*J Biosci* (2020) 45:145
DOI: 10.1007/s12038-020-00118-2


# Expression of H19 long non-coding RNA is down-regulated in oral squamous cell carcinoma

SUPRIYA VISHWAKARMA[1], RITU PANDEY[1], RENU SINGH[2], RAGINI GOTHALWAL[3] and ASHOK KUMAR[1]*

[1]*Department of Biochemistry, All India Institute of Medical Sciences (AIIMS), Saket Nagar, Bhopal 462 020, India*

[2]*Jawaharlal Nehru Cancer Hospital and Research Centre (JNCHRC), Idgah Hills, Bhopal, India*

[3]*Department of Biotechnology, Barkatullah University, Hoshangabad Road, Bhopal, India*

**Corresponding author (Email, ashok.biochemistry@aiimsbhopal.edu.in)*

*MS received 13 June 2020; accepted 23 October 2020; published online 23 November 2020*

Long non-coding RNAs (lncRNAs) are a group of non-protein-coding RNAs which are longer than 200 nucleotides. LncRNAs play important roles in epigenetic modification, transcription and post-transcriptional regulation, maintenance of normal tissue development and differentiation. LncRNA could serve as a biomarker for diagnosis and prognosis as well as a molecular target for therapy in oral squamous cell carcinoma (OSCC). Therefore, we have determined the expression profile of 5-lncRNAs namely UCA1, TUG1, HOTAIR, MALAT1, and H19 by quantitative real-time PCR in tumor tissues and adjacent normal tissue of 32 OSCC patients. To determine the expression, methylation status and genomic alterations in lncRNAs across pan-cancer, TCGA datasets were analyzed by UALCAN, MEXPRESS and cBioPortal database. Then, we determined the association between lncRNA expression and clinicopathological attributes of patients by Spearman's rank test. Expression of *UCA1* and *TUG1* genes was up-regulated in 54.83% and 53.12% OSCC tumors, respectively. Importantly, expression of *MALAT1* and *H19* was down-regulated in tumor tissues of 62.5% and 81.25% respectively of OSCC patients. Except for MALAT1, our experimental data showed concordance with the TCGA analysis. Expression of HOTAIR in OSCC tumors was positively correlated with tumor volume, whereas MALAT1 and H19 negatively correlated with the smoking status of patients.

**Keywords.** H19; long non-coding RNAs; oral cancer; oral squamous cell carcinoma; TCGA analysis

## 1. Introduction

Head and neck cancer (HNC) is the sixth most common cancer in the world. HNC arises in the oral cavity, pharynx, larynx, paranasal cavity, nasal cavity and salivary glands (Bray *et al.* 2018). As per GLOBOCON 2018 data, the age-standardized rate per 100,000 for cancer of the lip and oral cavity in males is twice in developing countries compared to developed countries (Bray *et al.* 2018). The majority of oral cavity, oropharyngeal, hypopharyngeal and laryngeal cancers are squamous cell carcinoma (SCC) in histology and are referred to as oral squamous cell carcinoma (OSCC). OSCC is the most common cancer in Indian men, accounting for one-sixth of all cancers (Mallath *et al.* 2014). The global incidence of OSCC is more than 350,000 (Bray *et al.* 2018), out of which approximately one-fourth of the cases are contributed by South-central Asia (Mallath *et al.* 2014). Rapid advances in diagnosis, early management protocols, and widespread availability of prognostic markers have not increased the 5-year survival rate. Depending on the tumor site and Tumor Node Metastasis (TNM) staging at the time of diagnosis, the average 5-year survival rate for OSCC patients is between 40% to 50% (*www.cancer.org*). A study from northern India has

*Electronic supplementary material: The online version of this article (https://doi.org/10.1007/s12038-020-00118-2) contains supplementary material, which is available to authorized users.*

reported the 3-year disease-free survival of 50%, and the 5-year overall survival (OS) of 42% (Jain *et al.* 2019).

Long non-coding RNAs (lncRNAs) are a group of non-protein-coding RNAs which are >200 nucleotides long. Approximately only 2% of the human genome is transcribed into mRNAs, while the rest of the genes are transcribed into non-coding RNAs. LncRNAs regulates epigenetic modification, transcription, post-transcriptional processing and differentiation (Qiu *et al.* 2019). They are found to appear in the genome in three major forms: antisense lncRNAs, intronic lncRNAs, and intergenic lncRNAs (also termed large intervening non-coding RNAs or lincRNAs). LncRNAs regulate gene activity through interactions with chromatin, especially to suppress gene expression (Zhang *et al.* 2019a). Aberrant expression of some lncRNAs has been shown to be closely correlated with cancer prognosis.

Urothelial carcinoma associated 1 (UCA1) is an oncogenic lncRNA which was initially discovered in bladder cancer and a member of the human endogenous retrovirus H family (Fang *et al.* 2017). UCA1 promotes tumorigenesis mainly via binding to tumor-suppressive microRNAs (miRNAs), activating several pivotal signaling pathways and alteration of epigenetic and transcriptional regulation (Yang *et al.* 2016; Yao *et al.* 2019). It contributes to cancer development via regulating cell proliferation, apoptosis, migration, and invasion in diverse tumors, such as breast cancer, colorectal cancer, tongue squamous cell carcinoma (TSCC) and bladder cancer. This suggests that UCA1 may be utilized as a biomarker for the diagnosis and prognosis of these cancers (Yang *et al.* 2016; Fang *et al.* 2017). Hox transcript antisense intergenic RNA (HOTAIR) is found to be located within the homeobox C (HOXC) gene cluster on chromosome 12 (Yu and Li 2015). This when interacts with polycomb repressive complex (PRC2), enhances the trimethylation of lysine 27 of histone H3 (H3K27), which leads to the repression of multiple genes (Yu and Li 2015). Another lncRNA called Taurine Up-regulated Gene 1 (TUG1) was originally identified as a transcript that is up-regulated by taurine (Han *et al.* 2013). It is 7.1-kb in size and was initially detected in a genomic screen for genes up-regulated in response to taurine treatment of developing mouse retinal cells (Zhang *et al.* 2014).

Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) was originally described as a prognostic marker of lung cancer metastasis, emerging evidence has linked this lncRNA to other cancers, such as breast cancer, prostate cancer, pancreatic cancer, glioma, and leukemia (Sun and Ma 2019). H19 lncRNA is associated with diverse cancer types and is located on human chromosome 11p15.5 having a length of 2.3 kb, and consists of 5 exons and 4 introns (Brannan *et al.* 1990; Ghafouri-Fard *et al.* 2020). It was initially thought to possess tumor-suppressive properties because of its ability to inhibit tumorigenesis, but now recent studies have shown that H19 possesses both tumor promoter and suppressive functions (Lv *et al.* 2014; Zhu *et al.* 2014; Raveh *et al.* 2015; Zhang *et al.* 2017; Lan *et al.* 2018; Yan *et al.* 2019; Ghafouri-Fard *et al.* 2020). However, the role of H19 in OSCC has not been studied yet.

