miR-486-3p, miR-139-5p, and miR-21 as Biomarkers for the Detection of Oral Tongue Squamous Cell Carcinoma.

Zujian Chen, University of Illinois at Chicago
Tianwei Yu, Emory University
Robert J. Cabay, University of Illinois at Chicago
Yi Jin, University of Illinois at Chicago
Ishrat Mahjabeen, University of Illinois at Chicago
Xianghong Luan, University of Illinois at Chicago
Lei Huang, University of Illinois at Chicago
Yang Dai, University of Illinois at Chicago
Xiaofeng Zhou, University of Illinois at Chicago

Journal Title: Biomarkers in Cancer
Volume: Volume 9
Publisher: Libertas Academica | 2017, Pages 1-8
Type of Work: Article | Final Publisher PDF
Permanent URL: https://pid.emory.edu/ark:/25593/rwmgb

Final published version:
http://insights.sagepub.com/journal-biomarkers-in-cancer-j154

Copyright information:
© 2017 the author(s), publisher and licensee Libertas Academica Ltd. This is an Open Access work distributed under the terms of the Creative Commons Attribution-Noncommercial 3.0 Unported License (http://creativecommons.org/licenses/by-nc/3.0/).

Accessed January 2, 2019 6:32 PM EST
miR-486-3p, miR-139-5p, and miR-21 as Biomarkers for the Detection of Oral Tongue Squamous Cell Carcinoma

Zujian Chen1, Tianwei Yu2, Robert J. Cabay3, Yi Jin1, Ishrat Mahjabeen1,4, Xianghong Luan5, Lei Huang6, Yang Dai6,7 and Xiaofeng Zhou1,7

1Center for Molecular Biology of Oral Diseases, Department of Periodontics, College of Dentistry, University of Illinois at Chicago, Chicago, IL, USA. 2Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, USA. 3Department of Pathology, College of Medicine, University of Illinois at Chicago, Chicago, IL, USA. 4Department of Biosciences, COMSATS Institute of Information and Technology, Islamabad, Pakistan. 5Department of Oral Biology, College of Dentistry, University of Illinois at Chicago, Chicago, IL, USA. 6Department of Bioengineering, College of Engineering, University of Illinois at Chicago, Chicago, IL, USA. 7UIC Cancer Center, Graduate College, University of Illinois at Chicago, Chicago, IL, USA.

ABSTRACT: Oral tongue squamous cell carcinoma (TSCC) is a complex disease with extensive genetic and epigenetic defects, including microRNA deregulation. The aim of the present study was to test the feasibility of performing the microRNA profiling analysis on archived TSCC specimens and to assess the potential diagnostic utility of the identified microRNA biomarkers for the detection of TSCC. TaqMan array-based microRNA profiling analysis was performed on 10 archived TSCC samples and their matching normal tissues. A panel of 12 differentially expressed microRNAs was identified. Eight of these differentially expressed microRNAs were validated in an independent sample set. A random forest (RF) classification model was built with miR-486-3p, miR-139-5p, and miR-21, and it was able to detect TSCC with a sensitivity of 100% and a specificity of 86.7% (overall error rate = 6.7%). As such, this study demonstrated the utility of the archived clinical specimens for microRNA biomarker discovery. The feasibility of using microRNA biomarkers (miR-486-3p, miR-139-5p, and miR-21) for the detection of TSCC was confirmed.

KEYWORDS: microRNA, biomarker, miR-486-3p, miR-139-5p, miR-21, formalin-fixed paraffin-embedded, tongue squamous cell carcinoma

Introduction

Head and neck/oral cancer (HNOC) is the sixth most common cancer worldwide, accounting for 4% of cancers in men and 2% of cancers in women,1 with an incidence of approximately 600,000 cases per year and a mortality rate of approximately 50%.2 In some parts of the world, including Southern China and the Indian subcontinent, head and neck squamous cell carcinoma (HNSCC) is still a major cancer problem. Tongue squamous cell carcinoma (TSCC) is significantly more aggressive compared to other HNOCs, with a propensity for rapid spreading and local invasion3 with a distinct pattern of lymph nodal metastasis,4,5 and high recurrence rate.6 Despite recent advances in clinical strategies for treating TSCCs, the overall survival has improved only marginally. This is because TSCCs are often detected at later stages. Advances in cancer screening strategies and early detection methods are required for improvement in the prognosis of TSCC patients.

