Comprehensive analysis of lncRNA-associated competing endogenous RNA network in tongue squamous cell carcinoma

Shusen Zhang 1,2, Ruoyan Cao 1, Qiulan Li 3, Mianfeng Yao 4, Yu Chen 1, Hongbo Zhou Corresp. 1

1 Department of Prosthodontics, Xiangya Stomatological Hospital, Xiangya School of Stomatology, Central South University, Changsha, China
2 Department of Stomatology, Hunan University of Medicine, Hunan, China
3 Department of Oral Medicine, The Second Xiangya Hospital, Central South University, Changsha, China
4 Department of Oral Medicine, Xiangya Hospital, Central South University, Changsha, China

Corresponding Author: Hongbo Zhou
Email address: zhb2540@csu.edu.cn

Background. Increasing evidence has demonstrated that long non-coding RNAs (lncRNAs) play an important role in the competitive endogenous RNA (ceRNA) networks in that they regulate protein-coding gene expression by sponging microRNAs (miRNAs). However, the understanding of the ceRNA network in tongue squamous cell carcinoma (TSCC) remains limited.

Methods. Expression profile data regarding mRNAs, miRNAs and lncRNAs as well as clinical information on 122 TSCC tissues and 15 normal controls from The Cancer Genome Atlas (TCGA) database were collected. We used the edgR package to identify differentially expressed mRNAs (DEmRNAs), lncRNAs (DElncRNAs) and miRNAs (DEmiRNAs) between TSCC samples and normal samples. In order to explore the functions of DEmRNAs, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was performed. Subsequently, a ceRNA network was established based on the identified DElncRNAs–DEmiRNAs and DEmiRNAs–DEmRNAs interactions. The RNAs within the ceRNA network were analyzed for their correlation with overall disease survival. Finally, lncRNAs were specifically analyzed for their correlation with clinical features in the included TSCC patient samples.

Results. A total of 1867 mRNAs, 828 lncRNAs and 81 miRNAs were identified as differentially expressed in TSCC tissues (|log2 fold change| ≥ 2; adjusted P value < 0.01). The resulting ceRNA network included 16 mRNAs, 56 lncRNAs and 6 miRNAs. Six out of the 56 lncRNAs were found to be associated with the overall survival in TSCC patients (P < 0.05); 10 lncRNAs were correlated with TSCC progression (P < 0.05).

Conclusion. Our study deepens the understanding of ceRNA network regulatory mechanisms in TSCC. Furthermore, we identified ten lncRNAs (PART1, LINC00261, AL163952.1, C2orf48, FAM87A, LINC00052, LINC00472, STEAP3-AS1, TSPEAR-AS1 and ERVH48-1) as novel, potential prognostic biomarkers and therapeutic targets for TSCC.
Comprehensive analysis of lncRNA-associated competing endogenous RNA network in tongue squamous cell carcinoma

Shusen Zhang1,2*, Ruoyan Cao1*, Qiulan Li3, Mianfeng Yao4, Yu Chen1, Hongbo Zhou1

1 Department of Prosthodontics, Xiangya Stomatological Hospital, Xiangya School of Stomatology, Central South University, 72 Xiangya Road, Changsha, 410000, China
2 Department of Stomatology, Hunan University of Medicine, Huaihua 418000, Hunan, China
3 Department of Oral Medicine, The Second Xiangya Hospital, Central South University, 139 Middle Renmin Road, Changsha, 410011, China
4 Department of Oral Medicine, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, 410000, China

* These authors contributed equally to this work.

Corresponding Author:
Dr. Hongbo Zhou, PhD, DDS. Department of Prosthodontics, Xiangya Stomatological Hospital, Xiangya School of Stomatology, Central South University, 72 Xiangya Road, Changsha, 410083, China

Email: zhb2540@csu.edu.cn
Abstract

Background. Increasing evidence has demonstrated that long non-coding RNAs (lncRNAs) play an important role in the competitive endogenous RNA (ceRNA) networks in that they regulate protein-coding gene expression by sponging microRNAs (miRNAs). However, the understanding of the ceRNA network in tongue squamous cell carcinoma (TSCC) remains limited.

Methods. Expression profile data regarding mRNAs, miRNAs and lncRNAs as well as clinical information on 122 TSCC tissues and 15 normal controls from The Cancer Genome Atlas (TCGA) database were collected. We used the edgR package to identify differentially expressed mRNAs (DEmRNAs), lncRNAs (DElncRNAs) and miRNAs (DEmiRNAs) between TSCC samples and normal samples. In order to explore the functions of DEmRNAs, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was performed. Subsequently, a ceRNA network was established based on the identified DElncRNAs–DEmiRNAs and DEmiRNAs–DEmRNAs interactions. The RNAs within the ceRNA network were analyzed for their correlation with overall disease survival. Finally, lncRNAs were specifically analyzed for their correlation with clinical features in the included TSCC patient samples.

Results. A total of 1867 mRNAs, 828 lncRNAs and 81 miRNAs were identified as differentially expressed in TSCC tissues (|log₂fold change| ≥ 2; adjusted P value < 0.01). The resulting ceRNA network included 16 mRNAs, 56 lncRNAs and 6 miRNAs. Six out of the 56 lncRNAs were found to be associated with the overall survival in TSCC patients (P < 0.05); 10 lncRNAs were correlated with TSCC progression (P < 0.05).

Conclusion. Our study deepens the understanding of ceRNA network regulatory mechanisms in
Furthermore, we identified ten lncRNAs (PART1, LINC00261, AL163952.1, C2orf48, FAM87A, LINC00052, LINC00472, STEAR-AS1, TSPEAR-AS1 and ERVH48-1) as novel, potential prognostic biomarkers and therapeutic targets for TSCC.

