel for ESCC.

Keywords: Esophageal squamous cell carcinoma, miRNA, Biomarker

**Background**

Esophageal squamous cell carcinoma (ESCC) is the most common histopathological subtype of various esophageal malignancies(PMID: 24834141). The global incidence of squamous cell cancer has more or less remained stable and it represented 87% of all cases of esophageal cancer in 2012(PMID: 25320104). In China, it accounts for approximately 90% of oesophageal carcinomas, which is one of the areas with the highest morbidity of esophageal cancer(PMID: 25651787). It was estimated that about 477,900 new diagnoses and 375,000 deaths occurred in China in 2015(PMID: 26808342). Patients tend to present with dysphagia at a late stage and as a result the overall 5-year survival is less that 15%. In contrast, tumors detected at an early stage prior to lymph node spread have a survival in excess of 80%(PMID: 24013023). Hence, novel and reliable biomarkers to detect ESCC are urgently needed for early intervention with the potential to reduce mortality of the disease.

MicroRNAs (miRNAs) are small, well-conserved, non-coding RNAs of 20-24 nucleotides that regulate gene expression by binding to 3’ UTR of their specific target mRNAs, which leads to mRNA degradation or inhibition of translation. Studies suggested that miRNAs can regulate about 30 % of genes in human genome [reference], and it was found that about 50 % of miRNAs are located in cancerassociated genomic region. Increasing investigations have confirmed that miRNAs can function as tumor suppressors or oncogenes and play essential roles in a wide variety of physiological and pathological processes in various cancers, including ESCC. Further evidence supports that miRNAs are useful diagnostic and prognostic markers in human cancer(PMID: 20595229). For example, miR-31 could serve as a potential diagnostic and prognostic biomarker for ESCC(PMID: 21658006).

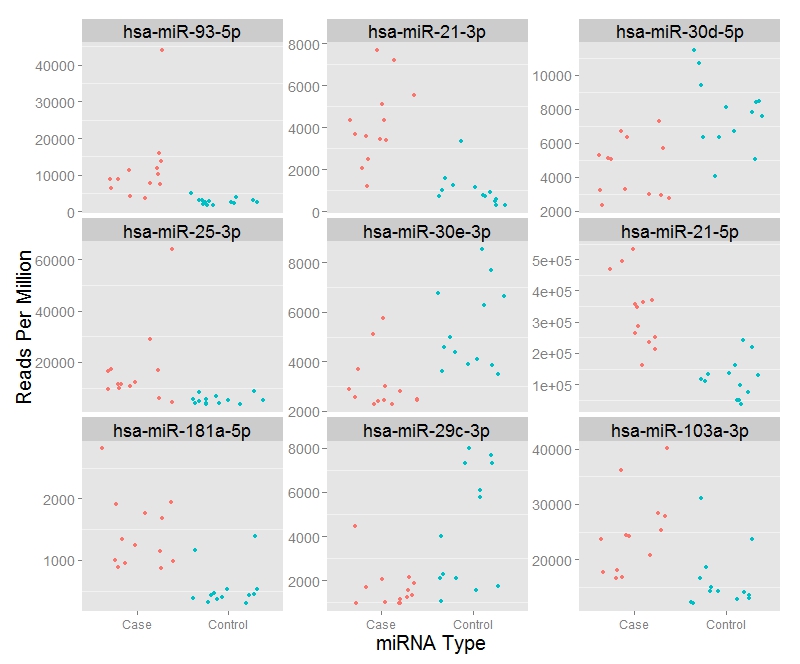
Many researches have showed that miR-21 play an important role in many cancers, not only in diagnosis but also in therapy. Studies have shown that aberrantly high miR-21 expression can regulate cell cycle, apoptosis, invasion, proliferation, and metastasis in several kinds of cancer. However, miR-21-3p which formed by its complementary chain was less studied. miR-21-3p can repressed the process of cancers which is similar to the miR-21-5p. However, it is unclear whether the expression of miR-21-3p is altered in ESCC. Regarding to the miR-30e-3p, study has shown that the expression of which was related to the prognosis in esophageal cancer. Otherwise, the diagnosis in esophageal cancer has not been reported.

In this study, we screened diﬀerentially expressed miRNAs in EC by bioinformation. Then we conducted the miRNA profiles by qRT-PCR to identify potential biomarkers for dianosis of ESCC. As a result, we confirmed that the three miRNAs were significant differences in expression between ESCC and adjacent tissues. The relationship between clinical information and miRNAs expression was analyzed.In addition, through performing ROC, we established a panel combining three miRNAs for ESCC diagnosis. We hoped to determine molecular markers for early diagnosis.