Like most of the other human cancers, TSCC is a disease involving multistep dynamic changes in the genome. While many recent studies have attempted to identify molecular biomarkers for the screening and early diagnosis of TSCC, most of these studies are based on coding genes. The potential value of utilizing microRNAs as biomarkers for TSCC detection is still not entirely clear. MicroRNAs are endogenous noncoding single-stranded small RNA molecules (18–25 nt). A number of microRNA genes have recently been characterized as oncogenes or tumor-suppressor genes, and their deregulations have been detected in many cancers, including HNOCs.7–11 MicroRNAs are important gene expression regulators, and they control the expression of their target genes by posttranscriptional mechanisms. Deregulation of these cancerous microRNA genes (eg, overexpression or deletion) contributes to tumorigenesis by promoting proliferation, survival, and invasion.12,13 MicroRNA deregulation is a frequent event in HNSCC. A number of...
recent reports demonstrated the feasibility of utilizing micro-RNAs as biomarkers to detect cancer cases from noncancerous specimens, with varying degrees of success.\textsuperscript{14,15} This type of microRNA biomarker-based approach can enhance the standard diagnostic technique of histopathologic examination.

In this study, we identified a panel of differentially expressed microRNAs based on the TaqMan array microRNA profiling analysis of archived TSCC samples and normal matched tissue samples. Using a statistical model based on three microRNA biomarkers (miR-486-3p, miR-139-5p, and miR-21), we were able to identify TSCC cases from an independent validation set, with a sensitivity of 100% (15/15), a specificity of 86.7% (13/15), and an overall error rate of 6.7%.

**Patients and Methods**

**Patient cohorts.** The following two TSCC patient cohorts were used in this study: (1) the training set: specimens from 10 TSCC cases and the adjacent matched normal tissue samples were used for microRNA differential expression profiling analysis and (2) the validation set: microRNA expression data (deep sequencing-based profiling results) and the associated demographic and clinical information of 15 TSCC cases that have matched normal samples were obtained from The Cancer Genome Atlas repository (TCGA; https://tcga-data.nci.nih.gov). Demographic and clinical information of the patients is presented in Table 1. This investigation was approved by the Institutional Review Board (IRB) of the University of Illinois at Chicago.

**Laser-capture microdissection and RNA isolation.** Laser-capture microdissection (LCM) procedure was performed using the TaqMan low-density array (TLDA; Applied Biosystems), following the manufacturer’s protocol with minor modifications as previously described.\textsuperscript{18} In brief, 20 ng of RNA was used as the input for cDNA generation. Eight distinct pools of RT primers were used for analysis of 370 distinct microRNAs. Following dilution, 14 cycles of preamplification with the Megaplex pool protocol for the array were performed on the cDNA. Following dilution, the cDNAs were loaded onto the arrays (Human miRNA Array v1.0; Applied Biosystems). This facilitated analysis of 386 wells; 370 distinct microRNAs were analyzed in singlicate and two housekeeping snoRNAs with eight replicates for each. Individual TaqMan assays were also performed for selected microRNAs in triplicates for the validation study. To control for potential variations in RNA samples isolated from each case,\textsuperscript{19} we also assessed U6 snRNA with TaqMan assay (Thermo Fisher Scientific). The polymerase chain reaction (PCR) was performed on an ABI 7900HT real-time PCR system (Thermo Fisher Scientific). Ct (crosses threshold) values were determined for all samples and genes, and delta Ct (\(\Delta Ct\)) was computed using U6 snRNA as an internal control.\textsuperscript{20}

**Data analysis and statistical methods.** MicroRNA differential expression analysis was performed using Cyber-T,\textsuperscript{21,22} and hierarchical clustering and principal component (PC) analysis were performed using ClustVis.\textsuperscript{23} Other statistical analyses were performed using the S-Plus 6.0 and R 3.2.2. The differences between groups were evaluated by Wilcoxon signed-rank test. Receiver-operating characteristic (ROC) curve analysis was performed, and the area under the ROC (AUROC) was computed for assessing the predictive power of the selected biomarkers. To select the combination of biomarkers that provides the best prediction, the random forest (RF) classification model was utilized. The Mean Decrease Gini was computed to assess the relative importance of microRNA biomarkers toward an RF classification model. Since the structure of the data for the training set and validation set are different (\(\Delta Ct\) value for the quantitative PCR measurement, and reads per million miRNA mapped for the deep sequencing-based quantification, respectively), a simple transformation was performed on all data before the statistical analysis.
Results

LCM was performed on 10 archived TSCC samples and matched noncancerous tissue samples, and microRNA profiling analysis was performed using TaqMan arrays. As shown in Table 2 and Supplementary Table 1, we identified a panel of 12 differentially expressed microRNAs. Hierarchical clustering (Fig. 1A) and PC analysis (Fig. 1B) suggested that this panel of microRNAs could provide classification values distinguishing TSCC samples from normal tissue samples.

