

Long non-coding RNA expression profiles and related regulatory networks in areca nut chewing-induced tongue squamous cell carcinoma

PANCHUN LI¹, SHANSHAN ZHANG², YONGZHEN MO³, LISHEN ZHANG³,
YUMIN WANG³, FANG XIONG², SHUAI ZHANG⁴, JIANG LIU¹, YUMING XU¹,
ZHAOYANG ZENG³, WEI XIONG³, YONG LI⁵ and ZHAOJIAN GONG¹

¹Department of Oral and Maxillofacial Surgery, The Second Xiangya Hospital, Central South University, Changsha, Hunan 410011; ²Department of Stomatology, The Key Laboratory of Carcinogenesis and Cancer Invasion of The Chinese Ministry of Education, Xiangya Hospital, Central South University, Changsha, Hunan 410008; ³NHC Key Laboratory of Carcinogenesis, Cancer Research Institute and School of Basic Medical Sciences, Central South University, Changsha, Hunan 410078; ⁴Department of Otolaryngology, Xiangya Hospital, Central South University, Changsha, Hunan 410008, P.R. China; ⁵Department of Medicine, Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, Texas 77030, USA

Received March 1, 2020; Accepted August 24, 2020

DOI: 10.3892/ol.2020.12165

Abstract. Areca nut chewing is an important risk factor for developing tongue squamous cell carcinoma (TSCC), although the underlying molecular mechanism is unknown. To determine the potential molecular mechanisms of areca nut chewing-induced TSCC, the present study performed whole-genome detection with five pairs of TSCC and adjacent normal tissues, via mRNA- and long non-coding (lnc)RNA-gene chip analysis. A total of 3,860 differentially expressed genes were identified, including 2,193 lncRNAs and 1,667 mRNAs. Gene set-enrichment analysis revealed that the differentially expressed mRNAs were enriched in chromosome 22q13, 8p21 and 3p21 regions, and were regulated by nuclear factor kappa B (NF- κ B) and interferon regulatory factors (IRFs). The results

of ingenuity pathway analysis revealed that these mRNAs were significantly enriched for inflammatory immune-related signaling pathways. A co-expression network of mRNAs and lncRNAs was constructed by performing weighted gene co-expression network analysis. The present study focused on NF- κ B-, IRF- and Th cell-signaling pathway-related lncRNAs and the corresponding mRNA-lncRNA regulatory networks. To the best of our knowledge, the present study was the first to investigate differential mRNA- and lncRNA-expression profiles in TSCCs induced by areca nut chewing. Inflammation-related mRNA-lncRNA regulatory networks driven by IRFs and NF- κ B were identified, as well as the Th cell-related signaling pathways that play important carcinogenic roles in areca nut chewing-induced TSCC. These differentially expressed mRNAs and lncRNAs, and their regulatory networks provide insight for further analysis on the molecular mechanism of areca nut chewing-induced TSCC, candidate molecular markers and targets for further clinical intervention.

Correspondence to: Dr Zhaojian Gong, Department of Oral and Maxillofacial Surgery, The Second Xiangya Hospital, Central South University, 139 Renmin Road, Changsha, Hunan 410011, P.R. China
E-mail: gongzhaojian4458@csu.edu.cn

Abbreviations: cDNA, complementary DNA; cRNA, complementary RNA; GSEA, gene set-enrichment analysis; IPA, Ingenuity Pathway Analysis; IRF, interferon regulatory factor; lncRNA, long non-coding RNA; OSCC, oral squamous cell carcinoma; NF- κ B, nuclear factor-kappa B; OSF, oral submucous fibrosis; PCA, principal component analysis; SAM, significance analysis of microarray; TSCC, tongue squamous cell carcinoma; TOM, topological overlap matrix; WGCNA, weighted gene co-expression network analysis

Key words: areca nut, long non-coding RNA, network, expression profile, tongue squamous cell carcinoma

Introduction

Oral squamous cell carcinoma (OSCC) is one of the most common cancers worldwide. Annually, over 350,000 new patients are diagnosed with oral cancer worldwide, and over 170,000 patients die from this disease (1). Tongue squamous cell carcinoma (TSCC) is the most common type of OSCC (2,3). For example, in Japan, TSCC accounted for 30.4-43.8% of OSCC between 2000 and 2014 (4). TSCC presents with early metastasis and poor prognosis due to the abundance of blood and lymphatic vessels in the tongue, as well as the continuous movement of the tongue (5). Although surgery, radiotherapy and chemotherapy have demonstrated great progress in cancer treatment in recent decades, the 5-year survival rate (40-50%)

remains low, and the prognosis of most patients is poor (6-10). Among the pathogenic factors of several TSCCs, reports have demonstrated an association between areca nut chewing and TSCC (11,12).

The areca nut is classified as a group 1 carcinogen, which can induce oral precancerous lesions and oral cancer (13). Oral submucous fibrosis (OSF) is a precancerous disorder caused by areca nut chewing (14). Areca nut extract can stimulate oral keratinocyte to secrete fibrosis-related cytokines, promote fibroblast proliferation and collagen synthesis, cause oral mucosal fibrosis, and ultimately induce OSF (15). Previous studies have demonstrated that OSF transforms into a malignant tumor in 1.5-15% of all cases (14,16) and increases the risk of oral cancer by 19.1-fold (17,18). In India, Pakistan and other countries or regions, oral cancer is at the forefront of malignant tumors due to the high prevalence of areca nut chewing, and consequently, the high morbidity of OSF (19,20). For example, 36.5% of men and 8.4% of women in India have areca nut chewing habits between the ages of 15 and 49 years in 2006, which led to a prevalence of OSF increased from 0.03% in 1968 to 6.42% in 2006, and >100,000 cases of OSCC are registered each year in India (20-24). Approximately 13.3% of residents in the Hunan province of China chew areca nuts, which has significantly increased the local oral cancer mortality (age-standardized mortality rates in 2013: 5.50 per 100,000) (25-27). However, the molecular mechanisms whereby areca nut chewing induces OSF and TSCC remain unclear.

A large number of genes are abnormally expressed during carcinogenesis, which promotes tumor initiation and progression, including p53 and MYC (28-30). Increasing evidence indicates that long non-coding (lnc)RNAs play important roles in carcinogenesis (31-40). However, the differential lncRNA expression profiles and related gene-regulatory networks associated with carcinogenesis of areca nut chewing-induced TSCC have not yet been reported. Thus, the present study collected five pairs of TSCC and adjacent normal tissues and monitored the resultant lncRNA- and mRNA-expression profiles using a lncRNA microarray. In addition, lncRNAs associated with areca nut chewing-induced TSCC and their potential gene expression-regulatory networks were screened, and the functions of these lncRNAs in TSCC were investigated.

Materials and methods

Sample collection. A total of five pairs of TSCC and adjacent normal tissues were randomly collected for a chip study at The Second Xiangya Hospital of Central South University (Changsha, China) between August 2015 and October 2015. All patients were male, with a mean age of 44.6 years (age range, 35-57 years). The inclusion criteria were as follows: Patients who reported a history of areca nut chewing, newly diagnosed with TSCC, who had not undergone preoperative radiotherapy or chemotherapy, and who did not have any other systemic disease. The exclusion criteria were as follows: Patients without a history of areca nut chewing and patients who had undergone preoperative radiotherapy or chemotherapy, or who had any other systemic diseases. In addition, the functions of the vital organs of all patients were in the normal range. Based on the International Union Against Cancer Eighth Edition TNM classification system for malignant tumors (41), 1 case was

identified as T1N1M0, 1 case as T1N2M0, 2 cases as T2N0M0 and 1 case as T2N1M0 (Table SI). All tissue specimens were immediately frozen and stored in liquid nitrogen until subsequent experiments. The present study was approved by the Ethics Committee of Cancer Research Institute of Central South University (approval no. 2015-068; Changsha, China) and written informed consent was provided by all patients.

Gene chip. The present study used a 4x180K lncRNA array chip (Agilent Technologies, Inc.), which contains 78,243 lncRNA microarray probes and 32,776 mRNA microarray probes, and every microarray probe consists of 60 oligonucleotides. Data were uploaded to the Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>; accession no. GSE139869).

Obtaining lncRNA expression profile data. Tissue RNA was extracted using TRIzol reagent (Invitrogen; Thermo Fisher Scientific, Inc.; cat. no. 15596026), and all materials used were processed using RNAase-free diethyl pyrocarbonat water (Sigma-Aldrich; Merck KGaA; cat. no. D5758). Briefly, 50-100 mg of human tissue samples were ground into tissue homogenate in liquid nitrogen. After volatilizing the liquid nitrogen, 1 ml TRIzol reagent was immediately added. RNA was subsequently extracted according to the manufacturer's instructions (Invitrogen; Thermo Fisher Scientific, Inc.).

Sample labeling, microarray hybridization, and washing were performed according to the manufacturer's protocols (42). Briefly, total RNA was reverse transcribed into double-strand cDNA by using Low Input Quick Amp Labeling Kit (Agilent Technologies, Inc.; cat. no. 5190-2305). The reverse transcription procedure was as follows: 65°C for 10 min, 40°C for 2 h, 70°C for 15 min, at 4°C. Next complementary RNA (cRNA) was synthesized and labeled with cyanine-3-CTP. The labeled cRNAs were hybridized onto the microarray. After washing by using Gene Expression Wash Buffer Kit (Agilent Technologies, Inc.; cat. no. 5188-5327) twice, the arrays were scanned using a G2505C Agilent Scanner (Agilent Technologies, Inc.) and raw data was obtained using Feature Extraction software (version 10.7.1.1; Agilent Technologies, Inc.).

Microarray data were filtered to remove 'background noise' for non-expressed or very poorly expressed targets. After filtering the data, 64,556 probes remained, of which 24,149 corresponded to mRNAs and 40,407 corresponded to lncRNAs.

Principal component analysis (PCA) and differential expression analysis. To understand whether there are different characteristics between the areca nut chewing and the non-areca nut chewing TSCC expression profiles, PCA was performed on the two groups of data. PCA is commonly used for data dimensionality reduction, which can easily reveal the characteristics of different types of samples (43,44). The non-areca nut chewing TSCC lncRNA expression profile GSE34105 dataset was downloaded from the GEO database (45). These data are from Swedish samples, where there is no habit of chewing areca nut. PCA was performed using the programming language R (version 3.6.0, <https://www.r-project.org>).