LncRNAs could serve as a biomarker for diagnosis and prognosis as well as a molecular target for therapy in OSCC. Therefore, in the current study, we have determined the gene expression profile of 5-lncRNAs namely HOTAIR, UCA1, TUG1, MALAT1 and H19 in the tumor and adjacent normal tissues from OSCC patients. Expression of *UCA1, TUG1* and *HOTAIR* genes was marginally up-regulated in 54.83%, 53.12% and 51.61% OSCC tumors, respectively. Further, MALAT1 was found to be down-regulated in 62.5% of OSCC tumors as compared to adjacent normal tissue. Importantly, the expression of *H19* was down-regulated in tumor tissues of 81.25% of OSCC patients. Except for MALAT1, our experimental data showed concordance with the TCGA analysis. Up-regulation of UCA1 in HNSCC was accompanied by demethylation of its promoter. TCGA analysis also revealed gene amplification of MALAT1 in 4% of HNSCC patients.

## 2. Materials and methods

### 2.1 *Subjects and tissue collection*

Patients (N = 32) diagnosed with cancer of the lip, tongue and oral cavity (ICD-10: C00-08) over the age of 18 years were enrolled in a prospective study at Jawaharlal Nehru Cancer Hospital and Research Centre, Bhopal. Patients who had not received radiotherapy and/or chemotherapy were enrolled in the study. Written informed consent was taken from each participating subject. Tumor and adjacent normal tissues (at least 1 cm away from tumor margin) were collected in microcentrifuge tubes containing 500µl RNA later (ThermoFisher Scientific, Waltham, MA) immediately after surgical resection. The above protocol was approved by the Institutional Human Ethics Committee of AIIMS Bhopal.

### 2.2 *RNA extraction, reverse transcription and qRT-PCR*

Total RNA extraction, cDNA synthesis and qRT-PCR were performed as described previously (Vishwakarma *et al.* 2017). The following set of primers were used: *MALAT1* Forward 5′-GACGGAGGTTGAGATGAAGC-3′ Reverse 5′-ATTCGGGGCTCTGTAGTCCT-3′; *HOTAIR* Forward 5′- AAATATGGCGGCGTCTACACGGA-3′ Reverse 5′- TCCAGAACCCTCTGACATTTGCCT-3′; *UCA1* Forward 5′- GACCCTACCCGGTCATTTATAG-3′ Reverse 5′- CTGATGGGCATGGCTTTATTC-3′; *TUG1* Forward 5′- CCTCCACTCAGCACAGTTCA-3′, Reverse 5′- TCCTGGGGTCAAAGACATGC-3′; *H19* Forward 5′ ACTCAGGAATCGGCTCTGGAA-3′, Reverse 5′- CTGCTGTTCCGATGGTGTCTT-3′; *GAPDH* Forward 5′-AATCCCATCACCATCTTC-CAG- 3′; Reverse 5′-AAATGAGCCCCAGCCTTC-3′. Relative RNA expression of genes was determined by ΔΔCq method as described previously (Vishwakarma *et al.* 2017). GAPDH was used as a reference gene.

### 2.3 *Analysis of The Cancer Genome Atlas (TCGA)*

UALCAN (*http://ualcan.path.uab.edu/*) is a user-friendly, comprehensive and interactive web resource for analyzing TCGA OMICS data (Chandrashekar *et al.* 2017). It allows to prepare plots and graphs depicting gene expression, DNA methylation, correlation among genes and survival analysis of cancer patients (Chandrashekar *et al.* 2017). UALCAN has clinicopathological and OMICS data from patients with 31 different cancer types/subtypes. It has data from 50 mucosae from healthy subjects and primary tumors from HNSCC patients (N = 528). The UALCAN database was used with the set conditions for filtering and data mining to compare gene expression and methylation between normal mucosa and primary tumors from HNSCC patients as well as pan-cancer. LncRNA expression has been shown as the number of transcripts per million total transcripts. Beta values indicate the level of DNA methylation ranging from 0 (unmethylated) to 1 (fully methylated). Beta values in the range of 0.7 to 0.5 indicate hypermethylation, whereas beta values between 0.3 to 0.25 indicates hypomethylation of DNA.

To correlate the lncRNA expression with clinicopathological features of HNSCC patients, MEXPRESS, a web-based bioinformatics tool (*http://mexpress.be*) (Koch *et al.* 2015) was used. MEXPRESS has gene expression data along with clinicopathological features for 612 HNSCC patients.

cBioPortal (*http://cbioportal.org/*) is a publicly accessible cancer genomics portal to explore, visualize and analyze multidimensional cancer genomics data (Gao *et al.* 2013). The portal contains genomic sequencing datasets from 1478 HNSCC patients from six previously published studies (Gao *et al.* 2013). Genomic alterations in the UCA1, TUG1, HOTAIR, MALAT1 and H19 lncRNA genes were analyzed by cBioportal.

The "Kaplan-Meier plotter" (KM plotter) is a widely used database for the real-time meta-analysis of published cancer microarray datasets to identify biomarkers related to survival (Gyorffy *et al.* 2013). KM plotter contains the cancer datasets from The Cancer Genome Atlas (TCGA), the Gene Expression Omnibus (GEO) and Cancer Biomedical Informatics Grid (caBIG). The association between gene-specific lncRNA expression and overall survival (OS) was analyzed by the KM plotter. Currently, gene expression and survival data from 1927 patients for a follow-up period of 20 years is available from TCGA lung cancer datasets (Gyorffy *et al.* 2013).

### 2.4 *Statistical analysis*

Statistical analysis was performed as described previously (Vishwakarma *et al.* 2017) using SPSS 21.0 software (IBM, Armonk, New York). Demographic and patient clinical history including age, gender, tobacco chewing, type of tobacco/betel nut consumed, history of smoking, alcohol consumption, site of tumor, tumor volume, TNM staging, lymph node ratio (LNR), lympho-vascular invasion (LVI) and perinodal extension (PNE) were recorded into the SPSS database. Mean fold change in individual lncRNA expression was determined and the difference in the mean fold change of gene expression among the groups and the association of fold change in lncRNA expression with clinicopathological attributes was analyzed as described previously (Vishwakarma *et al.* 2017). Further, to determine the correlation among different lncRNAs expression in OSCC, gene expression data were categorized as up-regulation or down-regulation for each gene, then analyzed by chi-square test.

## 3. Results

### 3.1 *Clinicopathological features of subjects*

A total of 32 OSCC patients who were undergoing surgical resection of tumor were enrolled in the study.

All the tumors were histopathologically proven cases of OSCC. The median age of the subjects was 45 years (23–72 years), and a majority of patients (81.25%) were male. Two-third of the subjects had a history of tobacco/betel nut chewing (78.12%). Approximately, one-fourth (25%) of the subjects had a history of smoking or alcohol consumption. Buccal mucosa was the most common (59.4%) site for oral cancer in the subjects, followed by tongue (18.7%) (table 1). Clinicopathological features of OSCC patients are shown in table 2.

### 3.2 *Expression of UCA1 and TUG1 lncRNA was up-regulated in OSCC patients*

In our study, higher levels of RNA expression for UCA1 and TUG1 was noted in the tumors of OSCC patients compared to normal adjacent tissue (figure 1A–D). Expression of UCA1 and TUG1 was also found to be up-regulated in the tumors compared to adjacent normal tissues from 54.83% and 53.12% OSCC patients, respectively (figure 1). Importantly, more than 30% increase in the mean expression of UCA1 was observed in the tumors, compared to normal tissues (figure 1B). No consistent pattern was observed for HOTAIR expression. Half of the OSCC patients enrolled in the study showed down-regulation, whereas in almost similar numbers up-regulation of HOTAIR lncRNA was noted (figure 1E–F).