Introduction

Tongue squamous cell carcinoma (TSCC) is the most common type of oral squamous cell carcinoma (OSCC) with remarkable invasiveness, early lymph node metastasis and a poor prognosis (Sano & Myers 2007; Zhou et al. 2015). The quality of life for TSCC survivors is often reduced due to speech dysfunction, mastication, and deglutition. Recently, the incidence and mortality of TSCC have steadily risen in the United States (Siegel et al. 2014; Siegel et al. 2015; Siegel et al. 2016; Siegel et al. 2017; Siegel et al. 2018). Despite significant advancements in surgical excision, radiotherapy and chemotherapy, mortality rates and recurrence rates for this cancer remain high (Adeel & Suhail 2016; Chen et al. 2018). For these reasons, the molecular mechanisms of TSCC tumorigenesis urgently require further study; potential biomarkers as well as therapeutic targets in this cancer should be identified in order to improve clinical outcomes.

Long noncoding RNAs (lncRNAs) are a subclass of noncoding RNAs longer than 200 nucleotides (Ponting et al. 2009). Of late, lncRNAs have been a new focus of cancer research and were found to be involved in tumorigenesis and metastasis (Chen et al. 2017; Gutschner et al. 2013; Yang et al. 2017; Yang et al. 2016b; Yuan et al. 2014). In TSCC, lncRNAs were reported to act as oncogenes or tumor suppressors and affect patient prognosis. For example, knockdown of AFAP1-
AS1 could suppress cell proliferation, migration and invasion in TSCC (Wang et al. 2018). Additionally, the overexpression of IncRNA MEG3 inhibited cell proliferation and induced apoptosis in TSCC (Jia et al. 2014). The current literature has demonstrated that IncRNAs regulate gene expression via genetic imprinting, splicing regulation, chromatin remodeling, mRNA decay, and translational regulation (Zhu et al. 2013). However, the formation and development of tumors is a complex pathophysiological process. The mechanisms by which IncRNAs affect TSCC biology remain unelucidated.

Salmena et al. proposed the competing endogenous RNA (ceRNA) hypothesis, which stated that IncRNA could crosstalk with mRNA by sharing common microRNA response elements (MREs) with miRNA (Salmena et al. 2011). More and more studies have validated the involvement of ceRNA crosstalk in the development and progression of various tumors, such as those of breast cancer, hepatocellular cancer and pancreatic cancer. Some of the few relevant studies on such crosstalk in TSCC have been verified, such as LINC00511/ miR-765/ LAMC2 (Ding et al. 2018) and H19/ let-7a/ HMGA2 (Kou et al. 2018). Furthermore, comprehensive analysis of TSCC-associated IncRNAs and miRNAs in a whole genome wide context is lacking, especially based on high-throughput sequencing with a large-scale sample size.

To better understand how IncRNAs regulate gene expression by sponging miRNAs in TSCC, we build a ceRNA network based on the TCGA database, including 6 mRNAs, 56 IncRNAs and 6 miRNAs. In addition, we found 10 IncRNAs to be associated with survival and 10 IncRNAs having
an association with carcinogenesis. Results of these analyses are a starting point to analyze ceRNA
crosstalk and gain insight into the molecular mechanisms participating in the tumorigenesis and
progression of TSCC.

Materials and Methods

Patients and samples

RNA sequencing data and the corresponding clinical information for our TSCC dataset were
retrieved from the TCGA data portal. The inclusion criteria were set as follows: 1) patients with
follow-up survival times less than 2000 days; 2) patients with detailed clinicopathological
information including age, gender, survival time, survival status, pathological stage, TNM stage.
As most patients were missing data about their metastatic states, we did not analyze this
information. After filtering available data with our inclusion criteria, a total of 122 TSCC patients
and 15 normal controls were included in our analysis. The clinical and pathological characteristics
of the TSCC patients are summarized in Table 1. This study conformed with the publication
guidelines provided by TCGA (https://cancergenome.nih.gov/publications/publicationguidelines)
and as our data was obtained from TCGA database, approval by an ethics committee was not
required.

RNA sequence data processing

Level 3 RNASeq and miRNASeq data from TSCC samples up to June 30,2018, including 122
TSCC tissues and 15 normal controls, were downloaded from the TCGA data portal. The sequence
data originated from IlluminaHiSeq_RNASeq and IlluminaHiSeq_miRNASeq sequencing platforms; all the data are publicly available.

**Analysis of differential expression profiles**

The Ensembl database (http://www.ensembl.org/index.html, version 89) (Aken et al. 2016) was used to identify lncRNA genes from the raw expression data. We discarded previously identified lncRNAs that were not included in this database. Differential expression analysis of mRNAs (DEmRNAs), miRNAs (DEmiRNAs) and lncRNAs (DElncRNAs) between TSCC and normal tissues was carried out using the edgeR package (Robinson et al. 2010). For all p values, false discovery rate (FDR) was applied for multiple testing correction. Absolute log$_2$FC $\geq$ 2 and the FDR < 0.01 were used as cut-off criteria.

**Functional enrichment analysis**

In order to better understand the mechanisms involved in the tumorigenesis of TSCC, we conducted Gene Ontology (GO) functional enrichment analysis using DAVID (the Database for Annotation, Visualization and Integrated Discovery) with FDR < 0.01 as the cut-off value. KEGG analysis was performed using the ClusterProfiler package in the R language with a cut-off value of adjusted p value < 0.05.

**Construction of ceRNA network**

We used the miRcode database (Jeggari et al. 2012) to predict lncRNA-miRNA interactions, which
were then combined with selected miRNAs. Secondly, TargetScan (Fromm et al. 2015), miRTarBase (Chou et al. 2016) and miRDB (Wong & Wang 2015) were used to retrieve and predict the targeted mRNAs of miRNA. In order to enhance the validity of this ceRNA network, we only included miRNA-targeted mRNAs present in all three databases and DEmRNAs. Finally, the ceRNA network was visualized using Cytoscape 3.6.1 software. A flowchart of the ceRNA network is presented in Fig 1. We also performed linear regression analysis to explore the correlation between ceRNA expression levels.

Statistical analysis

For overall survival analysis, the log-rank test was employed to compare the difference between TSCC and normal samples using Kaplan Meier survival curve. The cut-off point of expression was identified using survminer package (Li et al. 2018b). The edgeR package was used to screen out DElncRNAs associated with clinical features, by setting cut-off criteria of absolute log$_2$FC $\geq$ 1 and the FDR < 0.05. Unless specifically stated, a p value < 0.05 was considered to represent statistical significance. All statistical analyses were performed using R software (version: 3.3.2).

