**Pathway analysis of a genome-wide association study on a long non-coding RNA expression profile in oral squamous cell carcinoma**

YONG-LE QIU, YUAN-HANG LIU, JIAN-DONG BAN, WEN-JING WANG, MEI HAN, PENG KONG and BING-HUI LI

1Department of Surgery, Fourth Affiliated Hospital, Hebei Medical University, Shijiazhuang, Hebei 050017; 2Department of Stomatology, Second Hospital of Shijiazhuang, Shijiazhuang, Hebei 050000; 3Department of Stomatology, Hebei Eye Hospital, Xingtai, Hebei 054000; 4Department of Biochemistry and Molecular Biology, Hebei Medical University, Shijiazhuang, Hebei 050000, P.R. China

Received June 8, 2018; Accepted November 1, 2018

DOI: 10.3892/or.2018.6870

Abstract. Long non-coding RNAs (IncRNAs) have been consistently demonstrated to be involved in oral squamous cell carcinoma (OSCC) as either tumor oncogenes or tumor suppressors. However, the underlying mechanisms of OSCC tumorigenesis and development have not yet been fully elucidated. The expression profiles of mRNAs and IncRNAs in OSCC were analyzed by a microarray assay. To verify the results of the microarray, 10 differentially expressed IncRNAs were randomly selected and measured by quantitative RT-PCR (qRT-PCR). Gene Ontology (GO) and metabolic pathway analyses were performed to analyze gene function and identify enriched pathways. Subsequently, two independent algorithms were used to predict the target genes of the IncRNAs. We identified 2,294 IncRNAs and 1,938 mRNAs that were differentially expressed in all three OSCC tissues by a microarray assay. Through the construction of co-expression networks of differentially expressed genes, 4 critical IncRNA nodes were identified as potential key factors in the pathogenesis of OSCC. Expression of the 4 critical IncRNA nodes was not associated with age, sex, smoking or tumor location (P>0.05) but was positively correlated with clinical stage, lymphatic metastasis, distant metastasis and survival status (P<0.05). Kaplan-Meier analysis demonstrated that low expression levels of these 4 critical IncRNA nodes contributed to poor median progression-free survival (PFS) and overall survival (OS) (P<0.05). GO and pathway analyses indicated that the functions and enriched pathways of many dysregulated genes are associated with cancer. Potential target genes of dysregulated IncRNAs were enriched in 43 metabolic pathways, with cancer pathways being the primary enrichment pathways. In summary, we analyzed the profile of IncRNAs in OSCC and identified the functions and enriched metabolic pathways of both dysregulated mRNAs and the target genes of dysregulated IncRNAs, providing new insights into molecular markers and therapeutic targets for OSCC.

Introduction

Oral squamous cell carcinoma (OSCC) is one of the most common malignancies in the head and neck region. OSCC is derived from oral mucosal epithelium and is characterized by its strong local infiltration and cervical lymph node metastasis. Many factors, including smoking, can lead to the tumorigenesis and development of OSCC, which has a poor prognosis and low overall survival rate (1,2). OSCC not only affects the lives of late-stage patients but also affects their ability to chew and swallow and their appearance. Although surgery, radiotherapy and chemotherapy have made considerable progress in terms of treatment, the prognosis of OSCC patients remains poor. Early detection and radical treatment of tumors both present challenges and opportunities for cancer researchers (3). The solution to this issue depends on a comprehensive understanding of the molecular mechanisms of tumorigenesis and development. Many research groups are actively involved in the study of OSCC pathogenesis; however, the underlying mechanisms of OSCC tumorigenesis and development have not been fully elucidated (4-6). Therefore, further studies focusing on the mechanisms of OSCC need to be performed to improve early diagnosis, targeted therapy and prognosis.