Differentially expressed mRNAs and lncRNAs were screened using the significance analysis of microarrays (SAM)

technique. The screening criteria were fold-change of ≥ 2.0 and $P < 0.05$. Heat maps were constructed to identify differentially expressed mRNAs and lncRNAs, using Genesis software.

Chromosomal localization of differentially expressed mRNAs and lncRNAs by gene set-enrichment analysis (GSEA). GSEA is a method for analyzing gene expression information, which helps determine the effects of gene expression changes on several biological functions and pathways, including T cell receptor signaling pathway (46,47). The present study used TSCC and adjacent normal tissues as group variables and used GSEA to perform enrichment analysis of positional gene sets (c1) with differentially expressed genes between TSCC and adjacent normal tissues, to predict the regional chromosomal changes (chromosomal deletions or amplifications). 'c1: Positional gene sets (c1.all.v6.2. symbols)' were used as the given sets of genes from the Molecular Signatures Database (MSigDB; <https://www.gsea-msigdb.org/gsea/msigdb>), and the number of permutations was set to repeat 1,000 times.

Establishment of mRNA and lncRNA co-expression networks using weighted gene co-expression network analysis (WGCNA). In addition to chromosomal locations of mRNAs and lncRNAs, the associations between mRNAs and lncRNAs can be analyzed via their expression trends (48,49). WGCNA is a commonly used algorithm for constructing gene co-expression networks by gene expression trends (50). First, soft threshold was set to 24 to make network construction obeyed a scale-free network distribution, and a heat map of the topological overlap matrix (TOM) between genes was constructed. The thresholds were set and highly correlated co-expression networks were selected and visualized using Cytoscape software (version 3.7.0; <http://www.cytoscape.org>).

Predicting upstream regulatory factors and selecting related lncRNAs that may be regulated by these transcription factors. GSEA was performed to determine whether some of the differentially expressed lncRNAs and mRNAs are regulated by the same transcription factors and to predict the upstream regulatory factors of these differentially expressed genes. GSEA was also performed to determine whether the related lncRNAs are regulated by these transcription factors. 'c3: Motif gene sets (c3. tft. v6.2. symbols. gmt)' was used as the reference gene set (from the GSEA MSigDB database; <https://www.gsea-msigdb.org/gsea/msigdb>) for enrichment analysis to predict the upstream regulatory factors of differentially expressed genes, as previously described (46). The number of permutations was set to repeat 1,000 times. Genes regulated by these transcription factors were identified in the GSEA reference gene set and selected lncRNAs that are highly related to these genes (the topological-overlap values between these lncRNAs and more than half of mRNAs of the GSEA gene set exceeded the threshold of 0.4) in the co-expression network constructed by WGCNA. The related regulatory networks were constructed by Cytoscape software (version 3.7.0, <http://www.cytoscape.org>).

Analyzing canonical pathways of differentially expressed genes via ingenuity pathway analysis (IPA) and selecting

pathway-related lncRNAs in the WGCNA constructed co-expression network. To identify functional lncRNAs in complex co-expression networks, the IPA tool was used to perform canonical pathway analysis on differentially expressed genes, and lncRNAs associated with significantly enriched pathways were screened in WGCNA-constructed co-expression networks. Based on the Ingenuity Knowledge Base (51), IPA can be used to identify biological networks and signaling pathways for specific datasets (51,52). By importing gene-expression profiles into the IPA program, signaling pathways that were significantly enriched for differentially expressed genes, as well as the genes involved in these pathways were identified. lncRNAs that were highly correlated with these genes (topological-overlap values between these lncRNAs and more than half of mRNAs of the IPA gene sets exceeded the threshold of 0.4) were selected in the WGCNA-constructed gene co-expression network.

Results

Differential expression of lncRNAs and mRNAs. Gene expression profiles were obtained from five pairs of TSCC and adjacent normal tissues, via mRNA- and lncRNA-gene chip analysis, and we eventually detected 78,243 lncRNAs and 32,276 mRNAs. Following preliminary screening to remove data for lowly expressed lncRNAs and mRNAs, 40,407 lncRNAs and 24,149 mRNAs remained. PCA analysis was subsequently performed and the results demonstrated that the areca nut chewing expression and non-areca nut chewing expression profiles obtained from the database (GSE34105) had different characteristics (Fig. S1). In addition, differentially expressed mRNAs and lncRNAs were screened via SAM, with the screening criteria set as follows: Fold-change ≥ 2.0 and $P < 0.05$. A total of 3,860 differentially expressed RNAs were screened (Fig. 1A), including 1,667 mRNAs (866 upregulated and 801 downregulated; Fig. 1B) and 2,193 lncRNAs (1,057 upregulated and 1,136 downregulated; Fig. 1C).

Chromosomal localization of differentially expressed mRNAs and lncRNAs via GSEA. GSEA positional gene sets can be used to identify regional chromosomal changes (chromosomal deletions or amplifications) that affect gene expression (53). The results of the present study demonstrated significant changes in three chromosomal segments via GSEA, including 22q13, 8p21 and 3p21. The upregulated genes were significantly enriched in the 22q13 segment (Fig. 2A), whereas the downregulated genes were significantly enriched in the 8p21 and 3p21 segments (Figs. S2 and S3). Taken together, these results suggest that chromosome amplification may occur in the 22q13 region, while chromosome deletion may occur in the 8p21 and 3p21 regions. In addition, according to the annotation information of the gene chip, 21 differentially expressed lncRNAs were in segment 22q13 (Fig. 2B). It is speculated that the mRNAs and lncRNAs expressed from this region were upregulated due to chromosomal amplification of the segment. Similarly, 13 downregulated lncRNAs from the 8p21 region and 51 downregulated lncRNAs from the 3p21 region were identified (Figs. S2 and S3), further indicating that chromosomal deletion occurs within these regions.

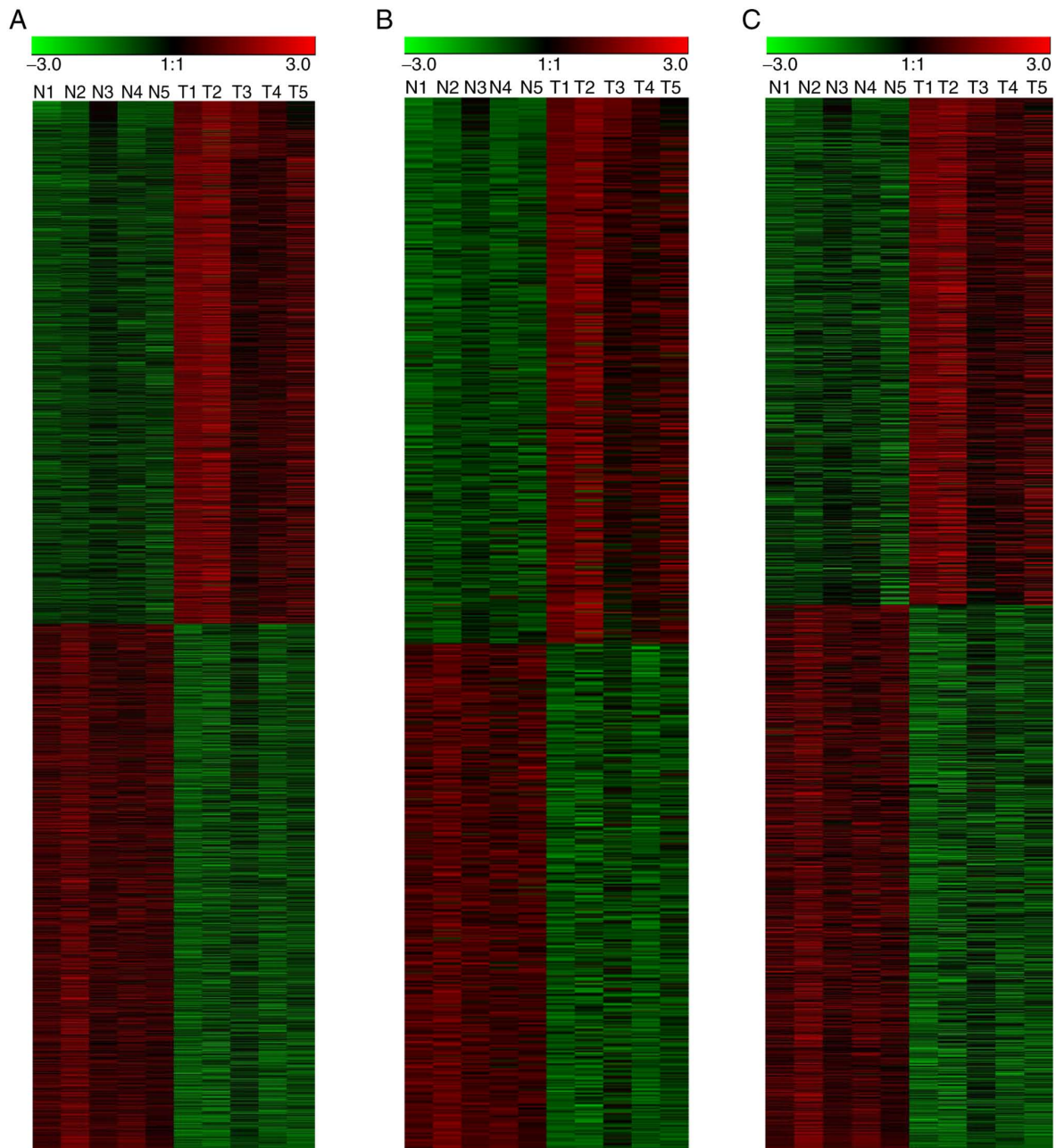


Figure 1. Differentially expressed molecular heat maps of TSCC and adjacent normal tissues. (A) Heat map of differentially expressed mRNAs and lncRNAs. (B) Heat map of differentially expressed mRNAs. (C) Heat map of differentially expressed lncRNAs. Fold-change ≥ 2 and $P < 0.05$. T represents TSCC tissue and N represents adjacent normal tissue. Red indicates upregulated genes and green indicates downregulated genes. TSCC, tongue squamous cell carcinoma; lncRNA, long non-coding RNA.