Results

DEmRNAs in TSCC

According to the cut-off threshold of log$_2$FC $\geq$ 2 and FDR < 0.01, 717 (38.40%) up-regulated and 1150 (61.60%) down-regulated genes were identified in TSCC (Supplementary Table 1). Fig. 2 shows the distribution of DEmRNAs between TSCC and normal controls. The expression heat
map of DEmRNAs is shown in Supplementary Fig. 1. Red or green represents significantly upregulated and downregulated genes, respectively.

A total of 45 significantly enriched GO terms are listed in Supplementary Table 2 that correspond to DEmRNAs. For “biological processes (BP)”, the top five terms were muscle filament sliding, collagen catabolic process, extracellular matrix organization, muscle contraction and skeletal system development; for the “cellular component (CC)” ontology the top five were, extracellular region, extracellular space, proteinaceous extracellular matrix, Z disc and collagen trimer; finally, the top five “molecular function (MF)” terms were, structural constituent of muscle, extracellular matrix structural constituent, calcium ion binding, heparin binding and cytokine activity (Fig. 3).

Additionally, a total of 20 significantly enriched KEGG pathways for the identified DEmRNAs are listed in Table 2, and the top 10 KEGG pathways are shown in Fig. 3D. The cytokine-cytokine receptor interaction pathway was found to harbor the largest number of DEmRNAs (Fig. 3E).

**DElncRNAs in TSCC**

Based on the cut-off criteria (log2FC ≥ 2 and FDR < 0.01), we identified 828 lncRNAs aberrantly expressed in TSCC compared to normal tissues, including 517 up-regulated (62.44%) and 311 down-regulated IncRNAs (37.56%) (Supplementary Table 3). The distribution of all the DElncRNAs are presented as a volcano plot in Fig.4 and an expression heat map of DElncRNAs is shown in Supplementary Fig. 2.
DEmiRNAs in TSCC

To build our lncRNA-miRNA-mRNA ceRNA network, we also compared miRNA expression profiles in tumor tissues with normal tissues. In total, 81 DEmiRNAs were identified, including 42 up- and 39 down-regulated miRNAs (Supplementary Table 4). A volcano plot of the related DEmiRNAs is shown in Fig. 5; a corresponding expression heat map is shown in Supplementary Fig. 3.

ceRNA network in TSCC

A dysregulated ceRNA network of lncRNA-miRNA-mRNA in TSCC was established based on the above data in order to better elucidate the role of DElncRNAs. First, the 828 DElncRNAs were retrieved from the miRcode, and 102 pairs of interacting lncRNAs and miRNAs were identified using the Perl language. Subsequently, we predicted that six DEmiRNAs could interact with 56 DElncRNAs. Then we found that these six DEmiRNAs targeted 221 mRNAs in all three databases (TargetScan, miRTarBase and miRDB). Among the 221 targeted mRNAs, only 16 mRNAs were found in the 1867 DEmRNAs (Supplementary Fig. 4). Finally, we constructed a ceRNA network relating to TSCC by incorporating 56 DElncRNAs, 6 DEmiRNAs and 16 DEmRNAs, as shown in Fig. 6. To confirm these findings, we performed smooth curve fitting between the expression levels of the DElncRNAs and DEmRNAs included in the ceRNA network. Our results indicated a positive correlation between ceRNA expression levels. For example, LINC00472 interacted with GREM2 mediated by mir-503 and SFTAIP regulated IL11 levels by sponging mir-211 (Fig. 7).
We also contract GO and KEGG analysis to reveal the functions of the 16 DEmRNAs that were involved in the ceRNA network. Only two GO terms were significantly enriched ($P < 0.05$) (Table 3).

RNAs in the ceRNA network are related to survival

LncRNAs, miRNAs and mRNAs associated with prognosis were identified using the expression profiles of 59 lncRNAs, 6 miRNAs and 16 mRNAs in the ceRNA network using Kaplan Meier Survival Curve. As a result, ten lncRNAs (PART1, LINC00261, AL163952.1, C2orf48, FAM87A, LINC00052, LINC00472, STEAR-AS1, TSPEAR-AS1 and ERVH48-1) were observed to be significantly related to overall survival rate ($P < 0.05$) (Fig. 8).

IncRNAs in the ceRNA network are related to clinical features

The 64 DEIncRNAs from the ceRNA network were further analyzed to identify their correlations with clinical features. TSCC patients were divided into subgroups according to pathological stage (Stage III + IV vs. stage I + II) and TNM stage (T3 + T4 vs. T1 + T2, N2 + N3 vs. N0 + N1). We found six IncRNAs with a high expression level (LINC00355, PSORS1C3, LINC00520, AC112721.1, AL139147.1, SFTA1P) and four IncRNAs with a low expression level (HCG22, LINC00492, AL035696.1, ERVH48-1) were significantly associated with the progression of TSCC (Table 4).

Discussion
TSCC is the most common form of oral cancer. Dysregulated genes are considered a major cause of oncogenesis and the development of TSCC. Recently, the crucial role of lncRNA in gene expression regulation at three levels including transcription, post-transcription and translation has attracted considerable interest. Accordingly, the ceRNA hypothesis was proposed, postulating that lncRNAs could act as part of post-transcriptional gene expression control. This conclusion generated new insights into the biology of cancer.

To better understand how lncRNA-associated ceRNA crosstalk affects TSCC, we exploited a large-scale TSCC data from the TCGA database and successfully established a dysregulated lncRNA-associated ceRNA network. In addition, growing evidence has indicated that lncRNAs have greater potential as prognostic biomarkers than protein-coding genes due to their stronger correlation with tumor status (Hauptman & Glavac 2013). Thus, we also identified ten lncRNAs (PART1, LINC00261, AL163952.1, C2orf48, FAM87A, LINC00052, LINC00472, STEAR-AS1, TSPEAR-AS1, and ERVH48-1) as prognostic biomarkers for TSCC. However, there is no research to clearly explain the function of AL163952.1, C2orf48, FAM87A, STEAR-AS1, TSPEAR-AS1, and ERVH48-1.