Long non-coding RNAs (IncRNAs) are a group of non-protein coding RNAs >200 nt (7). It is estimated that only 2% of the human genome is transcribed into mRNAs, while 70-90% of the genome is transcribed into IncRNAs (7). IncRNAs play important roles in epigenetic modification,
transcription and post-transcriptional regulation, maintenance of normal tissue development and differentiation (3,4). Recently, increasing evidence has indicated that lncRNAs exert vital roles in a number of biological processes, including cell metabolism and immune response, through comprehensive mechanisms (8). LncRNAs that affect tumorigenesis and development are considered novel candidates for targeted tumor therapy (9). According to previous research, lncRNAs are closely correlated with tumorigenesis and development of esophageal, liver, lung and breast cancer (10-13). LncRNAs affect the expression of mRNA by regulating the transcription and stability of their target genes (14). Delineating the lncRNA-mRNA coexpression network is an important method for analyzing the functional and regulatory mechanisms of lncRNAs.

The aim of the present study was to identify dysregulated lncRNAs and mRNAs in OSCC patients. The results of the present study indicated that abnormal expression of lncRNAs may contribute to the tumorigenesis and development of OSCC. Furthermore, the present study provides new insight into the molecular markers and therapeutic targets for OSCC.

Materials and methods

Samples. Seventy-two oral squamous cell carcinoma (OSCC) tissues and paired adjacent normal tissues (excised 2 cm from the tumor-free margin) were obtained from the Fourth Affiliated Hospital of Hebei Medical University between January 2015 and October 2016 and were pathologically confirmed to be OSCC (32 were <50 and 40 were ≥50 years; 35 male and 37 female patients). The present study was approved by the Ethics Committee of the Fourth Affiliated Hospital of Hebei Medical University (Shijiazhuang, China) and written informed consent was obtained from all subjects. None of the patients received radiotherapy, chemotherapy or other cancer treatment before tumor resection. Tumors were histologically graded according to the World Health Organization (WHO) standards. Classification of tumors was performed according to the TNM staging revised by the International Union Against Cancer (UICC) (https://www.uicc.org/).

All 72 patients were followed up by telephone and outpatient methods after discharge with assessments of their general condition and clinical symptoms and imaging examinations. The starting point of follow-up was the date of surgery or pathological biopsy, and the follow-up period ended on April 30, 2018. At the end of the follow-up period, 49 patients were still alive, and 23 had died; and no patient was lost to follow-up. Progression-free survival (PFS) was defined as the time between diagnosis and progression of the disease, and overall survival (OS) was defined as the time from diagnosis to death or last follow-up.

Microarray assay. Of these samples, three tissue pairs including tongue cancer (T1N2M0), gingival carcinoma (T2N0M0) and carcinoma of the buccal mucosa (T3N1M0) were used for microarray analysis. The microarray (SBC human 4*180K lncRNA array; Shanghai Biotechnology Corp., Shanghai, China) used in the present study was capable of detecting 77,103 lncRNAs and 18,853 mRNAs and covered core databases, such as GENCODE v21 (https://www.gencodegenes.org), Lncipedia v3.1 (https://lncipedia.org), Ensembl (http://asia.ensembl.org) and Agilent_ncRNA (https://earray.chem.agilent.com).

Total RNA extraction. Total RNA was extracted from frozen samples by TRIzol (Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA). Quantitation was performed using a NanoDrop 2000 spectrophotometer (NanoDrop Technologies; Thermo Fisher Scientific, Inc.). Denaturing agarose gel electrophoresis was used to assess the integrity of total RNA extracted from tissues.

cDNA synthesis, labeling and hybridization. Qualified total RNA was used for the synthesis of cDNA followed by fluorescent labeling according to the manufacturer’s instructions with the Agilent's Low Input Quick Amp WT Labeling kit (Agilent Technologies, Santa Clara, CA, USA). The labeled cRNA was purified with RNeasy Mini kit (Qiagen GmbH, Hilden, Germany) and hybridization was performed at 65°C for 17 h.