Construction of a co-expression network of differentially expressed lncRNAs and mRNAs. In addition to chromosomal co-localization of mRNAs and lncRNAs, a co-expression network based on similar expression trends between differentially expressed mRNAs and lncRNAs was constructed using WGCNA. First, the soft threshold for network construction was determined (Fig. S4), which made the constructed network conform to a scale-free distribution. Networks constructed in such a manner are similar to biological networks found in nature (50). Subsequently, the topological-overlap values among the molecules were calculated, and a hierarchical-clustering tree was constructed based on these values (Fig. 3A). Based on the topological-overlap values, genes with similar

expression patterns were grouped into the same module. A TOM of the differentially expressed lncRNAs and mRNAs was constructed using the associated topological-overlap values, and the TOM was visualized in the form of a heat map (Fig. 3A). The expression network was divided into seven color modules in the present study. By setting the threshold, the highly correlated co-expressed molecules were screened (topological-overlap values > 0.40) and a visual co-expression network was constructed using Cytoscape software (Fig. 3B). Due to the high threshold, four modules with fewer molecules and lower topological-overlap values between the molecules were almost filtered out. Thus, only three modules remained in the network, which consisted of 913 nodes, including

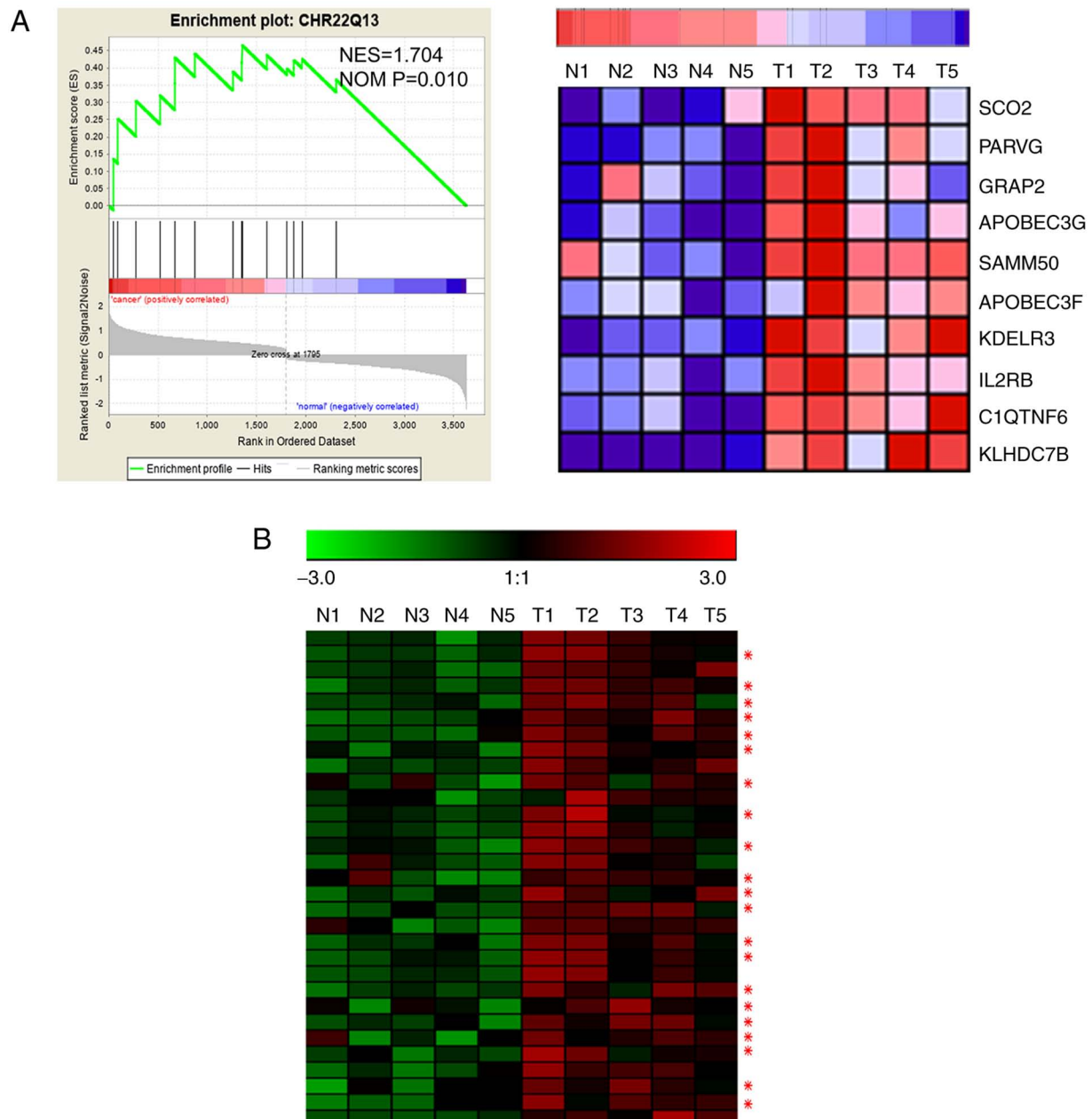


Figure 2. GSEA revealed that differentially expressed genes in TSCC and adjacent normal tissues are enriched in chromosome 22q13. (A) GSEA gene sets demonstrated that several genes located in the 22q13 segment were upregulated in TSCC. (B) Heat map of differentially expressed long non-coding RNAs (asterisks) and mRNAs from the 22q13 segment. GSEA, gene set-enrichment analysis; TSCC, tongue squamous cell carcinoma.

538 lncRNAs and 375 mRNAs, and 60,223 linkages (mRNA-lncRNA associations).

Prediction of upstream regulatory factors and related lncRNA-mRNA network that these factors may drive. To determine which transcription factors regulate the differentially expressed lncRNAs and mRNAs in TSCC, the c3: Motif gene sets (c3.tft.v6.2.symbols.gmt; GSEA) was used as a reference gene set for enrichment analysis to predict the upstream regulatory factors (Fig. 4A). The most significantly enriched molecules were interferon regulatory factors (IRFs; $P=0.003$) and nuclear factor-kappa B (NF- κ B; $P=0.006$; Fig. 4A). The downstream mRNA-lncRNA networks regulated by IRFs and NF- κ B were constructed (Fig. 4B). The results demonstrated that the transcription of

four differentially expressed genes (CXCL10, MSC, UBD and HOXB3) was simultaneously regulated by IRFs and NF- κ B (Fig. 4B).

In addition to these downstream mRNAs, the present study investigated whether IRFs and NF- κ B drive the expression of some lncRNAs in TSCC. The correlation between mRNAs and lncRNAs can be analyzed via their expression. Highly related molecules may participate in the same regulatory network and be regulated by the same regulatory factors. Thus, the present study screened for lncRNAs (Tables SII and SIII) whose expression trends were highly consistent with those of IRF-regulated or NF- κ B-regulated mRNAs in the WGCNA-constructed co-expression network. As presented in Fig. 4C, the expression trends of the lncRNAs (Tables SII) and IRF-regulated mRNAs were similar, suggesting that the lncRNAs may also

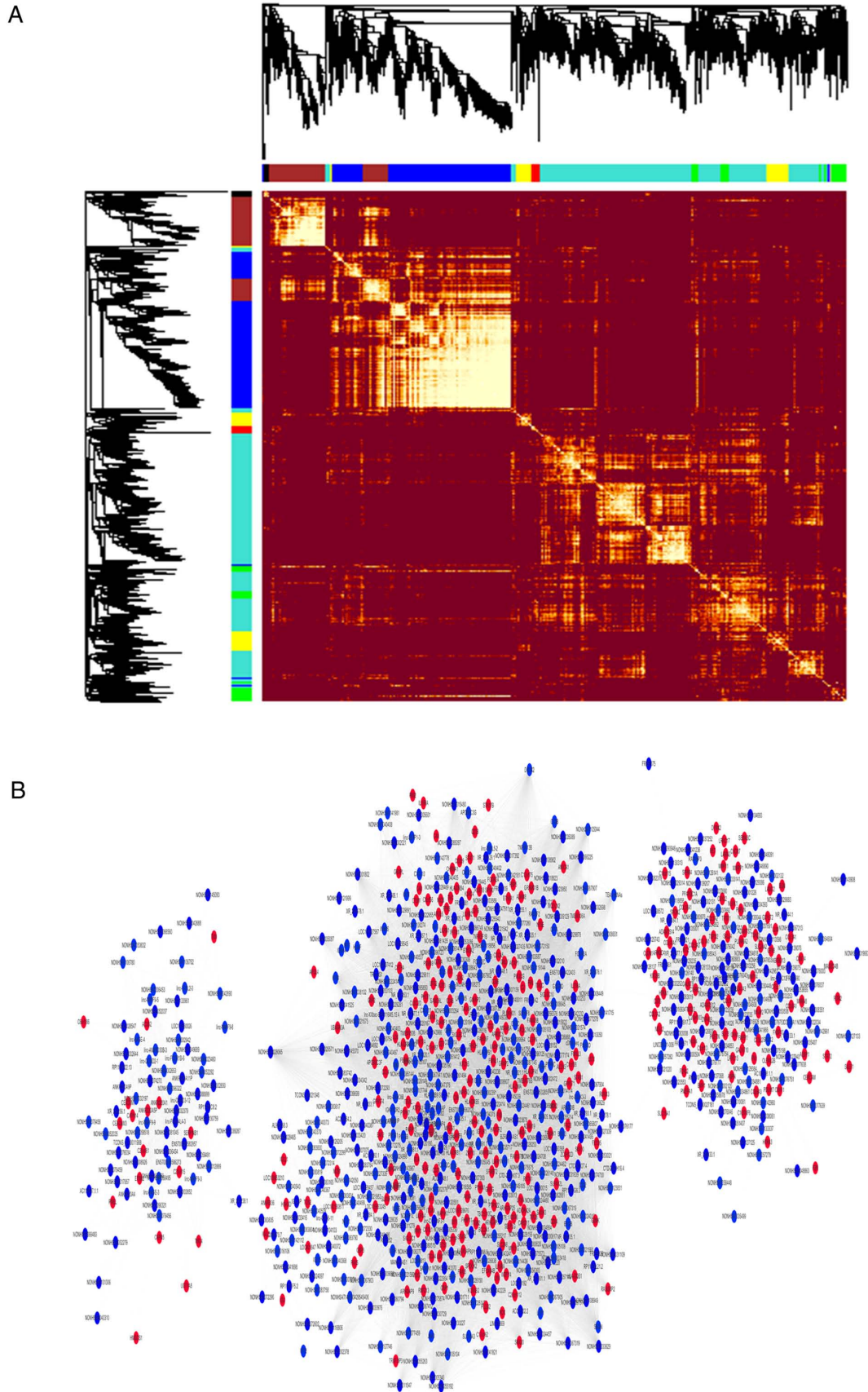


Figure 3. WGCNA-constructed mRNA-lncRNA co-expression network in TSCC. (A) Topological overlap matrix heat map of differentially expressed lncRNAs and mRNAs, which demonstrated that the network exhibits a scale-free topology. Red represents a lower overlap, while yellow represents a higher overlap. The top and left sides of the heat map are hierarchical cluster trees, with different branches of the cluster tree representing different gene modules, while the corresponding colors represent different modules. (B) Co-expression network of differentially expressed mRNAs and lncRNAs in TSCC and adjacent normal tissues, obtained via WGCNA and visualized using Cytoscape software. The network consisted of 913 nodes, including 538 lncRNAs and 375 mRNAs, and 60,223 linkages (mRNA-lncRNA associations). The topological-overlap value was higher than the threshold of 0.40. WGCNA, weighted gene co-expression network analysis; lncRNA, long non-coding RNA; TSCC, tongue squamous cell carcinoma.