PART1 is upregulated and its higher expression is associated with poor prognosis in prostate cancer and non-small cell lung cancer (Li et al. 2017b; Sun et al. 2018). Elevated PART1 promotes prostate cancer cell proliferation and inhibits cell apoptosis (Sun et al. 2018). On the contrary, we found that PART1 had lower expression and was negatively correlated to survival rate in TSCC,
which was consistent with findings in OSCC (Li et al. 2017c). This may be because PART1 is located on chromosome 5q12, a region that is usually lost in oral squamous cell carcinoma (OSCC) and head and neck squamous cell carcinoma (HNSCC) (Abou-Elhamd & Habib 2008; Noutomi et al. 2006). In our ceRNA network, low PART1 expression reduced levels of NR3C2 mediated by mir-301b, and reduced expression of NR3C2 promotes tumor cell proliferation, metastasis and epithelial-to-mesenchymal transition (Yang et al. 2018; Yang et al. 2016a; Zhang et al. 2017; Zhao et al. 2018c). In addition, patients with low levels of NR3C2 have a poor prognosis in pancreatic cancer and renal cell carcinoma (Yang et al. 2016a; Zhao et al. 2018c). The relationship between mir-301b and NR3C2 has also been validated in pancreatic cancer (Yang et al. 2016a). Therefore, PART1/mir-301b/NR3C2 axis may be an important mechanism that involves in TSCC development.

LINC00261 is a tumor suppressor positively associated with prognosis in many tumors, such as hepatocellular carcinoma (Zhang et al. 2018a), endometrial carcinoma (Fang et al. 2018) and non-small cell lung (Liu et al. 2017). Its functions mainly include inhibiting tumor cell proliferation, invasion and metastasis. LINC00261 was also down-regulated in our study. Notably, decreased expression of LINC00261 indicated a better prognosis in TSCC. Our study found LINC00261 regulated the expression of ENPP4 and ENPP5, however, these two genes have not been extensively studied. ENPP2 as their closely-related molecule has been widely reported to participate in tumor development. Low expression of ENPP2 increases reactive oxygen species (ROS) level, and high ROS level could promote tumor cell apoptosis (Cholia et al. 2018; Dawei
et al. 2018). The special function of LINC00261 still needs to be further investigated.

Several existing studies indicate that LINC00472 plays an important role in inhibiting tumor development (Shen et al. 2015; Su et al. 2018; Ye et al. 2018). Our study also showed the similar results that higher expression of LINC00472 was associated with better prognosis. In addition, LINC00472 may regulate the expression of GREM2 by sponging mir-503. GREM2 is an antagonist of bone morphogenetic proteins (BMP) and could activate Notch signaling pathway (Li et al. 2018a) and Wnt/β-catenin signaling (Wu et al. 2015), which may be the reasons that GREM2 involves in the development of TSCC. Given that GREM2’s functions have not been studied in cancer, our results need to be verified by experiments.

The function of LINC00052 varies depending on the location of tumors. It acts as a tumor suppressor through inhibiting cell proliferation, invasion and migration in hepatocellular carcinoma (Xiong et al. 2016; Yan et al. 2018; Zhu et al. 2017), whereas it could promote breast cancer growth (Salameh et al. 2017) and gastric cell metastasis and proliferation (Shan et al. 2017). In our study, it may act as an oncogene because it is significantly up-regulated and its high expression indicates a poor prognosis in TSCC. Unfortunately, our ceRNA network failed to find miRNAs that could interact with LINC00052. The function of LINC00052 is needed to be further investigated.

Considering the correlation between DElncRNAs and clinical characteristics, we found 10
DElncRNAs related to pathologic stage, T stage and N stage. Though the function of these lncRNAs are not well investigated, they also may be as therapeutic targets and present a new road to understand the pathogenesis of TSCC. HCG22 was negatively associated with tumor stage in our study. Similarly, Zhao et al. finds a significant inverse correlation between HCG22 and tumor size (Zhao et al. 2018a). Regarding OSCC, HCG22 was found to be correlated with poor survival basing on TCGA database. However, Feng et al. did not find a similar association using their clinical data (Feng et al. 2017). LINC00355 was positively correlated with distant metastasis, lymphatic metastasis and tumor stage, and negatively correlated with prognosis in colon adenocarcinoma (Zhang et al. 2018b). Upregulated LINC00355 was also associated with poor prognosis in prostate cancer (Jiang et al. 2018). In our study, LINC00355 was positively associated with tumor stage, however, it was not identified to be a prognosis biomarker.

Another DElncRNA, SFTA1P has been reported to be a tumor suppressor by inhibiting cell proliferation, invasion and migration in gastric cancer (Ma et al. 2018). It also increases cisplatin chemosensitivity in lung squamous cell carcinoma; similarly, elevated SFTA1P indicates a longer life (Li et al. 2017a). However, we found SFTA1P was up-regulated in TSCC, especially in lymph node metastasis tumor. Thus, we speculate that SFTA1P may act as an oncogene. Our ceRNA network indicated that SFTA1P up-regulated the expressions of IL-11 or HOXC8 by binding miR-211, and elevated IL-11 or HOXC8 contributes to development of cancer, such as breast cancer (Cai et al. 2018; Li et al. 2014) and non-small cell lung cancer (Liu et al. 2018; Zhao et al. 2018b). Therefore, it is reasonable that our results are not consistent with other studies. Though the
functions of IL-11 and HOXC8 are not validated in TSCC, our results also offer new ideas for the
development of TSCC.

Our study identified some valuable lncRNAs that are associated with carcinogenesis and survival. Few of them have been validated in vitro and in vivo, however, none of the lncRNAs were validated in TSCC. Hence, these valuable lncRNAs still need to be verified, and our ceRNA network, which was build based on high-throughput sequencing, requires further verification.

Conclusion

Taken together, we identified aberrantly expressed mRNAs, lncRNAs, and miRNAs and then successfully constructed a functional ceRNA network for TSCC tumorigenesis. Key lncRNAs should be check closely for association with survival and clinical features in TSCC patients, which provide novel lncRNAs as potential prognosis biomarkers and therapeutic targets.

ACKNOWLEDGEMENTS

We thank The Cancer Genome Atlas (TCGA) project and its contributors for this valuable public data set.