**Results**

**Identify miRNAs that were differentially expressed between ESCC and adjacent tissues**

Based on the microRNA-seq data of EC from TCGA database, we analyzed to obtain each mature miRNA of each sample and then screened using two criteria: abundance at least above 500 reads / millon; P< 0.05 (Wilcoxon's rank sum test) . The selected miRNAs were screened again by Random Forest. Finally, we obtained 9 differentially expressed miRNAs between ESCC and adjacent tissues after analysis by SAMR in R package, including 6 up-regulated and 3 down-regulated . The up-regulated miRNAs were hsa-miR-93-5p, hsa-miR-21-3p, hsa-miR-25-3p, hsa-miR-21-5p, hsa-miR-181a-5p, and hsa-miR-103a-3p. while the down-regulated miRNAs were hsa-miR-30d-5p, hsa-miR-30e-3p, and hsa-miR-29c-3p.(figure 1)



(临床信息与表达量之间的关系 table1)

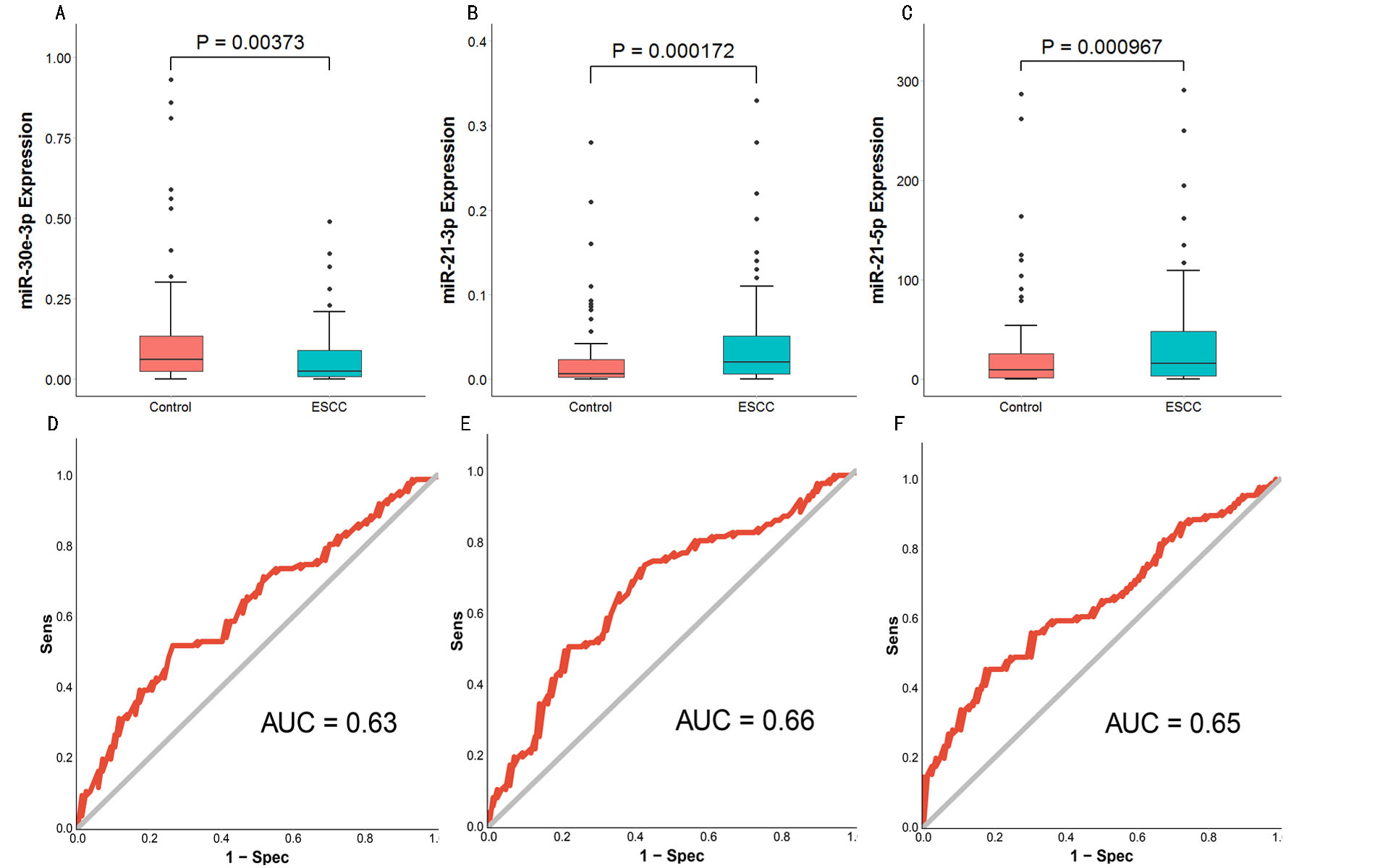
**Validation miRNAs by quantitative reverse transcription real-time PCR**