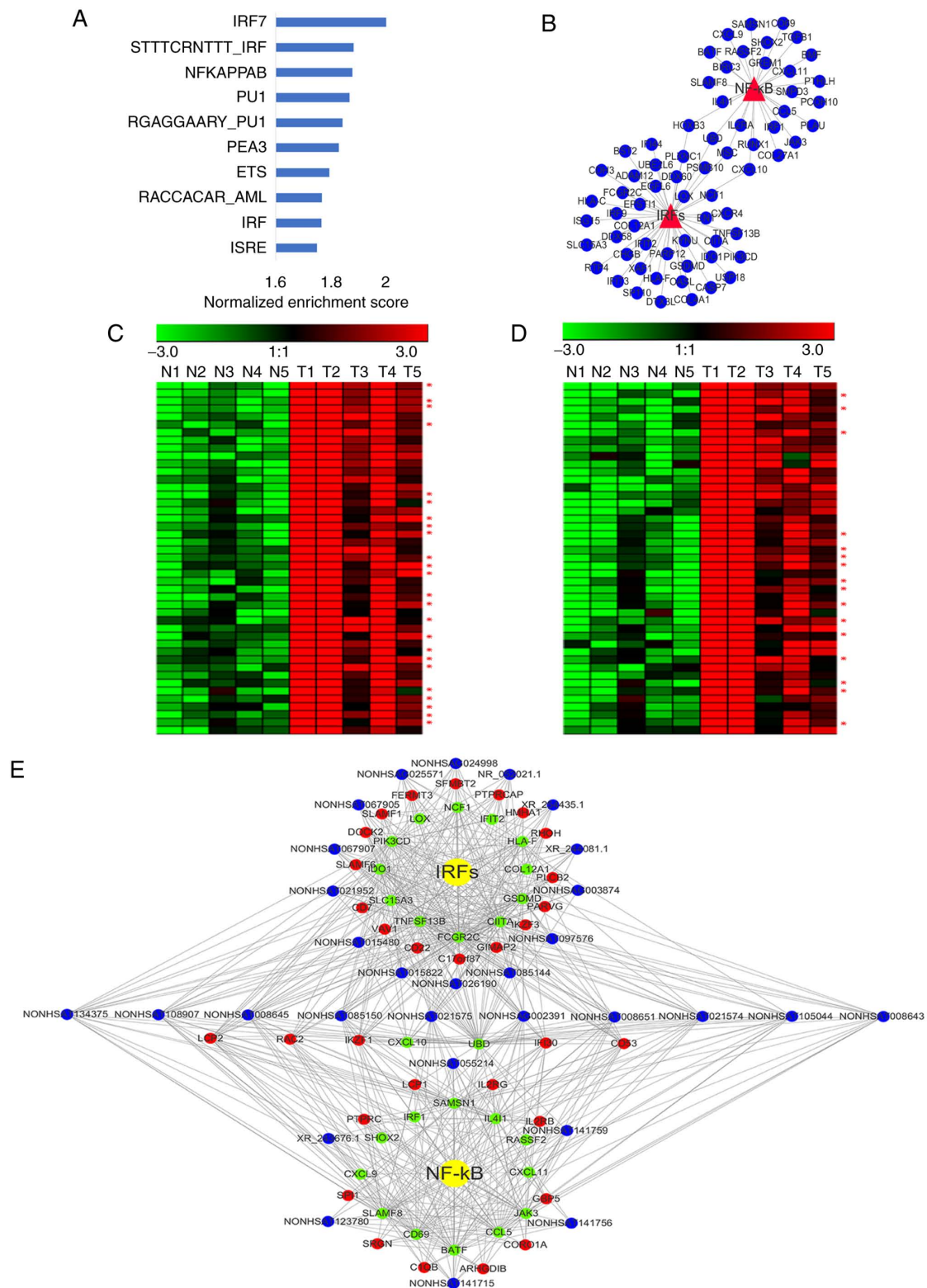


Figure 4. Prediction of IRF- and NF-κB-regulated mRNAs via GSEA, screening related lncRNAs and construction of regulatory networks. (A) The upstream regulators of differentially expressed mRNAs were predicted via GSEA. IRF- and NF-κB-regulated mRNAs were the most significantly enriched. (B) mRNA regulatory networks driven by IRFs and NF-κB. (C) The co-expression heat map of IRF-regulated mRNAs and lncRNAs (asterisks). In the heatmap, some mRNAs were derived from the reference gene set that was regulated by IRFs in the GSEA database, whereas others were significantly associated with IRFs in the WGCNA-constructed co-expression network. The lncRNAs that were significantly associated with IRFs in the WGCNA-constructed co-expression network are presented. (D) The co-expression heat map of NF-κB-regulated mRNAs and lncRNAs (asterisks). In the heatmap, some mRNAs were derived from the reference gene set that was regulated by NF-κB in the GSEA database, whereas others were significantly associated with NF-κB in the WGCNA-constructed co-expression network. The lncRNAs that were significantly associated with NF-κB in the WGCNA-constructed co-expression network are presented. (E) mRNA-lncRNA regulatory networks driven by IRFs and NF-κB. Yellow represents IRFs or NF-κB; green represents mRNAs derived from the reference gene set that was regulated by IRFs or NF-κB in the GSEA database; blue represents lncRNAs significantly associated with IRFs or NF-κB in the WGCNA-constructed co-expression network and red represents mRNAs significantly associated with IRFs or NF-κB in the WGCNA-constructed co-expression network. IRF, interferon regulatory factor; NF-κB, nuclear factor kappa B; GSEA, gene set-enrichment analysis; lncRNA, long non-coding RNA; WGCNA, weighted gene co-expression network analysis.

be regulated by IRFs. Similarly, as presented in Fig. 4D, the mRNAs and lncRNAs (Tables SIII) that were regulated by NF- κ B in TSCC have similar expression trends. Based on the IRF- and NF- κ B-regulated mRNAs enriched by GSEA and the mRNA-lncRNA co-expression network, the downstream mRNA-lncRNA networks regulated by IRFs and NF- κ B in TSCC were constructed (Fig. 4E).

Canonical pathway analysis of differentially expressed genes via IPA and construction of related regulatory-network pathways. To further understand what functions and pathways the differentially expressed lncRNAs participate in, the gene expression profile dataset was imported into IPA to identify significantly enriched biological pathways in areca nut chewing-induced TSCC (Fig. S5A). The majority of the top five pathways (Th cell activation pathway, antigen presentation pathway, role of NFAT in regulation of the immune response, GP6 signaling pathway and T cell exhaustion signaling pathway) were associated with inflammatory and immunity. Collectively, these results suggest that inflammatory immune responses play important roles in the carcinogenesis of areca nut chewing induced TSCC. Given that the Th cell activation pathway was the most significantly enrichment pathway, lncRNAs that were associated with the Th cell activation pathway genes in the WGCNA-constructed co-expression network were screened (Table SIV). The expression trends of these lncRNAs were similar to mRNAs in the Th cell activation pathway (Fig. S5B). A network of lncRNAs that potentially regulates the Th cell activation pathway was derived (Fig. 5) by integrating the lncRNAs into the Th cell activation pathway (Fig. S6).

Discussion

Areca nut has been identified as a group 1 carcinogen that is closely associated with the development of OSCC and its precancerous lesions (11). Epidemiological evidence has demonstrated that in the Indian mainland, Southeast Asia countries and Hunan province of China, extensive areca nut chewing has caused a high incidence of OSCC, including TSCC (10,24,54). In certain countries, such as Sri Lanka, India and Pakistan, OSCC is the most common cancer, accounting for 25% of all new cancer cases (24). For example, more than 100,000 cases of OSCC are registered each year in India. The age-adjusted cancer incidence rates of OSCC in India were between 3.4 per 100,000 and 6.0 per 100,000 in 2001-2002 (24). However, the molecular mechanisms underlying areca nut chewing induced TSCC remain unclear. lncRNAs are non-coding RNA transcripts that exceed 200 nucleotides, and regulate gene expression at the epigenetic, transcriptional and post-transcriptional levels (55). Increasing evidence has demonstrated that aberrant lncRNA expression and dysfunctional lncRNA activity are associated with tumor initiation and progression (56-64). However, a full transcriptome analysis of the interactions between mRNAs and lncRNAs, and related regulatory networks in areca nut chewing induced TSCC have not yet been reported.

The lncRNA and mRNA profiles from TSCC were obtained by lncRNA microarray technology to determine the potential molecular mechanisms of areca nut chewing-induced

TSCC. To determine whether there are different characteristics between the areca nut chewing and the non-areca nut chewing TSCC expression profiles, PCA analysis was performed and the results demonstrated that the components of the two data groups have obvious differences in clustering. This indicates that the expression profiles of patients with areca nut chewing and patients with non-areca nut chewing have different characteristics, and there may be specific differential expression genes involved in areca nut chewing induced TSCC. A total of 3,860 differentially expressed genes were identified, including 1,667 mRNAs and 2,193 lncRNAs via SAM analysis.