**Table 1.** Demographic details of OSCC patients (N = 32) m

| Parameters | Number | Frequency |
|---|---|---|
| *Median Age (yr)* | 45 (23–72) | |
| *Age group* | | |
| <50 | 21 | 65.60 |
| >50 | 11 | 34.40 |
| *Gender* | | |
| Male | 26 | 81.25 |
| Female | 06 | 18.75 |
| *Smoking status* | | |
| Yes | 09 | 28.12 |
| No | 23 | 71.87 |
| *Drinking status* | | |
| Yes | 08 | 25.00 |
| No | 24 | 75.00 |
| *Tumor chewing status* | | |
| Yes | 25 | 78.12 |
| No | 07 | 21.87 |

### 3.3 *MALAT1 and H19 are down-regulated in OSCC patients*

In our study, decreased RNA expression for MALAT1 and H19 was noted in the tumors of OSCC patients compared to normal adjacent tissue. The expression of *MALAT1* and *H19* was down-regulated in tumor tissues of 62.5 % and 81.25 % respectively (table 3) of OSCC patients (figure 2). More than 30% decrease ($P < 0.05$) in the mean expression of *MALAT1* was observed in the tumors, compared to normal tissues (figure 2B). Notably, OSCC tumors showed 80% decrease ($p = 5.3 \times 10^{-5}$) in the mean expression of H19 lncRNA, compared to adjacent normal tissues (figure 2D).

### 3.4 *TCGA Analysis*

The expression of various lncRNAs in the HNSCC patients was analyzed by the publicly available

**Table 2.** Clinico-pathological features of OSCC patients (N = 32) mRNA expression profile of lncRNAs in OSCC tumors

| Parameters | Number | Frequency |
|---|---|---|
| *Tumor Site* | | |
| Buccal Mucosa | 19 | 59.4 |
| Tongue | 6 | 18.1 |
| Gingivo-buccal sulcus | 3 | 9.4 |
| Lip | 2 | 6.3 |
| Alveoli | 1 | 3.1 |
| Hard Palate | 1 | 3.1 |
| *TNM Over-all staging* | | |
| Stage-I | 2 | 6.25 |
| Stage-II | 10 | 31.25 |
| Stage-III | 0 | 0 |
| Stage-IV | 20 | 62.5 |
| *TNM Staging (Combined)* | | |
| Stage-I+II | 12 | 37.5 |
| Stage-III+IV | 20 | 62.5 |
| *Tumor volume* | | |
| <4cm3 | 13 | 41.93 |
| >4cm3 | 18 | 58.06 |
| *Lymph node ratio* | | |
| <0.1 | 21 | 65.62 |
| >0.1 | 11 | 34.37 |
| *Lympho-vascular invasion* | | |
| Yes | 2 | 6.66 |
| No | 30 | 93.75 |
| *Perinodal extension* | | |
| Yes | 18 | 56.25 |
| No | 14 | 43.75 |

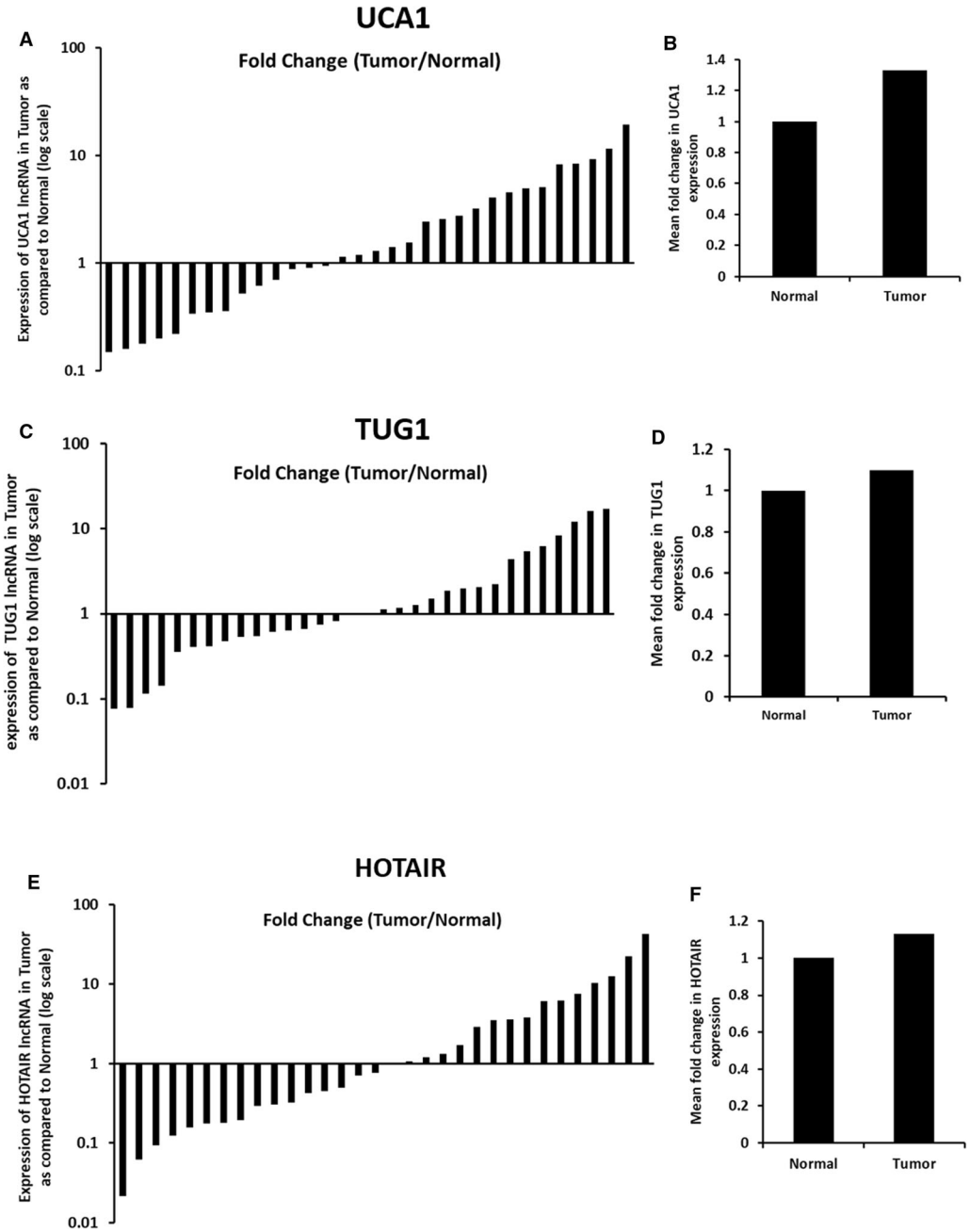


**Figure 1.** **(A–F)** Fold change in the expression of UCA1, TUG1 and HOTAIR lncRNA in tumors compared to adjacent normal tissues. Relative change in the expression of UCA1 **(A)**, TUG1 **(C)** and HOTAIR **(E)** lncRNA in tumors, compared to normal tissues was determined by qRT-PCR. Fold change in lncRNA expression is presented in log scale. Mean fold change in expression of UCA1 **(B)**, TUG1 **(D)** and HOTAIR **(F)** lncRNA was calculated by subtracting mean ΔCq of normal from mean ΔCq of tumours from all OSCC patients. Difference in means of ΔCq of tumor vs normal was calculated by independent sample t-test.