REFERENCES

Abou-Elhamd KE, and Habib TN. 2008. The role of chromosomal aberrations in premalignant and malignant lesions in head and neck squamous cell carcinoma. *Eur Arch Otorhinolaryngol* 265:203-207. 10.1007/s00405-007-0420-z

Adeel M, and Suhail A. 2016. Squamous cell carcinoma of oral tongue in young patients - A 10 years tertiary care experience. *J Pak Med Assoc* 66:155-158.

Aken BL, Ayling S, Barrell D, Clarke L, Curwen V, Fairley S, Fernandez Banet J, Billis K, Garcia Giron C, Hourlier T, Howe K, Kahari A, Kokocinski F, Martin FJ, Murphy DN, Nag R, Ruffier M, Schuster M, Tang YA, Vogel JH, White S,
Zadissa A, Flicek P, and Searle SM. 2016. The Ensembl gene annotation system. *Database (Oxford)* 2016. 10.1093/database/baw093

Cai WL, Huang WD, Li B, Chen TR, Li ZX, Zhao CL, Li HY, Wu YM, Yan WJ, and Xiao JR. 2018. microRNA-124 inhibits bone metastasis of breast cancer by repressing Interleukin-11. *Mol Cancer* 17:9. 10.1186/s12943-017-0746-0

Chen L, Yao H, Wang K, and Liu X. 2017. Long Non-Coding RNA MALAT1 Regulates ZEB1 Expression by Sponging miR-143-3p and Promotes Hepatocellular Carcinoma Progression. *J Cell Biochem* 118:4836-4843. 10.1002/jcb.26158

Chen Y, Tian T, Mao MJ, Deng WY, and Li H. 2018. CRBP-1 over-expression is associated with poor prognosis in tongue squamous cell carcinoma. *BMC Cancer* 18:514. 10.1186/s12885-018-4249-1

Cholia RP, Dhiman M, Kumar R, and Mantha AK. 2018. Oxidative stress stimulates invasive potential in rat C6 and human U-87 MG glioblastoma cells via activation and cross-talk between PKM2, ENPP2 and APE1 enzymes. *Metab Brain Dis* 33:1307-1326. 10.1007/s11011-018-0233-3

Dawei H, Honggang D, and Qian W. 2018. AURKA contributes to the progression of oral squamous cell carcinoma (OSCC) through modulating epithelial-to-mesenchymal transition (EMT) and apoptosis via the regulation of ROS. *Biochem Biophys Res Commun* 507:83-90. 10.1016/j.bbrc.2018.10.170

Ding J, Yang C, and Yang S. 2018. LINC00511 interacts with miR-765 and modulates tongue squamous cell carcinoma progression by targeting LAMC2. *J Oral Pathol Med* 47:468-476. 10.1111/jop.12677

Fang Q, Sang L, and Du S. 2018. Long noncoding RNA LINC00261 regulates endometrial carcinoma progression by modulating miRNA/FOXO1 expression. *Cell Biochem Funct*. 10.1002/cbf.3352

Feng L, Houck JR, Lohavanichbutr P, and Chen C. 2017. Transcriptome analysis reveals differentially expressed IncRNAs between oral squamous cell carcinoma and healthy oral mucosa. *OncoTarget* 8:31521-31531. 10.18632/oncotarget.16358

Fromm B, Billipp T, Peck LE, Johansen M, Tarver JE, King BL, Newcomb JM, Sempere LF, Flatmark K, Hovig E, and Peterson KJ. 2015. A Uniform System for the Annotation of Vertebrate microRNA Genes and the Evolution of the Human microRNAome. *Annu Rev Genet* 49:213-242. 10.1146/annurev-genet-120213-092023

Gutschner T, Hammerle M, Eissmann M, Hsu J, Kim Y, Hung G, Revenko A, Arun G, Stentrup M, Gross M, Zornig M, MacLeod AR, Spector DL, and Diederichs S. 2013. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res* 73:1180-1189. 10.1158/0008-5472.can-12-2850

Hauptman N, and Glavac D. 2013. Long non-coding RNA in cancer. *Int J Mol Sci* 14:4655-4669. 10.3390/ijms14034655

Jeggari A, Marks DS, and Larsson E. 2012. miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. *Bioinformatics* 28:2062-2063. 10.1093/bioinformatics/bts344

Jia LF, Wei SB, Gan YH, Guo Y, Gong K, Mitchelson K, Cheng J, and Yu GY. 2014. Expression, regulation and roles of miR-26a and MEG3 in tongue squamous cell carcinoma. *Int J Cancer* 135:2282-2293. 10.1002/ijc.28667

Jiang T, Guo J, Hu Z, Zhao M, Gu Z, and Miao S. 2018. Identification of Potential Prostate Cancer-Related Pseudogenes Based on Competitive Endogenous RNA Network Hypothesis. *Med Sci Monit* 24:4213-4239. 10.12659/msm.910886
Kou N, Liu S, Li X, Li W, Zhong W, Gu L, Chai S, Ren X, Na R, Zeng T, and Liu H. 2018. H19 facilitates tongue squamous cell carcinoma migration and invasion via sponging miR-let-7. *Oncol Res.* 10.3727/096504018x15202945197589

Li L, Yin JY, He FZ, Huang MS, Zhu T, Gao YF, Chen XY, Zhou DB, Chen X, Sun LQ, Zhang W, Zhou HH, and Liu ZQ. 2017a. Long noncoding RNA SFTA1P promoted apoptosis and increased cisplatin chemosensitivity via regulating the hnRNP-U-GADD45A axis in lung squamous cell carcinoma. *Oncotarget* 8:97476-97489. 10.18632/oncotarget.22138

Li M, Zhang W, Zhang S, Wang C, and Lin Y. 2017b. PART1 expression is associated with poor prognosis and tumor recurrence in stage I-III non-small cell lung cancer. *J Cancer* 8:1795-1800. 10.7150/jca.18848

Li S, Chen X, Liu X, Yu Y, Pan H, Haak R, Schmidt J, Ziebolz D, and Schmalz G. 2017c. Complex integrated analysis of lncRNAs-miRNAs-mRNAs in oral squamous cell carcinoma. *Oral Oncol* 73:1-9. 10.1016/j.oraloncology.2017.07.026