OSF is a potentially malignant disorder associated with areca nut chewing and many patients with OSCC who habitually chew areca nuts develop OSCC from OSF malignant transformation (19). Through microarray analysis, Li *et al* (65) discovered several differentially expressed genes between OSF and normal tissues, were as *Loricrin*, *COMP*, *CXCL9*, *KRT19* and *CYP3A5* were the most differentially expressed genes and may play important roles in OSF. In the present study, *CXCL9* and *CYP3A5* were also differentially expressed and the expression trend was consistent with that of the report by Li *et al* indicating that *CXCL9* and *CYP3A5* may be involved in the entire process leading from normal mucosa to OSF and TSCC. Qiu *et al* (66) also constructed a lncRNA-expression profile for OSCC tissues obtained from individuals who did not habitually chew areca nuts. A total of 2,294 differentially expressed lncRNAs were obtained and the results demonstrated that *lnc-TPP2-7:2*, *lnc-MANSC4-8:1*, *TMPRSS11BNL*, *GAS5* and *DANCR* are significantly downregulated in OSCC. The lncRNAs, *GAS5* and *DANCR* were also downregulated in the present study, suggesting that they all play causative roles in OSCC induced by areca nut chewing or other reasons. However, most lncRNAs in the present differentially expressed lncRNA profile have not been previously reported. Some of these lncRNAs may be specifically associated with areca nut chewing induced TSCC and warrant further study. Regarding the lncRNAs that have not been previously studied, it is vital to determine their functional roles.

Genome instability is one of the common characteristics of cancer (67). During carcinogenesis, exogenous stimulation may induce chromosomal rearrangement, amplification, deletion, ectopic recombination or inversion (68). GSEA can be used to identify regional chromosomal deletions or amplifications. The present study performed chromosomal-localization analysis via GSEA and demonstrated that the differentially expressed genes were predominantly enriched in the chromosomal regions 22q13, 8p21 and 3p21. The upregulated genes were significantly enriched in the 22q13 region, whereas the downregulated genes were significantly enriched in the 8p21 and 3p21 regions. Taken together, these results indicate that chromosome amplification may occur in the 22q13 region, while chromosome deletion may occur in the 8p21 and 3p21 regions. If chromosome amplification or deletion occur in these regions, the lncRNAs located in such regions should also be abnormally expressed. The annotated positional information for the lncRNA chip was analyzed and 21 additional lncRNAs that were significantly upregulated were identified in the 22q13 region. Thus, it is speculated that chromosome amplification occurred in 22q13 and several mRNAs and lncRNAs were upregulated simultaneously in this region. Consistent with this result, Janjetovic *et al* (69)

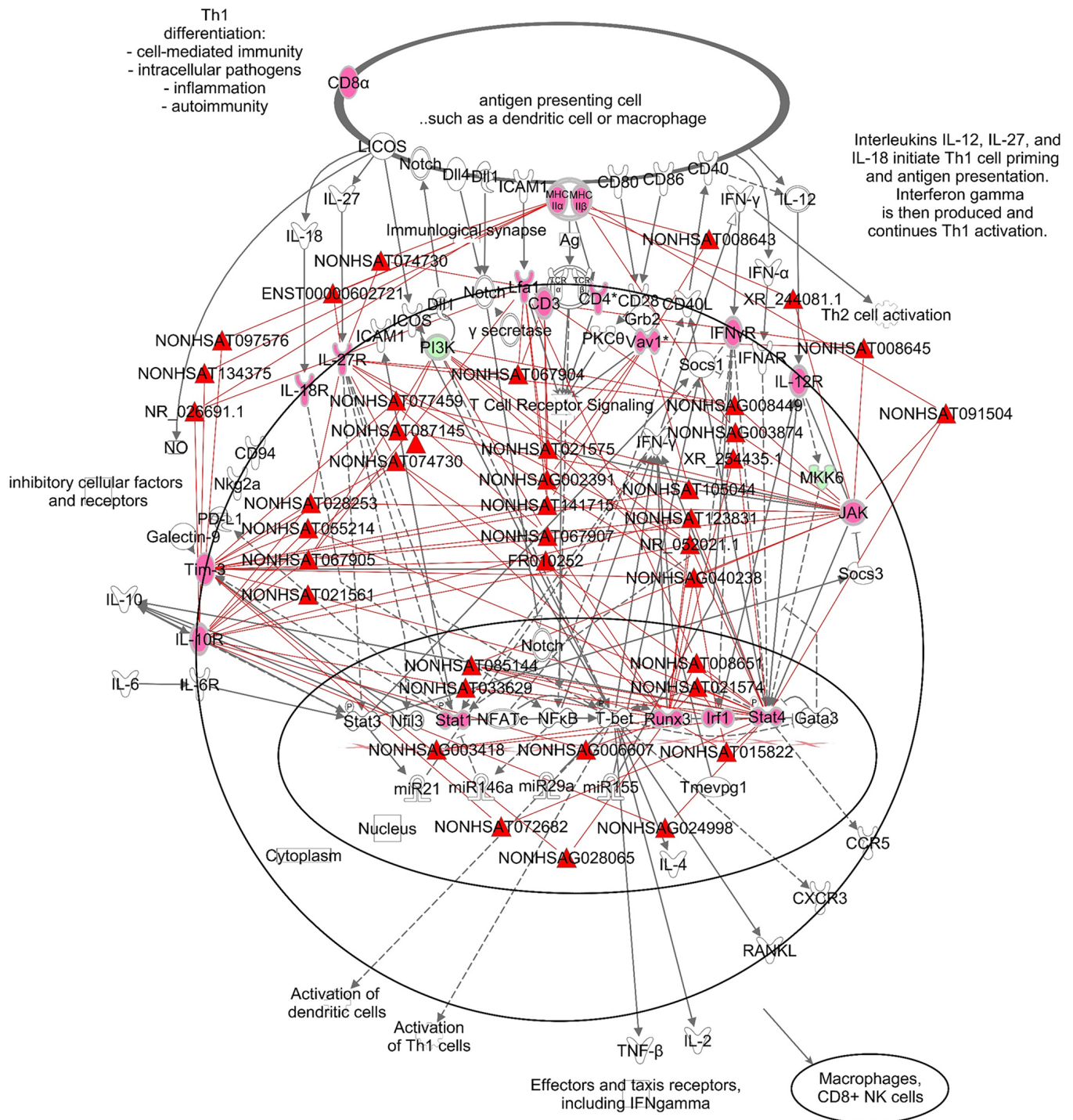


Figure 5. Network of lncRNAs involved in regulating the Th cell activation pathway in tongue squamous cell carcinoma. The lncRNAs were derived from the co-expression network and had similar expression trends compared with the mRNAs in the Th cell activation pathway. The triangles represent lncRNAs. Round and oval represent mRNAs. Red and pink represent upregulation, green represents downregulation and white represents no difference in expression. lncRNA, long non-coding RNA.

identified chromosomal amplification of 22q13 in patients with OSCC using comparative expressed sequence hybridization. In addition, Sakai *et al* (70) identified chromosomal amplification of 22q13 in patients with esophageal cancer by performing comparative genomic hybridization (CGH). The results of the present study demonstrated that multiple mRNAs and lncRNAs were significantly downregulated in the 8p21 and 3p21 regions, such as CCDC25 and FAM3D, which suggests that these 2

chromosomal segments may have been deleted in the areca nut chewing induced TSCC specimens, particularly considering that these deletions have been reported in different types of cancers, including OSCC (71-75). This finding needs to be further verified in more areca nut chewing induced TSCC tissues using CGH, fluorescence *in situ* hybridization or immunohistochemistry.

In addition to chromosomal colocalization analysis, the association between mRNAs and lncRNAs can be studied by

constructing a co-expression network based on their expression trends. WGCNA is commonly used for constructing gene co-expression networks (50,76). When constructing co-expression networks, thresholds usually need to be set to determine whether two genes have similar expression patterns (50). Unlike the traditional 'hard-threshold' algorithm, WGCNA generates connections between genes in the network that obey scale-free networks by setting soft thresholds, which is more in line with biological characteristics (50). Thus, the present study constructed a co-expression network of differentially expressed lncRNAs and mRNAs using WGCNA. However, this co-expression network only provided lncRNA-mRNA associations at the transcriptome level, and so the functions of these lncRNAs remained unclear. Thus, the set of differentially expressed genes were imported into data analysis software programs based on known gene functions, such as IPA (Ingenuity Systems, www.ingenuity.com) and GSEA (version 3.0; <http://software.broadinstitute.org/gsea/index.jsp>), to find common transcriptional regulators and cell-signaling pathways.

GSEA for known upstream transcription factor-regulated gene sets demonstrated that IRFs and NF- κ B that transcriptionally regulated gene sets were the most significantly enriched. IRFs and NF- κ B are immunoregulatory factors that regulate the expression of several genes, particularly those involved in inflammation and immune responses, such as IFN- α and TNF- α (77,78). In addition, they are closely associated with tumor initiation and progression (79-81). For example, the chemokine, CCL5 promotes the invasion and metastasis of oral cancer cells by activating NF- κ B (82). IRFs are also associated with the expression of immune evasion-related molecules, such as CTLA-4 and IDO (83,84). Jiang *et al* (84) reported that IRF-1 is an important factor for IFN- γ -induced IDO expression. These previous reports (82-84) are consistent with the results of the regulatory networks constructed in the present study, which proves that IRFs or NF- κ B regulate areca induced TSCC through these downstream mRNAs derived from the reference gene set in the GSEA database. We constructed an mRNA-lncRNA co-expression network to establish the relationship between mRNAs and lncRNAs. Thus, lncRNAs that are associated with the downstream mRNAs regulated by IRFs or NF- κ B in the co-expression network, may also be regulated by IRFs or NF- κ B. A lncRNA and mRNA co-expression network was constructed to determine which lncRNAs are regulated by IRFs or NF- κ B and the results identified 30 lncRNAs that were associated with IRFs or NF- κ B, including *NONHSAT085150*. However, the functions of these lncRNAs have not yet been reported, and how these lncRNAs participate in IRF- and NF- κ B-driven downstream mRNA-lncRNA-signaling pathways or promote TSCC merits further investigation.