To confirm the expression of selected candidated miRNAs, we used qRT-PCR to evaluate the nine miRNAs expression of 93 pairs of ESCC and adjacent tissues. The ΔCt analysis method was used to identify differentially expressed miRNAs between ESCC and paracancer tissues. Only three miRNAs, miR-30e-3p, miR-21-3p and miR-21-5p, were significantly differentially expressed between ESCC and adjacent tissues(P = 0.004, 0.0002 and 0.001, respectively). The mean relative expression of miR-30e-3p, miR-21-3p and miR-21-5p in adjacent tissues were 0.12, 0.18 and 25.19, respectively. The mean relative expression of miR-21-3p and miR-21-5p in esophageal squamous cell carcinoma 0.072, 0.44 and 40.80. The trends of up or down of these three miRNAs were consistent with the results analyzed by TCGA data.

There was no difference in expression of the remaining 6 miRNAs.(figure 2)

**Diagnostic Utility of miRNAs in ESCC**

The diagnostic performance of the three miRNAs was evaluated by receiver operating characteristic (ROC) curves. Firstly, the diagnostic effect of single miRNA was calculated by logistic regression analysis. The sensitivity of miR-30e-3p, miR-21-3p and miR-21-5p in the diagnosis of esophageal squamous cell carcinoma were 0.52, 0.74 and 0.45, respectively, and the specificity was 0.74, 0.57 and 0.83, respectively. The corresponding areas under the ROC curve (AUCs) were 0.63, 0.66 and 0.65 for miR-30e-3p, miR-21-3p and miR-21-5p, respectively. In order to improve the diagnostic effect, we combined these three miRNA as a panel. A representation of the data using an ROC curve showed a strong separation between the ESCC and paracancerous groups, with an AUC of 0.80. (figure 3)



**Discussion**

ESCC is a common gastrointestinal cancer. Accurate judgments of ESCC are important bases for clinicians to take a rational approach to the treatment of this disease. The current clinical biomarkers for ESCC diagnosis are not ideal, as there remains a lack of reliable biomarkers that can specifically distinguish between ESCC patients and healthy people.

Many studies have found that miRNAs can act as biomarkers for the diagnosis, prognosis, and treatment evaluation of cancers. A review showed that four miRNAs (miR-25, -99a, -133a and -133b) showed good potential as diagnostic markers. In the present study, we quantified the expression levels of nine miRNAs by qRT-PCR and 3 miRNAs were finally selected. miR-30e-3p, a member of the miR-30 family and negatively regulates Ubc9 and Bmi1 expression and suppresses growth of different cancer cells PMID: 24945821. Lost expression of miR-30e-3p was associated with lymph node metastasis, a tumor size more than 3 cm, and advanced stages of NSCLC. miR-21-3p had been studied in some cancer, such as lung cancer, cervical cancer, ovarian cancer, and oral cancer. In ovarian cancer, miR-21-3p is overexpressed and binds to the 3-UTR of RBPMS, RCBTB1, and ZNF608 to effect the biological characteristics, which indicating that the miR-21-3p has the same biological effect in ESCC. There are many studies about miR-21-5p in cancers. MiR-21 is an oncogenic miRNA that modulates the expression of multiple cancer-related target genes such as PTEN, TPM1, and PDCD and has been shown to be overexpressed in various human tumors PMC3687369 . Although our observation of increased miR-21-5p expression in ESCC tissues is consistent with previous reports, our results are the first to combine the three miRNAs(miR-30e-3p, -21-3p and -21-5p) as a panel to diagnose the ESCC.The expression levels of three miRNAs demonstrated AUC values of 0.80 , associated with 62% sensitivity and 91% specificity, for discerning ESCC.

In conclusion, we demonstrated that 1) the exprssion of miR-30e-3pis significantly overexpressed in ESCC compared with adjacent normal tissues, but the exprssion of miR-21-3p and miR-21-5p in ESCC tissues are significantly lower than in normal tissues, 2) the three miRNAs (miR-30e-3p, -21-3p and -21-5p) can be as a panel for the detction of esophageal squamous cell carcinoma in solid biopsies. Further research is needed to investigate if these three markers could reliably be analyzed in liquid biopsies, so that it may be developed as low invasive, early detection biomarker for esophageal squamous cell carcinoma.