UALCAN database, which contains information regarding clinicopathological attributes of patients along with gene expression and DNA methylation in the normal and primary tumors from a variety of cancer types (Chandrashekar *et al.* 2017). As shown in figure 3A, a significant increase in the UCA1 lncRNA expression was found in the primary tumors from HNSCC patients compared to normal tissue. The increase in UCA1 expression in tumor tissue was accompanied by the hypomethylation of its promoter (figure 3B). We also analyzed the expression of UCA1 lncRNA expression across various cancer types. The expression of UCA1 was higher in most of the tumors compared to corresponding normal tissues. Among

**Table 3.** RNA expression profile of lnc RNA's in OSCC tumors1 mRNA expression profile of lnc RNA's in OSCC tumors

| S.No. | Name of lnc RNA | Change in gene expression in tumors | Proportion of patients |
|---|---|---|---|
| 1 | MALAT1 | Down-regulation | 62.5% |
| 2 | HOTAIR | Up-regulation | 51.61% |
| 3 | UCA1 | Up-regulation | 54.83% |
| 4 | TUG1 | Down-regulation | 53.12% |
| 5 | H19 | Down-regulation | 81.25% |

Fold change in lncRNA expression in tumors compared with adjacent normal tissue of the same patient was calculated by the delta-delta Cq method

various cancer types, urothelial carcinomas, cholangiocarcinoma, colon adenocarcinoma, rectum adenocarcinomas exhibited the highest expression of UCA1 (figure 3C). Similarly, the expression of TUG1 lncRNA was found to be increased by almost 2-folds in the HNSCC tumors compared to normal tissues. However, no apparent difference was noted in the methylation status of TUG1 promoter in the tumors, as compared to normal tissue (figure 4A–B). Most of the cancer types showed the increased expression in tumors except renal clear cell carcinoma, thymomas, thyroid carcinoma and uveal melanomas (figure 4C). The TCGA analysis showed increased expression of HOTAIR in the HNSCC tumors compared to normal tissue and it was associated with its promoter

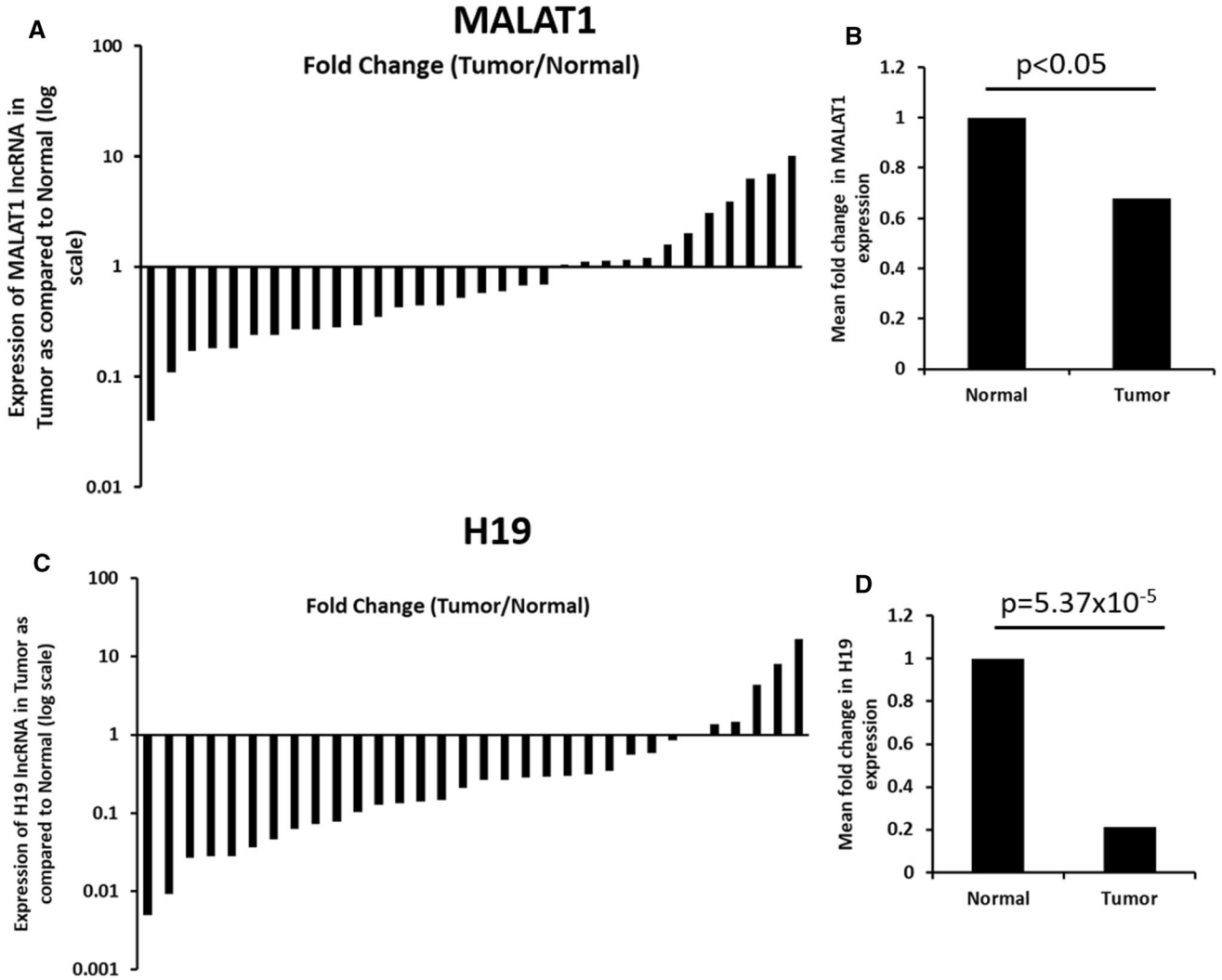


**Figure 2.** **(A–D)** Fold change in the expression of MALAT1 and H19 lncRNA in tumors compared to adjacent normal tissues. Relative change in the expression of MALAT1 **(A)** and H19 **(C)** in tumors, compared to normal tissues was determined by qRT-PCR. Fold change in the lncRNA expression is presented in log scale. Mean fold change in expression of MALAT1 **(B)** and H19 **(D)** lncRNA was calculated by subtracting mean ΔCq of normal from mean ΔCq of tumors from all OSCC patients. Difference in means of ΔCq of tumor vs normal was calculated by independent sample t-test.