Microarray data analysis. An Agilent Microarray Scanner (Agilent Technologies) was used in the present study. Data were obtained using Feature Extraction software 10.7 (Agilent Technologies). Raw data were normalized by Quantite algorithm, Gene Spring Software 11.0 (Agilent Technologies). The lncRNAs and mRNAs were considered to be differentially expressed when the fold change (FC) was >2 (P<0.05). A volcano plot was used to visualize differentially expressed genes and was subsequently processed for hierarchical clustering analysis using Gene Spring Software 11.0 (Agilent Technologies). Finally, Pearson correlation coefficients between differentially expressed lncRNAs and mRNAs were calculated, and co-expression networks of lncRNAs and mRNAs were constructed.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. GO terms were used to annotate and classify gene function. The differentially expressed genes were put into the Database for Annotation, Visualization and Integrated Discovery (DAVID; http://david.abcc.ncifcrf.gov/) v6.8, which utilizes GO to identify the molecular function represented in the gene profile. Furthermore, we used KEGG to analyze the potential functions of these genes in metabolic pathways. P<0.05 was recommended as a cut-off value.

Prediction of lncRNA target genes. Two independent algorithms were used to predict the target genes of dysregulated lncRNAs. The first algorithm was performed to predict cis-acting target genes using the University of California Santa Cruz (UCSC; Santa Cruz, CA, USA) genome browser (http://genome.ucsc.edu/). Genes transcribed within a 10-kb window upstream or downstream of lncRNAs were considered cis target genes. The second algorithm predicted trans-acting target genes using RNAplex 0.2 software (http://www.bioinf.uni-leipzig.de/Software/RNAplex) based on RNA duplex energy prediction and mRNA sequence complementation according to the previous reference (15).

qRT-PCR. Differentially expressed lncRNAs from microarray data were randomly selected for qRT-PCR. Reverse
transcription of total RNA was performed using PrimeScript RT reagent kit with gDNA Eraser (Takara Biotechnology Co., Ltd., Dalian, China). SYBR Green qPCR Master Mix (Takara Biotechnology Co., Ltd.) was used for qRT-PCR assay according to the manufacturer's protocol. The thermocycling conditions were as follows: A denaturation step 10 min at 95˚C, followed by 40 cycles of 15 sec at 95˚C and 30 sec at 60˚C (adjusted with the Tm of different lncRNAs), 30 sec at 72˚C. The housekeeping gene GAPDH was selected as an internal control. The 2−ΔΔCq method was used to measure relative expression levels, and each sample was analyzed at least in triplicate (16). Specific primers of each gene are listed in Table I.

**Table I. Primers used for qRT-PCR.**

<table>
<thead>
<tr>
<th>lncRNA</th>
<th>Forward primer (5'‑3')</th>
<th>Reverse primer (5'‑3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnc-MANSC4‑8:1</td>
<td>AAGGAAAACAACAGAAGAACAC</td>
<td>GCCAGCTTAAAGAGACAAATA</td>
</tr>
<tr>
<td>CXCR2P1</td>
<td>AGGGGAGTATGGGGAGGTGATG</td>
<td>GGGCCAAGGTTTCTTTTCTTTTA</td>
</tr>
<tr>
<td>NRIR</td>
<td>CCAAGAAAAGAGGGCTTAAATGAA</td>
<td>AAGGAGGTTAGAGGTGTCCTGTGTCG</td>
</tr>
<tr>
<td>lnc-CMPK2‑1:3</td>
<td>TCAATAGAGAGGCAGACATACACA</td>
<td>ACAAGAAACACAGCACTAAACAAAC</td>
</tr>
<tr>
<td>lnc-GLI3‑4:1</td>
<td>GATGTGGTGGTCTCCAGTGTGA</td>
<td>TTTCCATCTTGCTTCTATGGTTTT</td>
</tr>
<tr>
<td>NR_104048</td>
<td>AGTTTCCTTTTATTTCTTTTTTGCA</td>
<td>GATCCGTGGCTACTGCGAGA</td>
</tr>
<tr>
<td>MEG3</td>
<td>CTTCTGCTGGGGGAATGGGG</td>
<td>AGAGGGGTGGGGAGGGACT</td>
</tr>
<tr>
<td>lnc-WRN‑10:1</td>
<td>ACATCAAGCTGTAACCAACCCCAAC</td>
<td>TGCCCTTCATCCACACTACAAA</td>
</tr>
<tr>
<td>ENST00000583044</td>
<td>AAATAACCTATACAATCCAAAG</td>
<td>AGAGGAGAGAGATCGGGGAAACC</td>
</tr>
<tr>
<td>ENST00000527317</td>
<td>ACCAGAATGAGTAAAAGAAGA</td>
<td>TGAGAGGTGTGTGAGAAACAAAG</td>
</tr>
<tr>
<td>GAPDH</td>
<td>ATCTTCCAGGAGGAGATCC</td>
<td>TGAGTCTTCCACGATACCAA</td>
</tr>
</tbody>
</table>