**Conclusion**

Combined with the TCGA data and our experimental results, we demonstrated that the miR-30e-3p, miR-21-3p, miR-21-5p, KCNA3 and KCNA6 could form a diagnosis panel for ESCC.

**Materials and Methods**

Biomarker discovery based on the public datasets. The miRNA expression data and the corresponding medical information of the patients were downloaded from the TCGA database, and there were a total of 200 samples (187 esophageal cancer samples and 13 normal samples) and the medical information was for 169 patients. These miRNA expression data were sequenced by Illumina HiSeq system and the standardized miRNA data were level 3 data in the TCGA database. Firstly, these miRNA data were standardized, and then the data with expression values zero were removed. The miRNA data in level 3 were downloaded, which included a total of 1046 comments for the miRNA expression values. These miRNA data in level 3 had already been standardized within the sample and the standardization between samples was performed using generalized linear model in an R language Limma package to  
eliminate batch eﬀects between samples.

**Patients and samples**

ESCC samples and the paired adjacent normal tissueswere obtained from xx. All patients were pathologically diagnosed as ESCC using surgical specimens or biopsies. None of the ESCC patients had received any anticancer treatment prior to sampling. All data, including age, sex, invasion depth (T stage), lymph node metastasis (N stage), distant metastasis (M stage), smoking and drinking habit were obtained from clinical or pathologic records. All tissues were immediately frozen at -80 oC after surgical resection.

Written informed consent was obtained from all patients prior to the study. The study protocol was approved by the ethics committees of Soochow University. The study was performed in accordance with the Declaration of 1975 Helsinki and the REMARK guidelines for biomarker studies.

**DNA and RNA isolation**

DNA was extracted from the tissues of 94 paired samples using AllPrep DNA/RNA Mini kit(QIAGEN,Germany) according to the manufacturer,s instructions. Total RNA was extracted from the tissues of 93 paired samples using Trizol(Invitrogen) according to the manufacturer’s instructions. The purity and concentration of DNA and RNA were quantified using Nanodrop 2000 thermo scientific spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

**qRT-PCR**

cDNAs were synthesized from total RNA using miDETECT A TrackTM miRNA qPCR Start Kit (Bio-Rad, Hercules, CA). Total RNA was polyadenylated and then reverse transcribed with the miR reverse anchor primer(PMID: 18834857). Quantitative real-time PCR was performed using SYBR Green Real Time PCR Kit (TAKARA, Japan) and MiDETECT A TrackTM miRNA qPCR Primer (Bio-Rad, Hercules, CA) on a Roche LightCycler480 (Roche Diagnostics, USA) according to the manufacturer,s protocols. U6 was used as an internal control. The relative expression was calculated using the 2-ΔCt method. All experiments were performed in triplicate.

Briefly, the polyadenylation reaction was in 10 μl of solution, containing 1.5 μg of extracted RNA, 2 μl of 5x poly(A) polymerase buffer, 1 μl of poly(A) polymerase and added RNAase-free H2O up to 10 μl. The mixtures were incubated at 37°C for 60 min. Next, RT reactions were conducted at 42℃ for 60 min followed by 70℃ for 10 min with the mixture containing 4ul of RTase mix, 4ul of 5x RTase buffer, 2ul of miDETECT A TrackTM Uni-RT Primer and 10ul of poly(A) Tailing product. Then 2 μl of 10-fold dilution cDNA was amplified using 5 μl of 2× SYBR Green Mix (TAKARA, Japan), 1 μl of gene-specific primers, and 2 μl of nuclease-free water in a final volume of 10 μl. The qRT-PCR reactions were carried out at 95℃ for 3 min, followed by 40 cycles of 95℃ for 10 sec and 60℃ for 34 sec. The melting curve analysis was added to assess the specificity of PCR products.

**Statistical analysi**s

Wilcoxon test was used to compare the miRNA expression and methylation status of each CpGsite of paired tissue samples. Further, differential methylation status in tumor and normal tissues of the candidate CpGsites were tested with logistic regression method. False discovery rate (FDR) correction was used for multiple test correction. ROC curves and the area under the ROC curve analyzed by logistic regression were used to evaluate the value of the identified miRNA and methylation status in detecting ESCC.

**Abbreviations**

**Acknowledgements**

**Competing interests**

**Authors’ contributions**

**Reference**