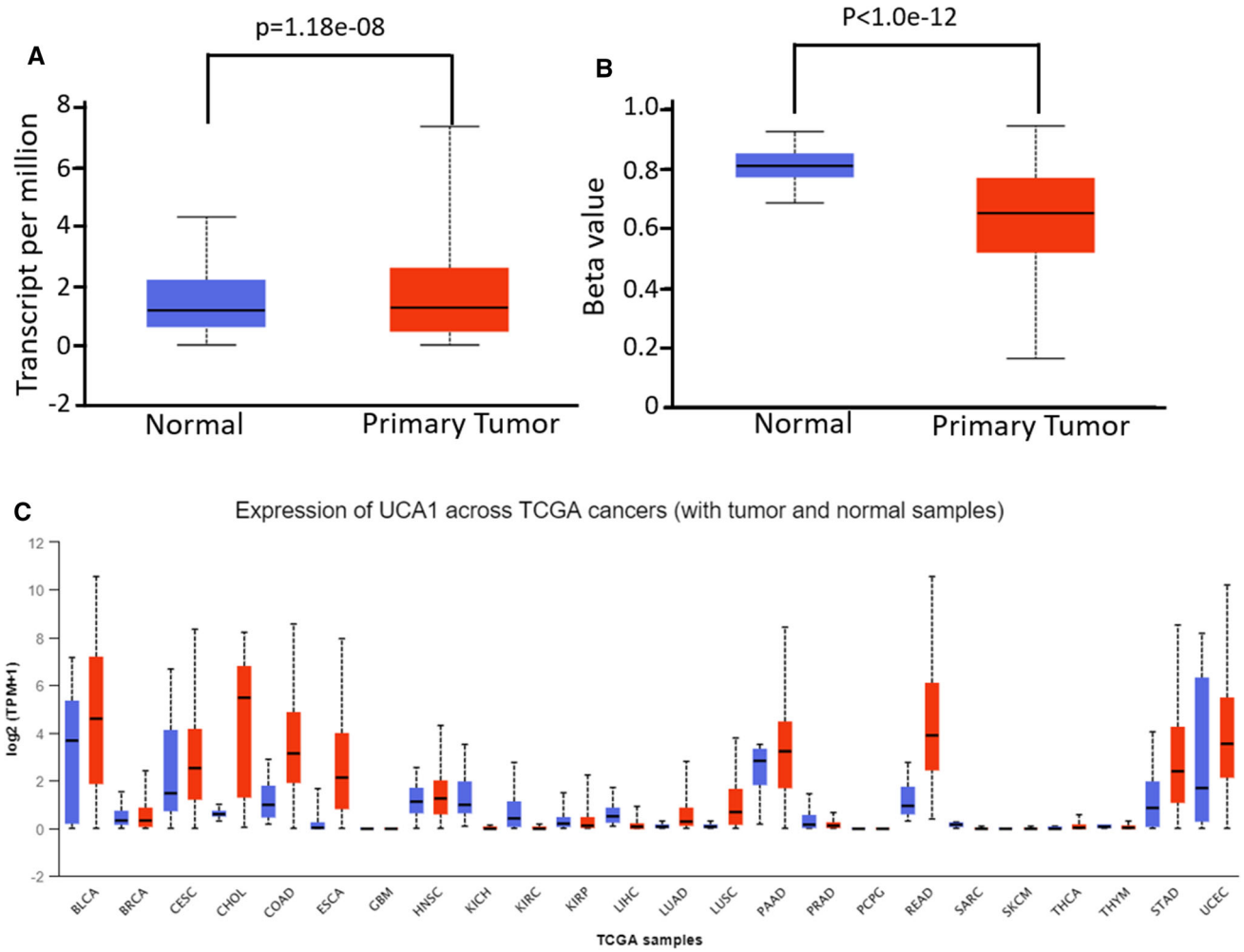


**Figure 3.** TCGA analysis of UCA1 lncRNA. (**A–B**) UCA1 lncRNA expression and methylation was analyzed in the normal oral mucosa (N = 44) and primary tumors from HNSCC (N = 520) patients from the publicly available UALCAN database. (**C**) UCA1 lncRNA expression was analyzed in the pan-cancers (normal tissues and primary tumors) from the publicly available UALCAN database.

methylation (figure 5A–B). Among various cancer types, esophageal carcinoma, glioblastoma multiforme, sarcoma and breast cancer showed the highest expression (figure 5C). The TCGA analysis showed the increased expression of MALAT1 in the HNSCC tumors compared to normal tissues (figure 6A). As compared to normal tissue, no difference was noted in the promoter methylation of MALAT1 lncRNA in the tumors (figure 6B). Unlike UCA1, TUG1 and HOTAIR; MALAT1 was differentially expressed in the different cancer types (figure 6C). Among various tissues, the esophagus showed the highest expression of MALAT1 (figure 6C). Few cancer types including breast cancer, esophageal carcinoma, sarcoma, thymoma and uveal melanoma showed the decreased expression of MALAT1 (figure 6C). Consistent with our experimental data, expression of H19 lncRNA was decreased (75%) in the HNSCC tumors compared to normal tissue (figure 7A); however, it does not correlate with the methylation status of its promoter (figure 7B). Similar to MALAT1, H19 lncRNA is also differentially expressed in various cancer types (table 3). Cervical cancer, cholangiocarcinoma, HNSCC, kidney chromophobe, renal clear cell carcinoma, hepatocellular carcinoma, pancreatic, prostate, pheochromocytoma, thyroid and uveal melanomas exhibited the decreased expression of H19 lncRNA (figure 7C).

Further, to analyze the genomic alterations in the lncRNA, the TCGA data were analyzed by cBioPortal (Gao *et al.* 2013). As shown in supplementary figure 1, genomic alterations are uncommon in UCA1, TUG1, HOTAIR and H19, but gene amplification in MALAT1 was noted in 4% HNSCC primary tumors