Of the 12 identified microRNAs, the differential expression was validated for 11 microRNAs (miR-486-3p, miR-21, miR-486-5p, miR-139-5p, miR-204, miR-489, miR-223, miR-196b, miR-31, miR-422a, and miR-146b-5p) using the individual TaqMan assays on the same sample set (Fig. 2). Differential expression of miR-328 was not validated ($P = 0.238$; Fig. 2K). To further validate the differential expression of the identified microRNAs, microRNA expression profiling results on 15 TSCC cases and the paired normal tissue samples were obtained from TCGA. Of the 12 microRNAs tested, the differential expression was validated for 8 microRNAs (miR-21, miR-486-5p, miR-139-5p, miR-204, miR-196b, miR-31, miR-422a, and miR-146b-5p) based on the TCGA dataset (Table 2 and Supplementary Table 2). Differential expression of miR-489, miR-223, and miR-328 was not validated. Apparent differential expression of miR-486-3p was also observed, but the difference is not statistically significant ($P = 0.0619$).

The ROC analysis was performed to assess the predictive powers of the identified microRNA biomarkers using both sample sets. The AUROC for miR-486-3p, miR-21, miR-486-5p, miR-139-5p, miR-204, miR-489, miR-223, miR-196b, miR-31, miR-422a, and miR-146b-5p were 0.9333, 0.9000, 0.9111, 0.9333, 0.8333, 0.8778, 0.8667, 0.8667, 0.7778, 0.8778, 0.8778, 0.8667, and 0.8000 for the training samples set, and 0.7022, 0.9911, 0.7778, 0.8756, 1.0000, 0.0074, 0.0018, 0.0009, 0.0001, 0.0003, 0.0021, and 0.0001 for the validation samples set.

Table 2. Receiver-operating characteristic (ROC) curve analysis of TSCC-associated microRNAs.

<table>
<thead>
<tr>
<th>microRNA</th>
<th>TRAINING SAMPLE SET</th>
<th>VALIDATION SAMPLE SET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WILCOXON (P VALUE)</td>
<td>ROC (AUC)</td>
</tr>
<tr>
<td>miR-486-3p</td>
<td>0.0006</td>
<td>0.9333</td>
</tr>
<tr>
<td>miR-21</td>
<td>0.0021</td>
<td>0.9000</td>
</tr>
<tr>
<td>miR-486-5p</td>
<td>0.0015</td>
<td>0.9111</td>
</tr>
<tr>
<td>miR-139-5p</td>
<td>0.0006</td>
<td>0.9333</td>
</tr>
<tr>
<td>miR-204</td>
<td>0.0133</td>
<td>0.8333</td>
</tr>
<tr>
<td>miR-489</td>
<td>0.0041</td>
<td>0.8778</td>
</tr>
<tr>
<td>miR-223</td>
<td>0.0057</td>
<td>0.8667</td>
</tr>
<tr>
<td>miR-196b</td>
<td>0.0057</td>
<td>0.8667</td>
</tr>
<tr>
<td>miR-31</td>
<td>0.0435</td>
<td>0.7778</td>
</tr>
<tr>
<td>miR-422a</td>
<td>0.0042</td>
<td>0.8778</td>
</tr>
<tr>
<td>miR-328</td>
<td>0.2428</td>
<td>0.6667</td>
</tr>
<tr>
<td>miR-146b-5p</td>
<td>0.0279</td>
<td>0.8000</td>
</tr>
</tbody>
</table>

Notes: *The microRNA data for the training sample set and validation sample set were assessed with different platforms (TaqMan assay and deep sequencing, respectively). To enable the comparison between these datasets, transformation was performed as described in “Patients and methods” section. *The miRSeq dataset for 15 TSCC and paired normal tissue samples was downloaded from TCGA data portal. The levels of microRNAs were extracted as reads per million miRNA mapped.

modeling $[(x - \mu_{\text{normal}})/\sigma_{\text{normal}}]$ where $\mu_{\text{normal}}$ is the mean of the normal group and $\sigma_{\text{normal}}$ is the standard deviation of the normal group.