IPA pathway enrichment analysis demonstrated that four of the top five pathways were associated with inflammation and immunity, including the Th cell activation pathway, antigen presentation pathway, role of *NFAT* in regulation of the immune response pathway and T cell exhaustion signaling pathway. These pathways differ from those of patients with OSCC that do not habitually chew areca nuts. Qiu *et al* (66) reported that the differentially expressed genes in patients with OSCC who do not chew areca nut are predominantly enriched in metabolic pathways, such as steroid biosynthesis,

glycosaminoglycan biosynthesis and the synthesis and degradation of ketone bodies. These findings suggest that immune inflammatory responses may play more important carcinogenic roles in patients with oral cancer who chew areca nuts compared with patients who develop OSCC due to other causes. Autoimmune reaction is one of the pathogenic factors of OSF which is the main precancerous lesion of areca nut chewing induced OSCC (14). Long-term stimulation of exogenous antigens, and being repeatedly damaged and repaired places the oral mucosa in a state of chronic inflammation in patients who have chewed areca nuts for a long period of time (14,85). Long-term unresolved inflammation triggers cancer progression (86,87), thus, inflammation-mediated carcinogenesis may be one of the pathogenic factors of areca nut chewing induced TSCC. In addition to chronic inflammation, immune escape may also be one of the molecular mechanisms underlying areca nut chewing induced TSCC (88,89). Immune evasion plays important roles in oncogenesis and tumor development by protecting tumors from attack by the immune system (90,91). Th cell activation, antigen extraction and T cell exhaustion play important roles in immune escape, particularly for exhausted T cells (90). Exhausted T cells participate in the negative regulation of tumor immunity through inhibitory molecules and cytokines on the cell surface, thus causing tumor immune escape (90-101). In addition, some T cell exhaustion-related immune checkpoints were upregulated in the present study, such as *CTLA4*, *TIGIT* and *Tim3*. The differential expression of these immune checkpoints further suggests that immune escape occurs during areca nut chewing induced carcinogenesis. The present study aimed to investigate whether lncRNAs play important roles in these pathways via construction of a WGCNA co-expression network, and the results demonstrated that 74 lncRNAs were significantly associated with Th cell-associated pathways, including *NR_026691.1*. These 74 lncRNAs have similar expression trends to those of genes involved in the Th cell pathway, indicating that they may have the same function (48,49). Thus, these 74 lncRNAs were selected to study the regulatory networks that they may be involved in. However, the functions of these lncRNAs have not yet been reported, and their underlying molecular mechanisms require further investigation.

To the best of our knowledge, the present study was the first to investigate the differential lncRNA-expression profiles and the lncRNAs involved in signaling pathways and regulatory networks in areca nut chewing induced TSCC. The present study identified 1,667 differentially expressed mRNAs and 2,193 differentially expressed lncRNAs in five pairs of areca nut chewing induced TSCC and adjacent normal tissues via lncRNA gene chip detection and analysis. The majority of the differentially expressed lncRNAs have not yet been studied. Construction and analysis of a lncRNA-mRNA co-expression network generated a preliminary set of lncRNAs involved in biological signaling pathways and regulatory networks. The results of the present study provided insight for further investigation to determine the molecular mechanisms underlying areca nut chewing induced TSCC, provided candidates for selecting potential molecular markers and provided potential targets for clinical interventions.

A potential limitation of this study is the small sample size. Some genes that are less different between cancer and adjacent

normal tissues but are important for the induction of cancer by areca nut may not have been screened. In addition, due to the small sample size, it is easy to screen some genes with false positive differences. With this regard, larger sample study will be carried out in the future to verify our results.

Acknowledgements

The authors would like to thank Mr. Yanshan Ge (NHC Key Laboratory of Carcinogenesis, Cancer Research Institute and School of Basic Medical Sciences, Central South University, Changsha, China) and Mr. Zhihao Qu (Key Laboratory of Tea Science of Ministry of Education, Hunan Agricultural University, Changsha, China) for their technical support.

Funding

The present study was supported by the National Natural Science Foundation of China (grant nos. 81803025 and 81772901) and the “225” Project of Hunan Province High Level Health Talents.

Availability of data and materials

The datasets used and analyzed during the present study are available from the corresponding author upon reasonable request.

Authors' contributions

ZG conceived and designed the experiments. JL, YX, PL, SZ and YM performed the experiments. PL, SZ, YM, LZ, YW, FX and SZ analyzed the data. PL, JL, YX and ZZ drafted the initial manuscript. ZZ, WX, YL and ZG contributed significantly to data analyses and manuscript revision. All authors read and approved the final 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 Cancer Research Institute of Central South University (Changsha, China) and written informed consent was provided by all patients.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68: 394-424, 2018.
- Johnson NW, Jayasekara P and Amarasinghe AA: Squamous cell carcinoma and precursor lesions of the oral cavity: Epidemiology and aetiology. *Periodontol* 57: 19-37, 2020.
- Taniguchi Y and Okura M: Prognostic significance of perioperative blood transfusion in oral cavity squamous cell carcinoma. *Head Neck* 25: 931-936, 2003.
- Koyama S, Tabuchi T, Okawa S, Morishima T, Ishimoto S, Ishibashi M and Miyashiro I: Oral cavity cancer incidence rates in Osaka, Japan between 2000 and 2014. *Oral Oncol* 105: 104653, 2020.
- Wu C, Li M, Meng H, Liu Y, Niu W, Zhou Y, Zhao R, Duan Y, Zeng Z, Li X, *et al*: Analysis of status and countermeasures of cancer incidence and mortality in China. *Sci China Life Sci* 62: 640-647, 2019.
- Leemans CR, Braakhuis BJ and Brakenhoff RH: The molecular biology of head and neck cancer. *Nat Rev Cancer* 11: 9-22, 2011.
- Yu J, Liu Y, Gong Z, Zhang S, Guo C, Li X, Tang Y, Yang L, He Y, Wei F, *et al*: Overexpression long non-coding RNA LINC00673 is associated with poor prognosis and promotes invasion and metastasis in tongue squamous cell carcinoma. *Oncotarget* 8: 16621-16632, 2017.
- Yu J, Liu Y, Guo C, Zhang S, Gong Z, Tang Y, Yang L, He Y, Lian Y, Li X, *et al*: Upregulated long non-coding RNA LINC00152 expression is associated with progression and poor prognosis of tongue squamous cell carcinoma. *J Cancer* 8: 523-530, 2017.
- De Angelis R, Sant M, Coleman MP, Francisci S, Baili P, Pierannunzio D, Trama A, Visser O, Brenner H, Ardanaz E, *et al*: Cancer survival in Europe 1999-2007 by country and age: Results of EURO-CARE-5-a population-based study. *Lancet Oncol* 15: 23-34, 2014.
- Shield KD, Ferlay J, Jemal A, Sankaranarayanan R, Chaturvedi AK, Bray F and Soerjomataram I: The global incidence of lip, oral cavity, and pharyngeal cancers by subsite in 2012. *CA Cancer J Clin* 67: 51-64, 2017.
- Chen PH, Mahmood Q, Mariottini GL, Chiang TA and Lee KW: Adverse health effects of betel quid and the risk of oral and pharyngeal cancers. *Biomed Res Int* 2017: 3904098, 2017.
- Chuerduangphui J, Ekalaksananan T, Chaiyapit P, Patarapadungkit N, Chotiyano A, Kongyingyoes B, Promthet S and Pientong C: Effects of arecoline on proliferation of oral squamous cell carcinoma cells by dysregulating c-Myc and miR-22, directly targeting oncostatin M. *PLoS One* 13: e0192009, 2018.
- Li YC, Cheng AJ, Lee LY, Huang YC and Chang JT: Multifaceted mechanisms of areca nuts in oral carcinogenesis: The molecular pathology from precancerous condition to malignant transformation. *J Cancer* 10: 4054-4062, 2019.
- Shih YH, Wang TH, Shieh TM and Tseng YH: Oral submucous fibrosis: A review on etiopathogenesis, diagnosis, and therapy. *Int J Mol Sci* 20: 2940, 2019.
- Arakeri G, Rai KK, Hunasgi S, Merks MAW, Gao S and Brennan PA: Oral submucous fibrosis: An update on current theories of pathogenesis. *J Oral Pathol Med* 46: 406-412, 2017.
- Wang YY, Tail YH, Wang WC, Chen CY, Kao YH, Chen YK and Chen CH: Malignant transformation in 5071 southern Taiwanese patients with potentially malignant oral mucosal disorders. *BMC Oral Health* 14: 99, 2014.
- Merchant A, Husain SS, Hosain M, Fikree FF, Pitiphat W, Siddiqui AR, Hayder SJ, Haider SM, Ikram M, Chuang SK and Saeed SA: Paan without tobacco: An independent risk factor for oral cancer. *Int J Cancer* 86: 128-131, 2000.
- Srinivasan M and Jewell SD: Evaluation of TGF- α and EGFR expression in oral leukoplakia and oral submucous fibrosis by quantitative immunohistochemistry. *Oncology* 61: 284-292, 2001.
- Ray JG, Ranganathan K and Chattopadhyay A: Malignant transformation of oral submucous fibrosis: Overview of histopathological aspects. *Oral Surg Oral Med Oral Pathol Oral Radiol* 122: 200-209, 2016.
- Arora S and Squier C: Areca nut trade, globalisation and its health impact: Perspectives from India and South-east Asia. *Perspect Public Health* 139: 44-48, 2019.
- Nigam NK, Aravinda K, Dhillon M, Gupta S, Reddy S and Srinivas Raju M: Prevalence of oral submucous fibrosis among habitual gutkha and areca nut chewers in Moradabad district. *J Oral Biol Craniofac Res* 4: 8-13, 2014.
- Pindborg JJ, Mehta FS, Gupta PC and Daftary DK: Prevalence of oral submucous fibrosis among 50,915 Indian villagers. *Br J Cancer* 22: 646-654, 1968.
- Hazarey VK, Erlewad DM, Mundhe KA and Ughade SN: Oral submucous fibrosis: Study of 1,000 cases from central India. *J Oral Pathol Med* 36: 12-17, 2007.