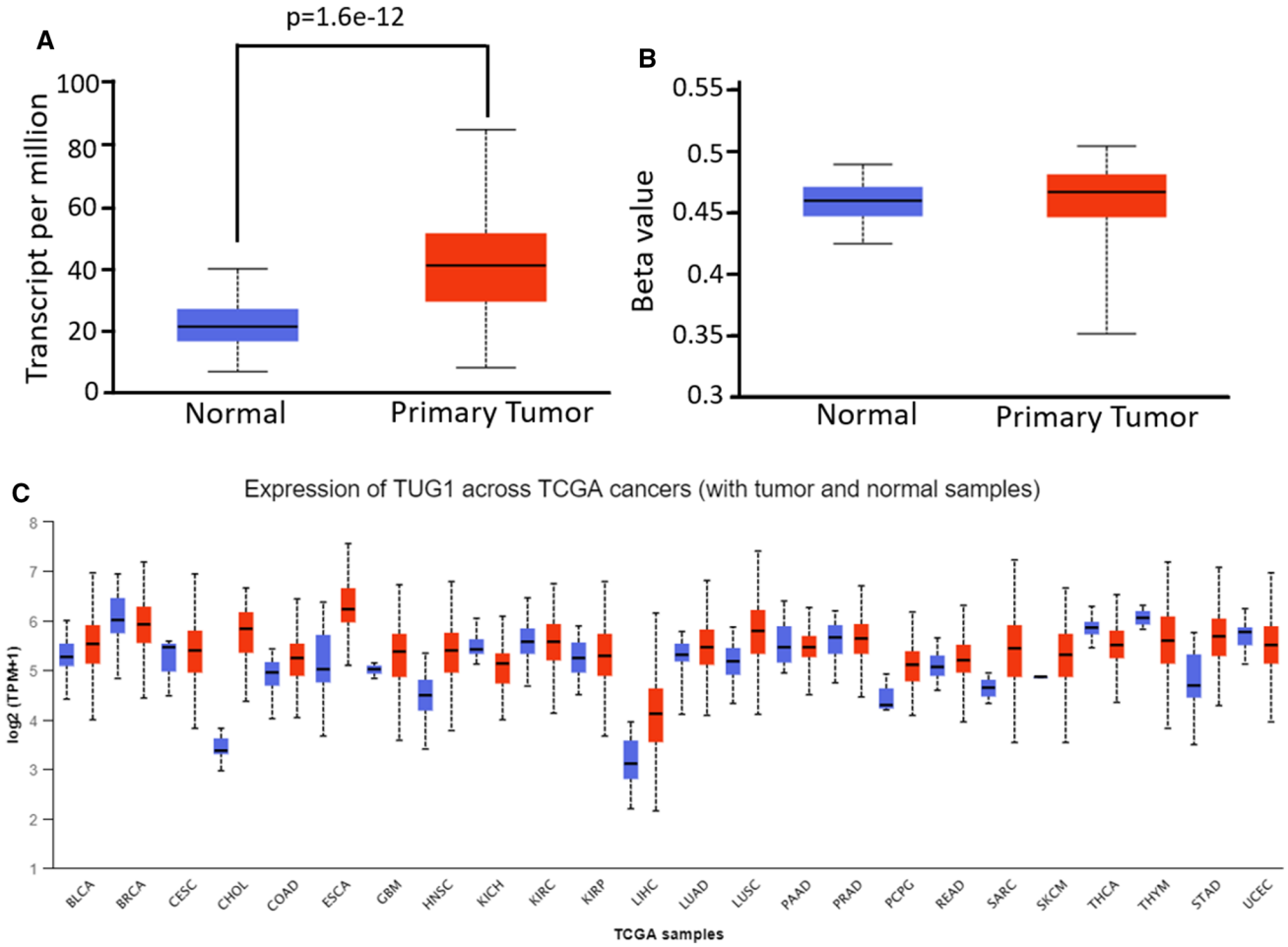


**Figure 4.** TCGA analysis of TUG1 lncRNA. (**A–B**) TUG1 lncRNA expression and methylation was analyzed in the normal oral mucosa (N = 44) and primary tumors from HNSCC (N = 520) patients from the publicly available UALCAN database. (**C**) TUG1 lncRNA expression was analyzed in the pan-cancers (normal tissues and primary tumors) from the publicly available UALCAN database.

(supplementary figure 1D). We also analyze the TCGA data for the relative abundance of lncRNA expression in the normal and HNSCC tumors. As shown in supplementary table 2, H19 was the most abundantly expressed lncRNA in oral mucosa followed by TUG1.

### 3.5 *Correlation of lncRNA expression profile in OSCC tumors*

To determine the correlation of lncRNA expression among each other in OSCC tumors, gene expression data were divided into two categories (down-regulation or up-regulation in tumors in comparison to normal tissues) for each lncRNA. Correlation between two genes was determined by the Pearson Chi-square test. The most significant negative correlation was observed between MALAT1 and HOTAIR (Pearson Chi-Square 4.8; R = 0.387; P = 0.028). However, MALAT1 showed positive correlation with TUG1 (Pearson Chi-Square 7.0; R = – 0.469; P = 0.008). No significant association was observed among other lncRNAs studied in this research paper.

### 3.6 *Correlation of lncRNA expression with clinicopathological features*

HOTAIR is a prognostic marker, and its overexpression correlates with poor survival of breast cancer patients. Therefore, to determine the clinical significance of lncRNAs in OSCC, correlations between fold-change in lncRNA expression in tumors and clinicopathological attributes (sex, age, tobacco chewing, smoking status, alcohol consumption, tumor site, TNM staging, tumor volume, LNR, PNE and LVI) were analyzed by Spearman's rank test. In the present study, expression of HOTAIR positively correlated with tumor volume

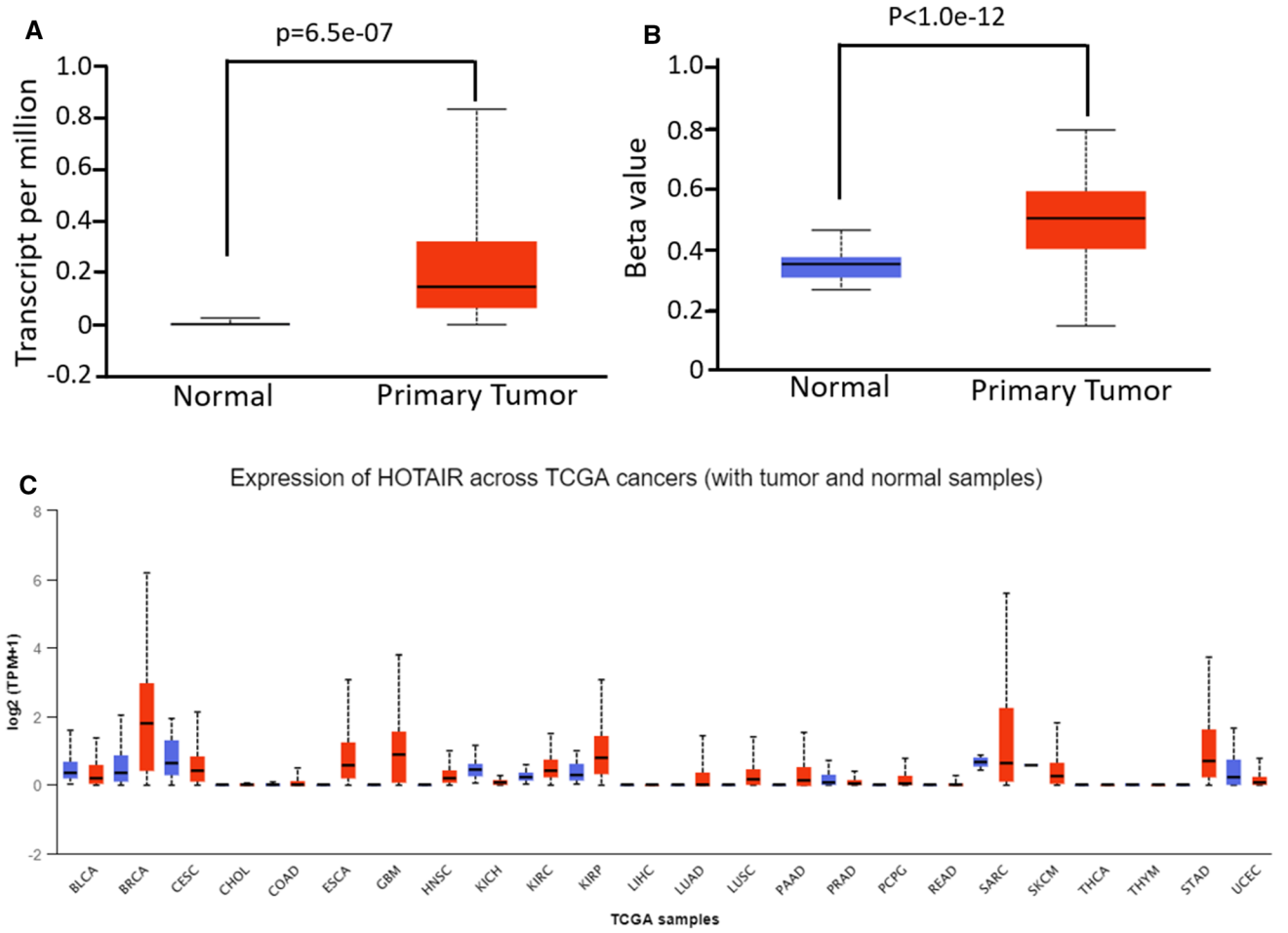


**Figure 5.** TCGA analysis of HOTAIR lncRNA. (**A–B**) HOTAIR lncRNA expression and methylation was analyzed in the normal oral mucosa (N = 44) and primary tumors from HNSCC (N = 520) patients from the publicly available UALCAN database. (**C**) HOTAIR lncRNA expression was analyzed in the pan-cancers (normal tissues and primary tumors) from the publicly available UALCAN database.