lncRNA, long non-coding RNAs.
Statistical analysis. Statistical analysis was performed using SPSS statistical software package (version 22.0; IBM Corp., Armonk NY, USA). Data are expressed as the mean ± standard deviation (SD). Student’s t-test was used to compare different groups. The correlation between the lncRNA expression levels and the clinicopathological factors was analyzed using the Chi-square tests. Survival plots were generated by Kaplan-Meier analysis, and the log-rank test was used to assess the significance of the differences. P-values <0.05 were considered to indicate a statistically significant result.

Results

Patient information. The clinical characteristics of the 72 patients are summarized in Table II.

Volcano plot and hierarchical clustering analysis. In the volcano plot, after normalization and standardization, different signals were distributed in corresponding regions. Gray signals indicated that the detected genes did not meet the screening criteria. The results indicated that there were many lncRNAs and mRNAs that were...
differentially expressed between OSCC and normal tissues (Fig. 1A and B). Hierarchical clustering analysis determined correlations among samples through grouping at the gene level. In the hierarchical clustering analysis, each column represented one sample, and each row represented one gene. Hierarchical clustering analysis showed that the expression profiles were significantly different between OSCC and paired adjacent normal tissues (Fig. 1C and D).

**Differentially expressed lncRNAs and mRNAs.** To explore the role of lncRNAs in OSCC, we performed a genome-wide analysis of lncRNA and mRNA expression in OSCC and normal tissues. The results of the microarray assays showed that 2,294 differentially expressed lncRNAs (accounting for 2.9% of all detectable lncRNAs) and 1,938 differentially expressed mRNAs (accounting for 10.3% of all detectable mRNAs) were identified. Furthermore, 933 IncRNAs and 891 mRNAs were upregulated and 1,361 IncRNAs and 1,047 mRNAs were downregulated. The most upregulated lncRNA and mRNA were MANSC4-8:1 (FC=201.36) and MMP7 (FC=2167.59), respectively. The most downregulated lncRNA and mRNA were NR_117092 (FC=418.62) and IL36A (FC=257.61), respectively. The top 20 dysregulated lncRNAs and mRNAs are summarized in Tables III and IV, respectively.

**Construction of the lncRNA-mRNA co-expression network.** Through construction of a co-expression network, we identified 306 differentially expressed lncRNAs interacting with other selected mRNAs and IncRNAs. According to the results, ENST00000583044, NR_104048, Inc-WRN-10:1 and ENST00000527317 were the four IncRNAs with the most frequent interactions (Fig. 2A and D). These four IncRNAs were node genes for the entire network, with relationship coefficients of 22, 22, 21 and 20. Other genes that did not directly interact with these four node genes interacted with them via indirect means through other relevant genes.

**GO and KEGG pathway analyses.** The differentially expressed mRNAs were processed for GO annotation. The results showed that the differentially expressed genes were enriched in molecular function (MF), biological process (BP) and cellular component (CC). Furthermore, three functions with the most enriched genes were interleukin-1 binding (GO:0019966; Ontology: molecular function; P=0.0003), response to interferon-α (GO:0035455; Ontology: Biological process; P=1.37E-10) and FHF complex (GO:0070695; Ontology: Cellular component; P=0.0035).

The results of the KEGG analysis demonstrated that differentially expressed mRNAs were mainly enriched in 38 biological pathways, including many cancer-related metabolic pathways, e.g., ‘pathways in cancer’ (enriched with 36 differentially expressed genes), ‘bladder cancer’ (enriched with 8 differentially expressed genes), ‘metabolic pathways’ (enriched with 109 differentially expressed genes), ‘pancreatic cancer’ (enriched with 11 differentially expressed genes) and ‘PPAR signaling pathway’ (enriched with 11 differentially expressed genes). The top 10 enriched GO and KEGG terms are summarized in the Tables V and VI.