Figure 1. MicroRNAs profiling on TSCC samples. Laser-capture microdissection was performed to acquire tumor cells from 10 cases of archived TSCC samples and matched normal samples. MicroRNA profiling was performed on these samples using TaqMan microRNA arrays. A signature gene set of 12 microRNAs was created as described in the “Patients and methods” section. Hierarchical clustering (A) and principal component (PC) analysis (B) were performed based on this signature set.
0.4844, 0.6756, 0.9822, 0.7956, 0.6333, 0.6178, and 0.8667 for the validation sample set (TCGA), respectively (Table 2). The RF model was built for selecting the best biomarker combination that can provide the highest prediction power. The Gini importance values were computed to evaluate the relative importance of the biomarkers, and miR-486-3p, miR-139-5p, and miR-21 have the highest relative importance (Fig. 3). The RF classification model based on these top three microRNA biomarkers was used to classify normal and TSCC cases of the validation sample set (n = 30).

As shown in Table 3, the overall error rate of this classification model is 6.7% [sensitivity = 100% (15/15) and specificity = 86.7% (13/15)].

**Discussion**

While mechanistic studies of the roles of microRNA in tumorigenesis are gaining momentum, translational studies utilizing microRNAs as biomarkers are still in their infancy. Nevertheless, microRNAs are remarkably stable in cells, tissue specimens (archived or fresh), and a number of biofluids, and as such, microRNA-based biomarkers are less prone to minor differences in sample processing, which offers a great advantage in clinical applications. The current study demonstrated the feasibility of using a panel of microRNAs as biomarkers for the detection of TSCC, with promising results that warrant further validation in larger, prospective studies.
Recent studies indicated that quantitative PCR and prostate cancer.

Downregulation of protein–microRNA complexes may prevent microRNA from degradation, especially during the formalin fixation process; the entire microRNA molecule is incorporated into the RNA-induced silencing complex (RISC).

It has been shown that RNA stability varies among different RNA species and that microRNAs are very sensitive to different methods of RNA extraction and purification. MicroRNA quantification, even with suboptimal starting materials, such as FFPE tissue blocks, has been suggested to have potential diagnostic value for lung cancer, and prognostic values for patients with esophageal squamous cell carcinoma (ESCC) or gastric adenocarcinoma (GC). MicroRNA-21 is one of the most well-documented oncogenic microRNAs and has been suggested as a biomarker for many types of cancer, including HNOC, Downregulation of miR-139 has been observed in HNOC, and a recent study suggested that miR-139-5p from saliva samples may be a molecular biomarker for the early diagnosis of TSCC. While miR-204 has been suggested as a tumor suppressor in HNOC, conflicting results regarding its expression pattern have been reported in various cancer types. The enhanced expression of miR-204 was observed in insulinas and acute lymphocytic leukemia; downregulation of miR-204 was reported in other cancer types, including HNOC, lung cancer, gastric cancer, endometrial cancer, and renal cancer. Furthermore, conflicting results regarding the miR-204 expression were reported in some other cancer types including breast cancer and prostate cancer. Our results here confirmed that miR-204 is downregulated in TSCC, which supports its tumor-suppressor role. Relatively little is known on the role of miR-489 in cancer, and conflicting results regarding the aberrant expression pattern of miR-489 have been reported in HNOC. We observed downregulation of miR-489 in our training set, but no statistical significant change in miR-489 level in the validation set. Thus, the observed differential expression of miR-489 in HNOC remains controversial, and additional studies with large sample sizes will be needed to fully explore this. Upregulation of miR-223 has been consistently observed in both tumor tissues and plasma samples of HNOC patients and the diagnostic value of circulating miR-223 as a biomarker for HNOC has recently been suggested. MicroRNA-196b (together with miR-196a, the other member of miR-196 family) has been suggested as an OncomiR in HNOC, which promotes invasive phenotype and metastasis and may serve as a biomarker for early detection and prognosis. The role of miR-31 appears to be cancer type specific; although miR-31 inhibits metastasis in breast cancer, upregulation

**Figure 3.** Ranking the microRNA biomarkers by random forest model. The relative importance of microRNA biomarkers toward a random forest classification model was assessed by computing the Gini importance, and the microRNAs were ranked by Mean Decrease Gini.