(rho (ρ) = 0.366, p = 0.043). Patients with tumor volume of >4 $cm^3$ had higher expression of HOTAIR. Interestingly, fold change in the expression of MALAT1 (ρ = − 0.496, p = 0.004). and H19 (ρ = − 0.439, p = 0.012) was negatively correlated with smoking status of OSCC patients. In our study, we have seen down-regulation of both the lncRNAs in the tumors, compared to normal tissue. Smokers are more likely to have low expression of both the lncRNAs.

Using a publicly available database MEXPRESS, we also analyzed the correlation between lncRNA expression and clinicopathological features. RNA expression of UCA1 was significantly associated with neoplasm histological grade ($p = 2.69 \times 10^{-5}$), clinical metastasis (p = 0.022), HPV-16 status (p = 0.016), pathological node (p = 0.041), gender (p = 0.003) and also with alcohol history (figure 8). The expression of H19 RNA was found to be highly significant with HPV-status (p = 0.002) and with the clinical stage (p = 0.007) (figure 8). The RNA expression of HOTAIR was significantly negatively correlated with the number of positive lymph nodes found by IHC (r = − 0.191) whereas positively correlated with the age at initial pathological diagnosis (r = 0.117) and also found to be significant with tumor stage (p = 0.039) (figure 8). The expression of MALAT1 and TUG1 was found to be significantly correlated with copy number r = 0.163 and r = 0.455 respectively (figure 8).

Survival outcome for HNSCC either is not available in a web-based bioinformatic public database such as KM Plotter, PrognoScan or sample size is very low, thus, to determine the prognostic role of lncRNA, Kaplan-Meier curves were plotted for the lncRNA expression with overall survival of non-small cell lung cancer (NSCLC) patients. As shown in supplementary figure 2, MALAT1 and TUG1 were significantly associated with the overall survival of NSCLC patients.

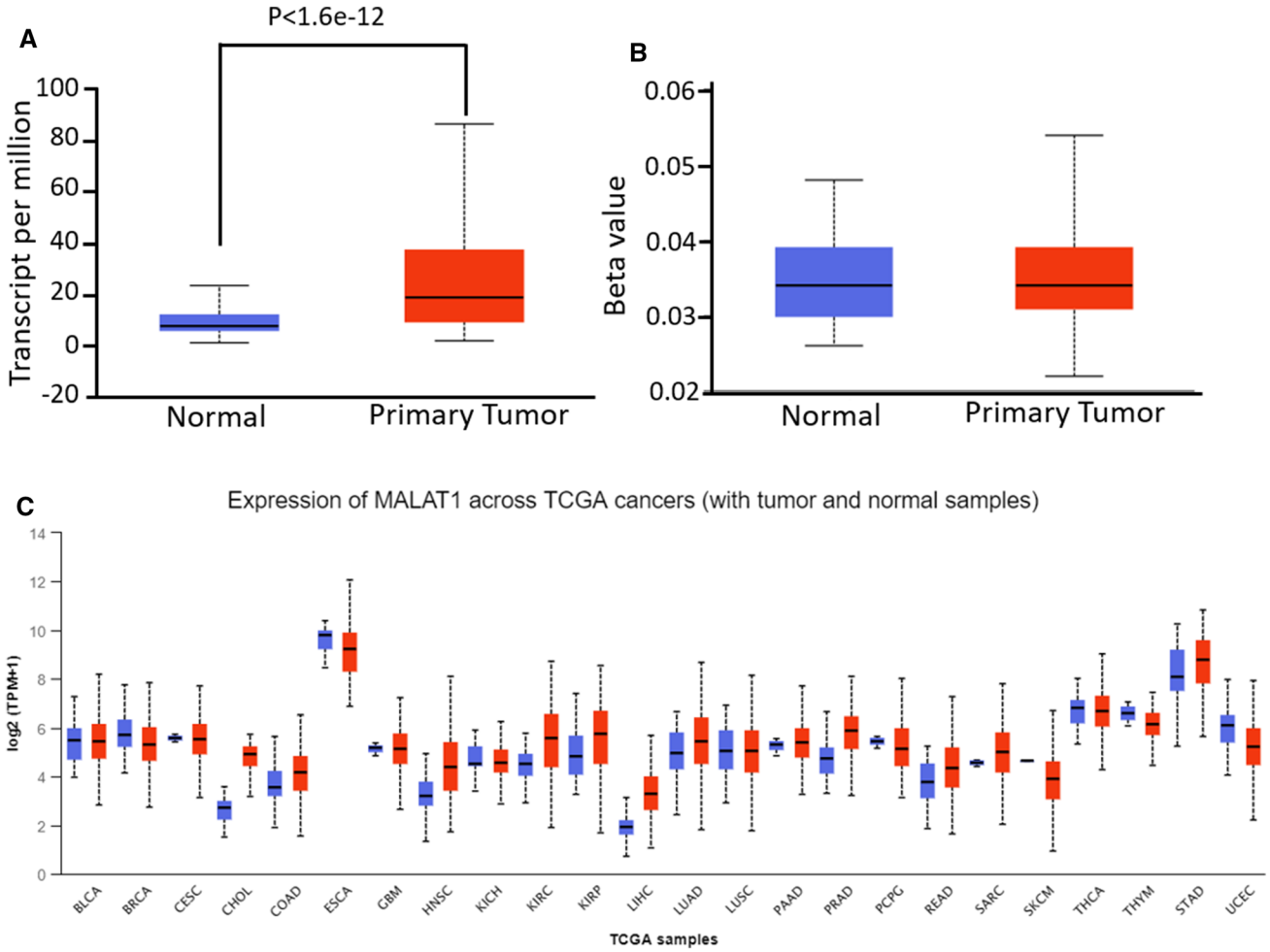


**Figure 6.** TCGA analysis of MALAT1 lncRNA. (**A–B**) MALAT1 lncRNA expression and methylation was analyzed in the normal oral mucosa (N = 44) and primary tumors from HNSCC (N = 520) patients from the publicly available UALCAN database. (**C**) MALAT1 lncRNA expression was analyzed in the pan-cancers (normal tissues and primary tumors) from the publicly available UALCAN database.