**Target gene prediction, GO analysis and KEGG pathway enrichment of lncRNAs.** To explore the role of dysregulated IncRNAs in OSCC-related gene regulation and metabolic pathways, the target genes of IncRNAs were predicted using two independent algorithms. The results revealed that 1,470 dysregulated IncRNAs were identified to have cis or trans target genes, including 1,356 IncRNAs targeting 1,250 cis-genes, 370 IncRNAs targeting 2,454 trans-genes and 256 IncRNAs targeting both cis and...
Table III. Top 20 differentially expressed lncRNAs between OSCC and paired adjacent normal tissues.

<table>
<thead>
<tr>
<th>Upregulated lncRNAs</th>
<th>Downregulated lncRNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>lncRNA Source</td>
<td>Fold change</td>
</tr>
<tr>
<td>lnc-MANSC4-8:1</td>
<td>201.357</td>
</tr>
<tr>
<td>ENST00000520185</td>
<td>86.973</td>
</tr>
<tr>
<td>Inc-CXCR3-5:2</td>
<td>64.771</td>
</tr>
<tr>
<td>NR_002712</td>
<td>61.104</td>
</tr>
<tr>
<td>Inc-IFI44-6:1</td>
<td>57.542</td>
</tr>
<tr>
<td>lnc-MANSC4-8:1</td>
<td>51.022</td>
</tr>
<tr>
<td>ENST00000522970</td>
<td>50.388</td>
</tr>
<tr>
<td>Inc-IFI44-5:1</td>
<td>43.841</td>
</tr>
<tr>
<td>ENST00000562027</td>
<td>42.530</td>
</tr>
<tr>
<td>ENST00000455557</td>
<td>38.310</td>
</tr>
<tr>
<td>Inc-G13-4:1</td>
<td>37.568</td>
</tr>
<tr>
<td>ENST00000597169</td>
<td>36.129</td>
</tr>
<tr>
<td>Inc-MSRB3-5:1</td>
<td>31.778</td>
</tr>
<tr>
<td>ENST00000433410</td>
<td>26.210</td>
</tr>
<tr>
<td>ENST00000422194</td>
<td>24.990</td>
</tr>
<tr>
<td>NR_038369</td>
<td>22.397</td>
</tr>
<tr>
<td>NR_110849</td>
<td>22.305</td>
</tr>
<tr>
<td>ENST00000560994</td>
<td>22.167</td>
</tr>
<tr>
<td>NR_126359</td>
<td>18.934</td>
</tr>
<tr>
<td>ENST00000365096</td>
<td>18.736</td>
</tr>
</tbody>
</table>

lncRNA, long non-coding RNAs; OSCC oral squamous cell carcinoma.
trans target genes. We listed the top 30 terms of GO and KEGG enrichment using cis and trans methods (Figs. 3-6).

In the GO analysis, target genes with MF were mainly enriched in molecular binding, e.g., ‘anion binding’ and ‘ion binding’. Target genes with CC functions were mainly enriched in cell organ and cell membrane, e.g., ‘aggresome’ and ‘organelle membrane’. Target genes with the function of BP were mainly enriched in the synthesis process, e.g., ‘cellular biosynthetic process’ and ‘organic substance biosynthetic processs’ (Figs. 3 and 4).
In the KEGG pathway enrichment, the target genes were enriched in 43 pathways. Among them, the cancer pathways were the main pathways. Target genes were also involved in other pathways, e.g., ‘biosynthesis’, ‘metabolism’ and ‘signal pathway’ (Figs. 5 and 6). These enriched metabolic pathways may be the key pathways involved in the regulation of tumorigenesis and development of OSCC by lncRNAs.

Validation by qRT-PCR. To verify the results of the microarray, 10 differentially expressed lncRNAs were randomly selected and assessed by qRT-PCR in 72 patient tissues. The data indicated that lnc-MANSC4-8:1, CXCR2P1, mRNA, messenger RNA.
NRIR, Inc-CMPK2-1:3 and Inc-GLI3-4:1 were significantly upregulated, while TMPRSS11BNL, MEG3, Inc-WRN-10:1, DANCR and Inc-TPP2-7:2 were significantly downregulated in OSCC (P<0.05, Fig. 7). The trend of dysregulated lncRNAs detected by qRT-PCR was consistent with those of the microarray assay.