Li W, Lu Y, Han R, Yue Q, Song X, Wang F, Wu R, Hou F, Yang L, Xu L, Zhao R, and Hu J. 2018a. Gremlin2 Regulates the Differentiation and Function of Cardiac Progenitor Cells via the Notch Signaling Pathway. *Cell Physiol Biochem* 47:579-589. 10.1016/j.cell.2009.02.006

Li X, Yuan Y, Ren J, Shi Y, and Tao X. 2018b. Incremental Prognostic Value of Apparent Diffusion Coefficient Histogram Analysis in Head and Neck Squamous Cell Carcinoma. *Acad Radiol* 25:1433-1438. 10.1016/j.acra.2018.02.017

Li Y, Chao F, Huang B, Liu D, Kim J, and Huang S. 2014. HOXC8 promotes breast tumorigenesis by transcriptionally facilitating cadherin-11 expression. *Oncotarget* 5:2596-2607. 10.18632/oncotarget.1841

Liu H, Zhang M, Xu S, Zhang J, Zou J, Yang C, Zhang Y, Gong C, Kai Y, and Li Y. 2018. HOXC8 promotes proliferation and migration through transcriptional up-regulation of TGFbeta1 in non-small cell lung cancer. *Oncogenesis* 7:1. 10.1038/s41389-017-0016-4

Liu Y, Xiao N, and Xu SF. 2017. Decreased expression of long non-coding RNA LINC00261 is a prognostic marker for patients with non-small cell lung cancer: a preliminary study. *Eur Rev Med Pharmacol Sci* 21:5691-5695. 10.26355/eurrev_201712_14014

Ma H, Ma T, Chen M, Zou Z, and Zhang Z. 2018. The pseudogene-derived long non-coding RNA SFTA1P suppresses cell proliferation, migration, and invasion in gastric cancer. *Biosci Rep* 38. 10.1042/bsr20171193

Noutomi Y, Oga A, Uchida K, Okafuji M, Ita M, Kawauchi S, Furuya T, Ueyama Y, and Sasaki K. 2006. Comparative genomic hybridization reveals genetic progression of oral squamous cell carcinoma from dysplasia via two different tumourigenic pathways. *J Pathol* 210:67-74. 10.1002/path.2015

Ponting CP, Oliver PL, and Reik W. 2009. Evolution and functions of long noncoding RNAs. *Cell* 136:629-641. 10.1016/j.cell.2009.02.006

Robinson MD, McCarthy DJ, and Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139-140. 10.1093/bioinformatics/btp616

Salameh A, Fan X, Choi BK, Zhang S, Zhang N, and An Z. 2017. HER3 and LINC00052 interplay promotes tumor growth in breast cancer. *Oncotarget* 8:6526-6539. 10.18632/oncotarget.14313

Salmena L, Poliseno L, Tay Y, Kats L, and Pandolfo PP. 2011. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 146:353-358. 10.1016/j.cell.2011.07.014

Sano D, and Myers JN. 2007. Metastasis of squamous cell carcinoma of the oral tongue. *Cancer Metastasis Rev* 26:645-662. 10.1007/s10555-007-9082-y
Shan Y, Ying R, Jia Z, Kong W, Wu Y, Zheng S, and Jin H. 2017. LINC00052 Promotes Gastric Cancer Cell Proliferation and Metastasis via Activating the Wnt/beta-Catenin Signaling Pathway. *Oncol Res* 25:1589-1599. 10.3727/096504017x14897896412027

Shen Y, Wang Z, Loo LW, Ni Y, Jia W, Fei P, Risch HA, Katsaros D, and Yu H. 2015. LINC00472 expression is regulated by promoter methylation and associated with disease-free survival in patients with grade 2 breast cancer. *Breast Cancer Res Treat* 154:473-482. 10.1007/s10549-015-3632-8

Siegel R, Ma J, Zou Z, and Jemal A. 2014. Cancer statistics, 2014. *CA Cancer J Clin* 64:9-29. 10.3322/caac.21208

Siegel RL, Miller KD, and Jemal A. 2015. Cancer statistics, 2015. *CA Cancer J Clin* 65:5-29. 10.3322/caac.21254

Siegel RL, Miller KD, and Jemal A. 2016. Cancer statistics, 2016. *CA Cancer J Clin* 66:7-30. 10.3322/caac.21332

Siegel RL, Miller KD, and Jemal A. 2017. Cancer statistics, 2017. *CA Cancer J Clin* 67:7-30. 10.3322/caac.21387

Siegel RL, Miller KD, and Jemal A. 2018. Cancer statistics, 2018. *CA Cancer J Clin* 68:7-30. 10.3322/caac.21442

Su C, Shi K, Cheng X, Han Y, Li Y, Yu D, and Liu Z. 2018. Long Noncoding RNA LINC00472 Inhibits Proliferation and Promotes Apoptosis of Lung Adenocarcinoma Cells via Regulating miR-24-3p/DEDD. *Technol Cancer Res Treat* 17:1533033818790490. 10.1177/1533033818790490

Sun M, Geng D, Li S, Chen Z, and Zhao W. 2018. LncRNA PART1 modulates toll-like receptor pathways to influence cell proliferation and apoptosis in prostate cancer cells. *Biol Chem* 399:387-395. 10.1515/hsz-2017-0255

Wang ZY, Hu M, Dai MH, Xiong J, Zhang S, Wu HJ, Zhang SS, and Gong ZJ. 2018. Upregulation of the long non-coding RNA AFAP1-AS1 affects the proliferation, invasion and survival of tongue squamous cell carcinoma via the Wnt/beta-catenin signaling pathway. *Mol Cancer* 17:3. 10.1186/s12943-017-0752-2

Wong N, and Wang X. 2015. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res* 43:D146-152. 10.1093/nar/gku1104

Wu Q, Tang SG, and Yuan ZM. 2015. Gremlin 2 inhibits adipocyte differentiation through activation of Wnt/beta-catenin signaling. *Mol Med Rep* 12:5891-5896. 10.3892/mmr.2015.4117