24. Warnakulasuriya S: Global epidemiology of oral and oropharyngeal cancer. *Oral Oncol* 45: 309-316, 2009.
25. Zhang SS, Li WH, Gao YJ, Liu ZW, Liu L, Tang JQ and Ling TY: Betel-quid and oral submucous fibrosis: A cross-sectional study in Hunan province, China. *J Oral Pathol Med* 41: 748-754, 2012.
26. Liu B, Shen M, Xiong J, Yuan Y, Wu X, Gao X, Xu J, Guo F and Jian X: Synergistic effects of betel quid chewing, tobacco use (in the form of cigarette smoking), and alcohol consumption on the risk of malignant transformation of oral submucous fibrosis (OSF): A case-control study in Hunan province, China. *Oral Surg Oral Med Oral Pathol Oral Radiol* 120: 337-345, 2015.
27. Hu Y, Zhong R, Li H and Zou Y: Effects of betel quid, smoking and alcohol on oral cancer risk: A case-control study in Hunan province, China. *Subst Use Misuse* 55: 1501-1508, 2020.
28. Levine AJ: p53: 800 million years of evolution and 40 years of discovery. *Nat Rev Cancer* 20: 471-480, 2020.
29. Baluapuri A, Wolf E and Eilers M: Target gene-independent functions of MYC oncoproteins. *Nat Rev Mol Cell Biol* 21: 255-267, 2020.
30. Hanahan D and Weinberg RA: The hallmarks of cancer. *Cell* 100: 57-70, 2000.
31. Zhang L, Meng X, Zhu XW, Yang DC, Chen R, Jiang Y and Xu T: Long non-coding RNAs in Oral squamous cell carcinoma: Biologic function, mechanisms and clinical implications. *Mol Cancer* 18: 102, 2019.
32. Yang L, Tang Y, He Y, Wang Y, Lian Y, Xiong F, Shi L, Zhang S, Gong Z, Zhou Y, *et al*: High expression of LINC01420 indicates an unfavorable prognosis and modulates cell migration and invasion in nasopharyngeal carcinoma. *J Cancer* 8: 97-103, 2017.
33. Fan C, Tang Y, Wang J, Xiong F, Guo C, Wang Y, Zhang S, Gong Z, Wei F, Yang L, *et al*: Role of long non-coding RNAs in glucose metabolism in cancer. *Mol Cancer* 16: 130, 2017.
34. He Y, Jing Y, Wei F, Tang Y, Yang L, Luo J, Yang P, Ni Q, Pang J, Liao Q, *et al*: Long non-coding RNA PVT1 predicts poor prognosis and induces radioresistance by regulating DNA repair and cell apoptosis in nasopharyngeal carcinoma. *Cell Death Dis* 9: 235, 2018.
35. Wang Y, Mo Y, Yang X, Zhou R, Wu Z, He Y, Yang X, Zhong Y, Du Y, Zhou H, *et al*: Long non-coding RNA AFAP1-AS1 is a novel biomarker in various cancers: A systematic review and meta-analysis based on the literature and GEO datasets. *Oncotarget* 8: 102346-102360, 2017.
36. Tang Y, He Y, Zhang P, Wang J, Fan C, Yang L, Xiong F, Zhang S, Gong Z, Nie S, *et al*: LncRNAs regulate the cytoskeleton and related Rho/ROCK signaling in cancer metastasis. *Mol Cancer* 17: 77, 2018.
37. Lian Y, Xiong F, Yang L, Bo H, Gong Z, Wang Y, Wei F, Tang Y, Li X, Liao Q, *et al*: Long noncoding RNA AFAP1-AS1 acts as a competing endogenous RNA of miR-423-5p to facilitate nasopharyngeal carcinoma metastasis through regulating the Rho/Rac pathway. *J Exp Clin Cancer Res* 37: 253, 2018.
38. Fan C, Tang Y, Wang J, Wang Y, Xiong F, Zhang S, Li X, Xiang B, Wu X, Guo C, *et al*: Long non-coding RNA LOC284454 promotes migration and invasion of nasopharyngeal carcinoma via modulating the Rho/Rac signaling pathway. *Carcinogenesis* 40: 380-391, 2019.
39. Bo H, Fan L, Gong Z, Liu Z, Shi L, Guo C, Li X, Liao Q, Zhang W, Zhou M, *et al*: Upregulation and hypomethylation of lncRNA AFAP1-AS1 predicts a poor prognosis and promotes the migration and invasion of cervical cancer. *Oncol Rep* 41: 2431-2439, 2019.
40. Wei F, Jing YZ, He Y, Tang YY, Yang LT, Wu YF, Tang L, Shi L, Gong ZJ, Guo C, *et al*: Cloning and characterization of the putative AFAP1-AS1 promoter region. *J Cancer* 10: 1145-1153, 2019.
41. Vuity D, McMahon J, Takhiuddin S, Slinger C, McLellan D, Wales C, MacIver C, Thomson E, McCaul J, Hislop S, *et al*: Is the 8th edition of the union for international cancer control staging of oral cancer good enough? *Br J Oral Maxillofac Surg* 56: 272-277, 2018.
42. van Schothorst EM, Pagmantidis V, de Boer VC, Hesketh J and Keijzer J: Assessment of reducing RNA input for Agilent oligo microarrays. *Anal Biochem* 363: 315-317, 2007.
43. Jolliffe IT and Cadima J: Principal component analysis: A review and recent developments. *Philos Trans A Math Phys Eng Sci* 374: 20150202, 2016.
44. DiLeo MV, Strahan GD, den Bakker M and Hoekenga OA: Weighted correlation network analysis (WGCNA) applied to the tomato fruit metabolome. *PLoS One* 6: e26683, 2011.
45. Rentoft M, Coates PJ, Laurell G and Nylander K: Transcriptional profiling of formalin fixed paraffin embedded tissue: Pitfalls and recommendations for identifying biologically relevant changes. *PLoS One* 7: e35276, 2012.
46. Subramanian A, Kuehn H, Gould J, Tamayo P and Mesirov JP: GSEA-P: A desktop application for Gene set enrichment analysis. *Bioinformatics* 23: 3251-3253, 2007.
47. He W, Fu L, Yan Q, Zhou Q, Yuan K, Chen L and Han Y: Gene set enrichment analysis and meta-analysis identified 12 key genes regulating and controlling the prognosis of lung adenocarcinoma. *Oncol Lett* 17: 5608-5618, 2019.
48. Jin X, Li J, Li W, Wang X, Du C, Geng Z, Geng Y, Kang L, Zhang X, Wang M and Tian S: Weighted gene Co-expression network analysis reveals specific modules and biomarkers in Parkinson's disease. *Neurosci Lett* 728: 134950, 2020.
49. Zhang B and Horvath S: A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 4: Article17, 2005.
50. Langfelder P and Horvath S: WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* 9: 559, 2008.
51. Krämer A, Green J, Pollard J Jr and Tugendreich S: Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* 30: 523-530, 2014.
52. Lv H, Liu L, Zhang Y, Song T, Lu J and Chen X: Ingenuity pathways analysis of urine metabolomics phenotypes toxicity of gentamicin in multiple organs. *Mol Biosyst* 6: 2056-2067, 2010.
53. Ooi AT, Gower AC, Zhang KX, Vick JL, Hong L, Nagao B, Wallace WD, Elashoff DA, Walser TC, Dubinett SM, *et al*: Molecular profiling of premalignant lesions in lung squamous cell carcinomas identifies mechanisms involved in stepwise carcinogenesis. *Cancer Prev Res (Phila)* 7: 487-495, 2014.
54. Kharbanda OP, Ivaturi A, Priya H, Dorji G and Gupta S: Digital possibilities in the prevention and early detection of oral cancer in the WHO South-East Asia Region. *WHO South East Asia J Public Health* 8: 95-100, 2019.
55. Marchese FP, Raimondi I and Huarte M: The multidimensional mechanisms of long noncoding RNA function. *Genome Biol* 18: 206, 2017.
56. Lin C and Yang L: Long Noncoding RNA in cancer: Wiring signaling circuitry. *Trends Cell Biol* 28: 287-301, 2018.
57. Wang D, Zeng Z, Zhang S, Xiong F, He B, Wu Y, Li W, Tang L, Wei F, Xiang B, *et al*: Epstein-Barr virus-encoded miR-BART6-3p inhibits cancer cell proliferation through the LOC553103-STMN1 axis. *FASEB J* 34: 8012-8027, 2020.
58. Ban Y, Tan P, Cai J, Li J, Hu M, Zhou Y, Mei Y, Tan Y, Li X, Zeng Z, *et al*: LNCAROD is stabilized by m6A methylation and promotes cancer progression via forming a ternary complex with HSPA1A and YBX1 in head and neck squamous cell carcinoma. *Mol Oncol* 14: 1282-1296, 2020.
59. Wu P, Mo Y, Peng M, Tang T, Zhong Y, Deng X, Xiong F, Guo C, Wu X, Li Y, *et al*: Emerging role of tumor-related functional peptides encoded by lncRNA and circRNA. *Mol Cancer* 19: 22, 2020.
60. Wang W, Zhou R, Wu Y, Liu Y, Su W, Xiong W and Zeng Z: PVT1 promotes cancer progression via MicroRNAs. *Front Oncol* 9: 609, 2019.
61. Jin K, Wang S, Zhang Y, Xia M, Mo Y, Li X, Li G, Zeng Z, Xiong W and He Y: Long non-coding RNA PVT1 interacts with MYC and its downstream molecules to synergistically promote tumorigenesis. *Cell Mol Life Sci* 76: 4275-4289, 2019.
62. Fan CM, Wang JP, Tang YY, Zhao J, He SY, Xiong F, Guo C, Xiang B, Zhou M, Li XL, *et al*: circMAN1A2 could serve as a novel serum biomarker for malignant tumors. *Cancer Sci* 110: 2180-2188, 2019.
63. Wang D, Tang L, Wu Y, Fan C, Zhang S, Xiang B, Zhou M, Li X, Li Y, Li G, *et al*: Abnormal X chromosome inactivation and tumor development. *Cell Mol Life Sci* 77: 2949-2958, 2020.
64. Tang T, Yang L, Cao Y, Wang M, Zhang S, Gong Z, Xiong F, He Y, Zhou Y, Liao Q, *et al*: LncRNA AATBC regulates Pinin to promote metastasis in nasopharyngeal carcinoma. *Mol Oncol* 14: 2251-2270, 2020.
65. Li N, Jian X, Hu Y, Xu C, Yao Z and Zhong X: Discovery of novel biomarkers in oral submucous fibrosis by microarray analysis. *Cancer Epidemiol Biomarkers Prev* 17: 2249-2259, 2008.
66. Qiu YL, Liu YH, Ban JD, Wang WJ, Han M, Kong P and Li BH: Pathway analysis of a genomewide association study on a long noncoding RNA expression profile in oral squamous cell carcinoma. *Oncol Rep* 41: 895-907, 2019.
67. Hanahan D and Weinberg RA: Hallmarks of cancer: The next generation. *Cell* 144: 646-674, 2011.