**Relationship between the expression of ENST00000583044, NR_104048, Inc-WRN-10:1, ENST00000527317 and the clinicopathological features of OSCC patients.** In 72 cases of OSCC, the expression levels of ENST00000583044, NR_104048, Inc-WRN-10:1 and ENST00000527317 were significantly lower than those in normal tissues (P<0.05) (Fig. 8). According to the average expression levels of ENST00000583044, NR_104048, Inc-WRN-10:1 and ENST00000527317 in OSCC, patients with OSCC were divided into a high ENST00000583044 expression group (≥3.15) (n=30) and a low ENST00000583044 expression group (<3.15) (n=42); a high NR_104048 expression group (≥2.99) (n=28) and a low NR_104048 expression group (<2.99) (n=44); a high Inc-WRN-10:1 expression group (≥3.35) (n=26) and a low Inc-WRN-10:1 expression group (<3.35) (n=46); and a high ENST00000527317 expression group (≥2.91) (n=30) and a low ENST00000527317 expression group (<2.91) (n=42).

We then analyzed the relationships among the expression of ENST00000583044, NR_104048, Inc-WRN-10:1, ENST00000527317 and age, sex, smoking status, tumor location, clinical stage, lymphatic metastasis, distant metastasis and survival status (P<0.05). However, no significant associations were detected among the expression of ENST00000583044, NR_104048, Inc-WRN-10:1 and ENST00000527317 and age, sex, smoking and tumor location (P>0.05, Table II).

Kaplan-Meier analysis demonstrated that the median OS for patients with low expression of ENST00000583044, NR_104048, Inc-WRN-10:1 and ENST00000527317 was
significantly lower than that in patients with high expression of these factors (P<0.05) (Fig. 9). Furthermore, the median PFS for patients with high expression of ENST00000583044, NR_104048 and Inc-WRN-10:1 was significantly higher than that in patients with low expression (P<0.05) (Fig. 10).
Discussion

Tumorigenesis and development of oral squamous cell carcinoma (OSCC) consists of a complex process, and the underlying mechanisms remain poorly understood. The aim of the present study was to explore the relationship between IncRNAs and OSCC. IncRNAs were initially considered 'noise' without any biological functions in the human genome. They are now known to play important roles in gene expression, and their differential expression may affect corresponding functional performance (4). Subsequently, many studies have demonstrated that IncRNAs are involved in many important regulatory processes, including X-chromosome silencing, genomic imprinting, chromatin modification, transcriptional activation, transcriptional interference and intranuclear transport (17,18). In the present study, we used an expression profile microarray to identify differentially expressed genes in OSCC at the whole genome level. The results revealed a large number of differently expressed IncRNAs and mRNAs, some of which may be important genes involved in tumorigenesis and the development of OSCC. For example, MALAT1 promotes the invasion and metastasis of lung cancer (19). MEG3 expression levels were highly correlated with invasion and metastasis of gastric cancer (20,21). GAS5 indicates a poor prognosis in ovarian cancer (22). Feng et al showed that IncRNAs were abnormally expressed in OSCC and metastatic tissue samples (5). Recent studies have shown that HOTAIR is highly expressed in OSCC and is associated with the biological behavior of tumor invasion and metastasis (6,23).