Xiong D, Sheng Y, Ding S, Chen J, Tan X, Zeng T, Qin D, Zhu L, Huang A, and Tang H. 2016. LINC00052 regulates the expression of NTRK3 by miR-128 and miR-485-3p to strengthen HCC cells invasion and migration. *Oncotarget* 7:47593-47608. 10.18632/oncotarget.10250

Yan S, Shan X, Chen K, Liu Y, Yu G, Chen Q, Zeng T, Zhu L, Dang H, Chen F, Ling J, Huang A, and Tang H. 2018. LINC00052/miR-101-3p axis inhibits cell proliferation and metastasis by targeting SOX9 in hepatocellular carcinoma. *Gene* 679:138-149. 10.1016/j.gene.2018.08.038

Yang S, He P, Wang J, Schetter A, Tang W, Funamizu N, Yanaka K, Uwagawa T, Satoskar AR, Gaedcke J, Bernhardt M, Ghadimi BM, Gaida MM, Bergmann F, Werner J, Ried T, Hanna N, Alexander HR, and Hussain SP. 2016a. A Novel MIF Signaling Pathway Drives the Malignant Character of Pancreatic Cancer by Targeting NR3C2. *Cancer Res* 76:3838-3850. 10.1158/0008-5472.can-15-2841

Yang W, Li X, Qi S, Li X, Zhou K, Qing S, Zhang Y, and Gao MQ. 2017. IncRNA H19 is involved in TGF-beta1-induced epithelial to mesenchymal transition in bovine epithelial cells through PI3K/AKT Signaling Pathway. *PeerJ* 5:e3950. 10.7717/peerj.3950

Yang YT, Wang YF, Lai JY, Shen SY, Wang F, Kong J, Zhang W, and Yang HY. 2016b. Long non-coding RNA UCA1 contributes to the progression of oral squamous cell carcinoma by regulating the WNT/beta-catenin signaling pathway. *Oncotarget* 7:30277-30287. 10.18632/oncotarget.11551
signaling pathway. Cancer Sci 107:1581-1589. 10.1111/cas.13058

Ye Y, Yang S, Han Y, Sun J, Xu L, Wu L, Wang Y, and Ming L. 2018. Linc00472 suppresses proliferation and promotes apoptosis through elevating PDCD4 expression by sponging miR-196a in colorectal cancer. Aging (Albany NY) 10:1523-1533. 10.18632/aging.101488

Yuan JH, Yang F, Wang F, Ma JZ, Guo YJ, Tao QF, Liu F, Pan W, Wang TT, Zhou CC, Wang SB, Wang YZ, Yang Y, Yang N, Zhou WP, Yang GS, and Sun SH. 2014. A long noncoding RNA activated by TGF-beta promotes the invasion-metastasis cascade in hepatocellular carcinoma. Cancer Cell 25:666-681. 10.1016/j.ccr.2014.03.010

Zhang HF, Li W, and Han YD. 2018a. LINC00261 suppresses cell proliferation, invasion and Notch signaling pathway in hepatocellular carcinoma. Cancer Biomark 21:575-582. 10.3233/cbm-170471

Zhang Z, Che X, Yang N, Bai Z, Wu Y, Zhao L, and Pei H. 2017. miR-135b-5p Promotes migration, invasion and EMT of pancreatic cancer cells by targeting NR3C2. Biomed Pharmacother 96:1341-1348. 10.1016/j.biopha.2017.11.074

Zhang Z, Qian W, Wang S, Ji D, Wang Q, Li J, Peng W, Gu J, Hu T, Ji B, Zhang Y, Wang S, and Sun Y. 2018b. Analysis of IncRNA-Associated ceRNA Network Reveals Potential IncRNA Biomarkers in Human Colon Adenocarcinoma. Cell Physiol Biochem 49:1778-1791. 10.1159/000493623

Zhao G, Fu Y, Su Z, and Wu R. 2018a. How Long Non-Coding RNAs and MicroRNAs Mediate the Endogenous RNA Network of Head and Neck Squamous Cell Carcinoma: a Comprehensive Analysis. Cell Physiol Biochem 50:332-341. 10.1159/000494009

Zhao M, Liu Y, Liu R, Qi J, Hou Y, Chang J, and Ren L. 2018b. Upregulation of IL-11, an IL-6 Family Cytokine, Promotes Tumor Progression and Correlates with Poor Prognosis in Non-Small Cell Lung Cancer. Cell Physiol Biochem 45:2213-2224. 10.1159/000488166

Zhao Z, Zhang M, Duan X, Deng T, Qiu H, and Zeng G. 2018c. Low NR3C2 levels correlate with aggressive features and poor prognosis in non-distant metastatic clear-cell renal cell carcinoma. J Cell Physiol 233:6825-6838. 10.1002/jcp.26550

Zhou Y, Zhang L, Pan H, Wang B, Yan F, Fang X, Munnee K, and Tang Z. 2015. Bmi1 essentially mediates podocalyxin-enhanced Cisplatin chemoresistance in oral tongue squamous cell carcinoma. PLoS One 10:e0123208. 10.1371/journal.pone.0123208

Zhu J, Fu H, Wu Y, and Zheng X. 2013. Function of IncRNAs and approaches to IncRNA-protein interactions. Sci China Life Sci 56:876-885. 10.1007/s11427-013-4553-6

Zhu L, Yang N, Chen J, Zeng T, Yan S, Liu Y, Yu G, Chen Q, Du G, Pan W, Li X, Zhou H, Huang A, and Tang H. 2017. LINC00052 upregulates EPB41L3 to inhibit migration and invasion of hepatocellular carcinoma by binding miR-452-5p. Oncotarget 8:63724-63737. 10.18632/oncotarget.18892
**Table 1 (on next page)**