<table>
<thead>
<tr>
<th>Prediction</th>
<th>NORMAL</th>
<th>TSCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>TSCC</td>
<td>2</td>
<td>15</td>
</tr>
</tbody>
</table>

**Table 3.** Random forest classification model for TSCC prediction.

**Notes:** Random forest classification model based on top three microRNA biomarkers (miR-486-3p, miR-139-5p, and miR-21) was used to predict normal and TSCC cases of the validation sample set (n = 30). The sensitivity of this classification model is 100% (15/15), and the specificity is 86.7% (13/15).
of miR-31 is essential to the TGF-beta-induced invasion and metastasis of colon cancer.73 The increases of miR-31 in tumor tissue, saliva, and plasma samples have all been suggested as potential biomarkers of HNOC.74-76 MicroRNA-422a has been suggested as a tumor suppressor, and the reinduction of miR-422a to cancer cells led to inhibition in cell proliferation, migration, invasion, metastasis, and enhanced chemosensitivity in various cancers.77-79 A recent study suggested that miR-422a deregulation may promote locoregional recurrence of HNOC.80 The potential of utilizing miR-422a as a biomarker has been suggested for several cancer types, including colorectal cancer and osteosarcoma.81-84 Several studies have suggested that miR-328 may serve as a biomarker for glioma, thyroid cancer, and non-small cell lung cancer (NSCLC).85-89 However, the differential expression of miR-328 was not validated in our training set, and no statistical significant change of miR-328 was observed in our validation set. The expression of miR-146b-5p appears to be cancer type specific. In glioma, miR-146b-5p expression is downregulated and acts as a tumor suppressor.90 However, in thyroid cancer and lung cancer, miR-146b-5p acts as an oncogene and has been suggested as a potential diagnostic marker.91,92

Our statistical analysis indicated that miR-486-3p, miR-139-5p, and miR-21 have the best classification power in terms of discriminating TSCC from normal control samples. We constructed an RF model by combining miR-486-3p, miR-139-5p, and miR-21, and we were able to achieve an excellent classification outcome (sensitivity = 100%, specificity = 86.7%, and overall error rate = 6.7%). These results clearly indicated that microRNA molecules acquired from archived TSCC specimens can be accurately measured and employed as molecular biomarkers for the early detection of TSCC. However, our sample size is relatively small. Additional studies with larger sample sizes are required to fully evaluate these microRNA biomarkers in detecting TSCCs.

MicroRNA deregulation has also been identified in premalignant lesions57,59,94 and in the field of cancerization.95-98 Thus, microRNA biomarker-based approaches can potentially be implemented as a cancer-screening tool for monitoring the oral premalignant lesions and the field of cancerization. In our future study, we anticipate to expend our investigation by utilizing cancer tissue, premalignant lesions, and histologically normal tissue samples adjacent to cancer, as well as other clinical samples (e.g., blood, serum, and/or saliva) from the same TSCC patients. This will allow us to fully explore the feasibility of utilizing microRNA biomarkers for cancer screening.

Conclusion
Our study demonstrated that microRNA deregulation can be accurately measured in archived clinical specimens and can be employed as molecular biomarkers for detecting TSCC. While utilizing microRNAs from snap-frozen specimens as biomarkers can produce better overall error rate,18 in the present study, microRNA biomarkers from archived clinical specimens also resulted in excellent levels of sensitivity and specificity. A specific combination of microRNA biomarkers (miR-486-3p, miR-139-5p, and miR-21) can achieve optimal outcomes in distinguishing TSCC from noncancerous tissue samples. Advances in in situ hybridization (ISH) probes for microRNA detection present us with potential strategies of integrating microRNA-based analysis with histopathologic examination,99-101 which may lead to improvement of early detection and prevention of oral cancer. One of the main limitations of our study is the small sample size. It is possible that there are still other microRNA biomarkers yet to be identified. Additional studies consisting of early-stage TSCC cases with larger sample sizes are required to fully evaluate the feasibility of this microRNA-based approach for early detection.

Author Contributions
Conceived and designed the experiments: RJC, YJ, XL, YD, XZ. Performed experiments: ZC, YJ, IM. Analyzed the data: TY, LH, YD, XZ. Wrote the first draft of the manuscript: TY, RJC, YD, XZ. Contributed to the writing of the manuscript: ZC, YJ, IM, XL, LH. Agree with manuscript results and conclusions: ZC, TY, RJC, YJ, IM, XL, LH, YD, XZ. Made critical revisions and approved final version: TY, RJC, YD. All authors reviewed and approved of the final manuscript.

Supplementary Materials
Supplementary table S1. The levels of 12 microRNAs on 10 cases of OTSCC and paired normal tissues.
Supplementary table S2. The levels of 12 microRNAs on 15 paired OTSCC and normal tissues (TCGA dataset).

REFERENCES
MicroRNA biomarkers for detecting tongue squamous cell carcinoma