68. Potapova TA, Zhu J and Li R: Aneuploidy and chromosomal instability: A vicious cycle driving cellular evolution and cancer genome chaos. *Cancer Metastasis Rev* 32: 377-389, 2013.
69. Janjetovic S, Sticht C, Knoepfle K, Joos S, Hofele C, Lichter P and Freier K: Comparative expressed sequence hybridization detects recurrent patterns of altered sequence expression in oral squamous cell carcinoma. *Oncol Rep* 24: 369-374, 2010.
70. Sakai N, Kajiyama Y, Iwanuma Y, Tomita N, Amano T, Isayama F, Ouchi K and Tsurumaru M: Study of abnormal chromosome regions in esophageal squamous cell carcinoma by comparative genomic hybridization: Relationship of lymph node metastasis and distant metastasis to selected abnormal regions. *Dis Esophagus* 23: 415-421, 2010.
71. Freier K, Knoepfle K, Flechtenmacher C, Pungs S, Devens F, Toedt G, Hofele C, Joos S, Lichter P and Radlwimmer B: Recurrent copy number gain of transcription factor SOX2 and corresponding high protein expression in oral squamous cell carcinoma. *Genes Chromosomes Cancer* 49: 9-16, 2010.
72. Cao Y, Liu Y, Yang X, Liu X, Han N, Zhang K and Lin D: Estimation of the survival of patients with lung squamous cell carcinoma using genomic copy number aberrations. *Clin Lung Cancer* 17: 68-74.e5, 2016.
73. Yoshikawa Y, Emi M, Hashimoto-Tamaoki T, Ohmuraya M, Sato A, Tsujimura T, Hasegawa S, Nakano T, Nasu M, Pastorino S, *et al*: High-density array-CGH with targeted NGS unmask multiple noncontiguous minute deletions on chromosome 3p21 in mesothelioma. *Proc Natl Acad Sci USA* 113: 13432-13437, 2016.
74. Dimova I, Orsetti B, Negre V, Rouge C, Ursule L, Lasorsa L, Dimitrov R, Doganov N, Toncheva D and Theillet C: Genomic markers for ovarian cancer at chromosomes 1, 8 and 17 revealed by array CGH analysis. *Tumori* 95: 357-366, 2009.
75. Uchida K, Oga A, Okafuji M, Mihara M, Kawauchi S, Furuya T, Chochi Y, Ueyama Y and Sasaki K: Molecular cytogenetic analysis of oral squamous cell carcinomas by comparative genomic hybridization, spectral karyotyping, and fluorescence in situ hybridization. *Cancer Genet Cytogenet* 167: 109-116, 2006.
76. Tang J, Kong D, Cui Q, Wang K, Zhang D, Gong Y and Wu G: Prognostic genes of breast cancer identified by gene co-expression network analysis. *Front Oncol* 8: 374, 2018.
77. Antonczyk A, Krist B, Sajek M, Michalska A, Piaszyk-Borychowska A, Plens-Galaska M, Wesoly J and Bluysen HAR: Direct inhibition of IRF-dependent transcriptional regulatory mechanisms associated with disease. *Front Immunol* 10: 1176, 2019.
78. Taniguchi K and Karin M: NF- κ B, inflammation, immunity and cancer: Coming of age. *Nat Rev Immunol* 18: 309-324, 2018.
79. Monisha J, Roy NK, Bordoloi D, Kumar A, Golla R, Kotoky J, Padmavathi G and Kunnumakkara AB: Nuclear Factor Kappa B: A potential target to persecute head and neck cancer. *Curr Drug Targets* 18: 232-253, 2017.
80. Alsamman K and El-Masry OS: Interferon regulatory factor 1 inactivation in human cancer. *Bioscience Rep* 38: BSR20171672, 2018.
81. Manzella L, Tirro E, Pennisi MS, Massimino M, Stella S, Romano C, Vitale SR and Vigneri P: Roles of interferon regulatory factors in chronic myeloid leukemia. *Curr Cancer Drug Targets* 16: 594-605, 2016.
82. Chuang JY, Yang WH, Chen HT, Huang CY, Tan TW, Lin YT, Hsu CJ, Fong YC and Tang CH: CCL5/CCR5 axis promotes the motility of human oral cancer cells. *J Cell Physiol* 220: 418-426, 2009.
83. Pick J, Arra A, Lingel H, Hegel JK, Huber M, Nishanth G, Jorch G, Fischer KD, Schlüter D, Tedford K and Brunner-Weinzierl MC: CTLA-4 (CD152) enhances the Tc17 differentiation program. *Eur J Immunol* 44: 2139-2152, 2014.
84. Jiang GM, Wang HS, Du J, Ma WF, Wang H, Qiu Y, Zhang QG, Xu W, Liu HF and Liang JP: Bortezomib relieves immune tolerance in nasopharyngeal carcinoma via STAT1 suppression and indoleamine 2,3-dioxygenase downregulation. *Cancer Immunol Res* 5: 42-51, 2017.
85. Haque MF, Harris M, Meghji S and Speight PM: An immunohistochemical study of oral submucous fibrosis. *J Oral Pathol Med* 26: 75-82, 1997.
86. Balkwill F and Mantovani A: Inflammation and cancer: Back to Virchow? *Lancet* 357: 539-545, 2001.
87. Demaria S, Pikarsky E, Karin M, Coussens LM, Chen YC, El-Omar EM, Trinchieri G, Dubinett SM, Mao JT, Szabo E, *et al*: Cancer and inflammation: Promise for biologic therapy. *J Immunother* 33: 335-351, 2010.
88. Huang W, Zhou X, Liao Q, Tang Y, Zuo L, Wang H, Chen X, Chen X, Li J, Zhu S, *et al*: Clinicopathological and prognostic significance of PD-1/PD-L1 axis expression in patients with tongue squamous cell carcinoma. *J Cell Physiol* 235: 6942-6953, 2020.
89. Kujan O, van Schaijk B and Farah CS: Immune checkpoint inhibitors in oral cavity squamous cell carcinoma and oral potentially malignant disorders: A systematic review. *Cancers (Basel)* 12: 1937, 2020.
90. Wherry EJ and Kurachi M: Molecular and cellular insights into T cell exhaustion. *Nat Rev Immunol* 15: 486-499, 2015.
91. Ren D, Hua Y, Yu B, Ye X, He Z, Li C, Wang J, Mo Y, Wei X, Chen Y, *et al*: Predictive biomarkers and mechanisms underlying resistance to PD1/PD-L1 blockade cancer immunotherapy. *Mol Cancer* 19: 19, 2020.
92. Zhao J, Guo C, Xiong F, Yu J, Ge J, Wang H, Liao Q, Zhou Y, Gong Q, Xiang B, *et al*: Single cell RNA-seq reveals the landscape of tumor and infiltrating immune cells in nasopharyngeal carcinoma. *Cancer Lett* 477: 131-143, 2020.
93. Peng M, Mo Y, Wang Y, Wu P, Zhang Y, Xiong F, Guo C, Wu X, Li Y, Li X, *et al*: Neoantigen vaccine: An emerging tumor immunotherapy. *Mol Cancer* 18: 128, 2019.
94. Xiong F, Deng S, Huang HB, Li XY, Zhang WL, Liao QJ, Ma J, Li XL, Xiong W, Li GY, *et al*: Effects and mechanisms of innate immune molecules on inhibiting nasopharyngeal carcinoma. *Chin Med J (Engl)* 132: 749-752, 2019.
95. Duan S, Guo W, Xu Z, He Y, Liang C, Mo Y, Wang Y, Xiong F, Guo C, Li Y, *et al*: Natural killer group 2D receptor and its ligands in cancer immune escape. *Mol Cancer* 18: 29, 2019.
96. Jiang X, Wang J, Deng X, Xiong F, Ge J, Xiang B, Wu X, Ma J, Zhou M, Li X, *et al*: Role of the tumor microenvironment in PD-L1/PD-1-mediated tumor immune escape. *Mol Cancer* 18: 10, 2019.
97. Wang YA, Li XL, Mo YZ, Fan CM, Tang L, Xiong F, Guo C, Xiang B, Zhou M, Ma J, *et al*: Effects of tumor metabolic microenvironment on regulatory T cells. *Mol Cancer* 17: 168, 2018.
98. Tang Y, He Y, Shi L, Yang L, Wang J, Lian Y, Fan C, Zhang P, Guo C, Zhang S, *et al*: Co-expression of AFAP1-AS1 and PD-1 predicts poor prognosis in nasopharyngeal carcinoma. *Oncotarget* 8: 39001-39011, 2017.
99. Zhou B, Yuan Y, Zhang S, Guo C, Li X, Li G, Xiong W and Zeng Z: Intestinal flora and disease mutually shape the regional immune system in the intestinal tract. *Front Immunol* 11: 575, 2020.
100. Zhang L, Sun S, Wang Y, Mo Y, Xiong F, Zhang S, Zeng Z, Xiong W, Li G, Chen H and Guo C: Gossypol induces apoptosis of multiple myeloma cells through the JUN-JNK pathway. *Am J Cancer Res* 10: 870-883, 2020.
101. Mo Y, Wang Y, Zhang L, Yang L, Zhou M, Li X, Li Y, Li G, Zeng Z, Xiong W, *et al*: The role of Wnt signaling pathway in tumor metabolic reprogramming. *J Cancer* 10: 3789-3797, 2019.