Clinicopathological characteristics of 122 patients with tongue squamous cell carcinoma
Table 1: Clinicopathological characteristics of 122 patients with tongue squamous cell carcinoma

| Characteristic      | Subtype | No. of cases (%) |
|--------------------|---------|------------------|
| Age (years)        | < 60    | 59 (48.4)        |
|                    | ≥ 60    | 63 (51.6)        |
| Gender             | Male    | 85 (69.7%)       |
|                    | Female  | 37 (30.3%)       |
| Pathologic stage   | Stage I | 13 (10.7%)       |
|                    | Stage II| 19 (15.6%)       |
|                    | Stage III | 30 (24.6%)    |
|                    | Stage IV | 60 (49.1%)     |
| Pathologic T       | T1      | 19 (15.6%)       |
|                    | T2      | 42 (34.4%)       |
|                    | T3      | 40 (32.8%)       |
|                    | T4      | 21 (17.2%)       |
| Pathologic N       | N0      | 49 (40.2%)       |
|                    | N1      | 17 (13.9%)       |
|                    | N2      | 51 (41.8%)       |
|                    | N3      | 1 (0.8%)         |
|                    | NX      | 4 (3.3%)         |
| Vital status       | Alive   | 72 (59%)         |
|                    | Dead    | 50 (41%)         |
Table 2 (on next page)

Significantly enriched KEGG pathways regulated by DEmRNAs in tongue squamous cell carcinoma
**Table 2:**

Significantly enriched KEGG pathways regulated by DEmRNAs in tongue squamous cell carcinoma

| ID     | Description                                         | pvalue     | p.adjust   | Count |
|--------|-----------------------------------------------------|------------|------------|-------|
| hsa04970 | Salivary secretion                                 | 2.93E-12   | 8.40E-10   | 31    |
| hsa04974 | Protein digestion and absorption                    | 9.63E-11   | 1.38E-08   | 29    |
| hsa04512 | ECM-receptor interaction                            | 3.42E-08   | 3.27E-06   | 24    |
| hsa04060 | Cytokine-cytokine receptor interaction              | 2.38E-06   | 0.000171   | 49    |
| hsa04020 | Calcium signaling pathway                           | 8.26E-06   | 0.000402   | 34    |
| hsa04510 | Focal adhesion                                      | 8.41E-06   | 0.000402   | 36    |
| hsa00500 | Starch and sucrose metabolism                       | 2.25E-05   | 0.000924   | 12    |
| hsa04261 | Adrenergic signaling in cardiomyocytes              | 5.94E-05   | 0.00213    | 27    |
| hsa05414 | Dilated cardiomyopathy (DCM)                        | 0.000147   | 0.004373   | 19    |
| hsa05410 | Hypertrophic cardiomyopathy (HCM)                   | 0.000152   | 0.004373   | 18    |
| hsa04976 | Bile secretion                                      | 0.00022    | 0.00573    | 16    |
| hsa03320 | PPAR signaling pathway                              | 0.000363   | 0.008688   | 16    |
| hsa04973 | Carbohydrate digestion and absorption               | 0.000833   | 0.018399   | 11    |
| hsa04971 | Gastric acid secretion                              | 0.001295   | 0.026541   | 15    |
| hsa04260 | Cardiac muscle contraction                          | 0.001957   | 0.035163   | 15    |
| hsa00910 | Nitrogen metabolism                                | 0.00196    | 0.035163   | 6     |
| hsa04964 | Proximal tubule bicarbonate reclamation             | 0.002199   | 0.035586   | 7     |
| hsa04610 | Complement and coagulation cascades                 | 0.002232   | 0.035586   | 15    |
| hsa05146 | Amoebiasis                                          | 0.002566   | 0.038758   | 17    |
| hsa00830 | Retinol metabolism                                  | 0.003408   | 0.048901   | 13    |
**Table 3** (on next page)

GO terms enriched by 16 DEmRNAs that were involved in the ceRNA network
Table 3: GO terms enriched by 16 DEmRNAs that were involved in the ceRNA network

| Category       | Term                                               | PValue | Genes                  |
|----------------|----------------------------------------------------|--------|------------------------|
| GOTERM_BP_FAT  | GO:0007167–enzyme linked receptor protein signaling pathway | 0.0161 | CHRDL1, PTPRT, GREM2   |
|                | GO:0030509–BMP signaling pathway                  | 0.0257 | CHRDL1, GREM2          |
Table 4 (on next page)

The correlations between DEIncRNAs in the ceRNA network and clinical characteristics of tongue squamous cell carcinoma
Table 4
The correlations between DElncRNAs in the ceRNA network and clinical characteristics of tongue squamous cell carcinoma

| Comparisons                  | Downregulated          | Upregulated               |
|------------------------------|-------------------------|---------------------------|
| Pathologic Stage             |                         |                           |
| (Stage III + IV vs. stage I + II) | HCG22, LINC00492, AL035696.1 | LINC00355, PSORS1C3 |
| Pathologic_T                 |                         |                           |
| (T3 + T4 vs. T1 + T2)        | ERVH48-1                | LINC00520, PSORS1C3      |
| Pathologic_N                 |                         |                           |
| (N2 + N3 vs. N0+ N1)         | LINC00492, ERVH48-1     | AC112721.1, AL139147.1, SFTA1P |
Figure 1

Flow chart of the ceRNA network construction.
Figure 2

Volcano map of DEMRNAs.

Red spots represent up-regulated genes, and green spots represent down regulated genes.
Figure 3 (on next page)

GO and KEGG pathway analyses

(A) Top 5 cellular component terms of dysregulated genes in the GO analysis. (B) Top 5 biological processes terms of dysregulated genes in the GO analysis. (C) Top 5 molecular function terms of dysregulated genes in the GO analysis. (D) Top 10 pathways of dysregulated genes in the pathway analysis. (E) Cytokine-cytokine receptor interaction map from KEGG analysis.
Figure 4

Volcano map of DEIncRNAs.

Red spots represent up-regulated genes, and green spots represent down regulated genes.
Figure 5

Volcano map of DEmiRNAs. Red spots represent up-regulated genes, and green spots represent downregulated genes.
Figure 6

DElncRNAs mediated ceRNA regulatory network in TSCC. The red nodes indicate expression up-regulation, and blue nodes indicate expression down-regulation. LncRNAs, miRNAs and mRNAs are represented by diamond, rounded rectangle, and ellipse, respectively.
Figure 7

Linear regression of PART1 vs SAMD5 expression level.
Figure 8 (on next page)

Kaplan-Meier curve analysis of DEIncRNAs and overall survival rate in tongue squamous cell carcinoma patients.