In view of the complex transcriptional regulatory mechanisms of IncRNAs and their ability to form a variety of secondary functional structures, their biological functions cannot be predicted based solely on nucleic acid sequence (24,25). IncRNA loci are often located in intronic regions of the coding gene and thus may affect expression of its adjacent genes (26-28). To investigate the relationship between IncRNAs and mRNAs, coexpression networks of dysregulated IncRNAs and mRNAs were constructed based on differential expression. Differentially expressed genes were assigned into subnetworks associated with phenotypic functions to roughly deduce the function of IncRNAs in this subnetwork and predict possible regulatory mechanisms. We identified 4 IncRNAs with high correlations with other genes, ENST00000583044, NR_104048, Inc-WRN-10:1 and ENST00000527317, which were found in the present study, belonged to downregulated IncRNAs and have never previously been reported in OSCC or any other solid tumors. Our data are the first to reveal these four critical node genes. We speculate that these results may depend on tumor heterogeneity. As with any cancer, OSCC is also highly heterogeneous and is characterized by different genetic backgrounds, different pathological types, different differentiation states, different gene mutation and transcriptional patterns and proteome expression profiles (29,30). We further examined the expression of these 4 genes in 72 patients with OSCC and normal tissues and analyzed their relationship with the clinico-pathological features and prognosis. We found that these 4 genes were downregulated in OSCC. Moreover, their expression was not correlated with age, sex, smoking, or tumor location but was related to clinical stage, lymphatic metastasis, distant metastasis and survival status. Furthermore, low expression levels of these 4 IncRNAs contributed to poor median PFS and OS.

In addition, by constructing a co-expression network of differentially expressed genes, we found that some IncRNAs, e.g., NR_002812, regulate many genes, including SPI100 and B2M. Previous studies have shown that the SPI100 protein is involved in viral infection, virus-related protein interaction and self-ubiquitination regulation, and plays an important role in interferon and p53 signaling pathways (31,32). p53 protein inhibits the growth and invasion of oral malignancy by regulating the phosphorylation of AKT (25,33,34). Although the specific mechanism and related signaling pathways need to be further studied, they can be used as molecular markers for early diagnosis, treatment and prognostic evaluation of OSCC.

Since IncRNAs do not encode protein, we studied the mechanism of pathogenesis from another angle by enriching the biological function of differentially expressed mRNAs. In the present study, GO and KEGG pathway analyses were performed to examine the biological function of dysregulated genes (35,36). GO functional annotations indicated that the differentially expressed genes were enriched in different BP, CC and MF categories. KEGG enrichment identified many metabolic pathways associated with cancer, e.g., ‘metabolic pathways’ and ‘pathways in cancer’, demonstrating that OSCC is associated with cell structure changes, metabolic process disorders, tumor suppressor genes and oncogene signaling pathway abnormalities and was a consequence of multiple intracellular and external factors. These results further verified the results of the microarray assay. IncRNAs can guide gene expression in either a cis or trans manner. In the present study, two independent algorithms were used to predict cis and trans target genes of differentially expressed IncRNAs in OSCC. Through GO functional annotation and KEGG pathway analysis, we found that these target genes regulate relevant OSCC proteins and affect OSCC tumorigenesis and development via their functions in organ, molecular binding, metabolism and cancer pathways. Target prediction of IncRNAs provides important information for further study of potential functional IncRNAs and target genes in OSCC.

The present study also has several limitations including its small sample size. In the future, our results need to be validated in large-scale samples.

In conclusion, in the present study, 2,294 dysregulated IncRNAs and 1,938 dysregulated mRNAs were identified by a microarray assay. We explored 4 critical IncRNAs nodes, which may play an important role in the pathogenesis of OSCC. GO and pathway analyses indicated that the functions and enriched pathways of many dysregulated genes were associated with cancer. The potential target genes of dysregulated IncRNAs were enriched in 43 KEGG pathways, and cancer pathways were the primary enrichment pathways. These results provide new insight into the molecular markers and therapeutic targets for OSCC.

Acknowledgements

Not applicable.

Funding

The present study was supported by the National Natural Science Foundation of China (no. 81372150).
Availability of data and materials
The datasets used during the present study are available from the corresponding author upon reasonable request.

Authors’ contributions
YLQ, YHL and BHL acquired the data and created a draft of the manuscript; JDB, WJW, MH and PK conducted the experiments and collected the data; YLQ and YHL interpreted the data, performed the statistical analysis and analyzed the results; BHL revised and approved the final version of the manuscript. All authors read and approved the manuscript and agree to be accountable for all aspects of the research in ensuring that the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Ethics approval and consent to participate
The present study was approved by the Ethics Committee of the Fourth Affiliated Hospital of Hebei Medical University (Shijiazhuang, China) and written informed consent was obtained from all subjects.

Patient consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References