## 4. Discussion

On the Indian subcontinent, cancers of the lip and oral cavity, histologically called oral squamous cell carcinoma (OSCC) are the leading cause of cancer-related death among men in India. LncRNA may serve as a biomarker for diagnosis and prognosis for OSCC. In the present study, we have found that UCA1 is up-regulated in 60% of OSCC tumors as compared to normal adjacent tissue. The up-regulation of UCA1 lncRNA expression in OSCC tumors observed in our study was in accordance with the HNSCC TCGA dataset (figure 3). Consistent with our findings, others have shown that UCA1 is up-regulated in TSCC tissues and was closely correlated with metastasis and TNM staging (Fang *et al.* 2014; Yang *et al.* 2016; Zhang *et al.* 2019b). Furthermore, the up-regulation of UCA1 in OSCC tumors might be due to the demethylation of its promoter, as it seems to be hypermethylated in the normal mucosa. UCA1 modulates TGFβ1-induced epithelial-mesenchymal transition and invasion of tongue cancer cells through JAG1/Notch signaling (Zhang *et al.* 2019b). The up-regulated UCA1 promotes cancer progression by regulating the mTOR or Wnt signaling pathway and the higher UCA1 expression was associated with poor prognosis of TSCC patients (Xu *et al.* 2017; Zhang *et al.* 2019b). TUG1 is known to be up-regulated in various types of cancer, and in the present study, up-regulation of TUG1 seen in OSCC tumors is also in accordance with TCGA analysis (figure 4). TUG1 lncRNA has been shown to promote OSCC progression through up-regulating formin-like 2 (FMNL2) by sponging miR-219 (Yan *et al.* 2017).

HOTAIR was known to be one of the first lncRNAs reported to be involved in the development of cancer

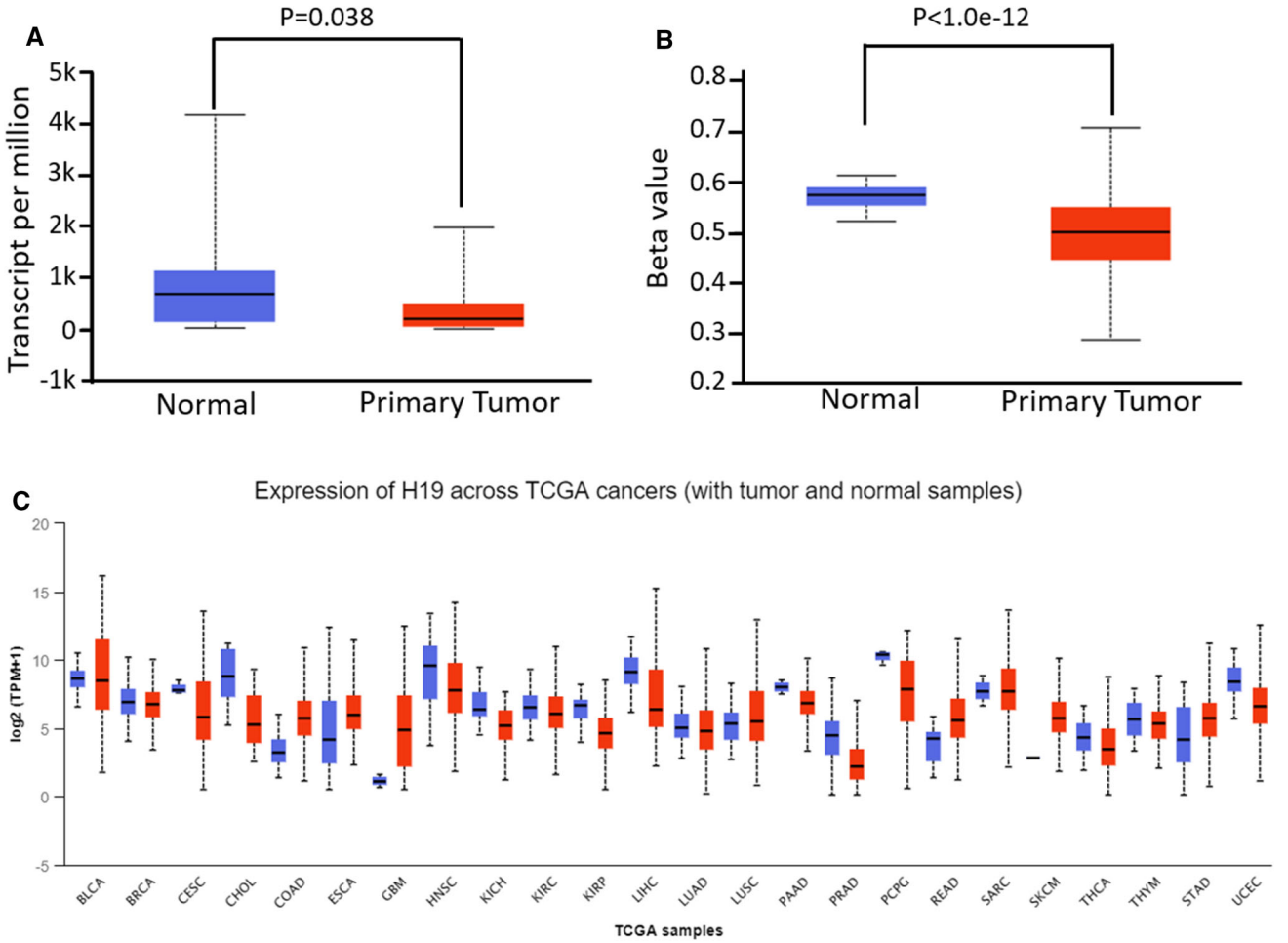


**Figure 7.** TCGA analysis of H19 lncRNA. (**A–B**) H19 lncRNA expression and methylation was analyzed in the normal oral mucosa (N = 50) and primary tumors from HNSCC (N = 528) patients from the publicly available UALCAN database. (**C**)H19 lncRNA expression was analyzed in the pan-cancers (normal tissues and primary tumors) from the publicly available UALCAN database.

(Yu and Li 2015). It has been demonstrated in *in-vitro* studies that HOTAIR can promote cancer cell proliferation, invasion and metastasis, while it is known to inhibit apoptosis. While in *in-vivo* studies also, especially in murine xenograft models, it has been found that *HOTAIR*-knockout can reduce tumor growth. Thus, HOTAIR has been suggested as a potential oncogene (Yu and Li 2015). It is also an independent prognostic marker of metastasis in estrogen receptor-positive primary breast cancer. HOTAIR is overexpressed in various types of cancers (figure 5), and its overexpression is associated with metastasis and poor survival rates and correlates positively with DNA methylation in primary breast cancer (Gupta *et al.* 2010). Here, in our study, no consistent pattern was observed for HOTAIR expression in OSCC tumors as compared to adjacent normal tissue. More than 40% of patients had a 10-fold increase in the HOTAIR expression. The TCGA dataset analysis showed the increased expression of HOTAIR in the HNSCC tumors, but the increase in the expression was not correlated with methylation of its promoter. Signal transducer and activator of transcription 3 (STAT3), a transcription factor, which is overexpressed in HNSCC has been shown to induce the expression of HOTAIR through EZH2 phosphorylation (Sun *et al.* 2018). A previous study showed that HOTAIR expression was correlated with tumor size, TNM stage, and prognosis of OSCC patients. A study had shown the presence of HOTAIR in saliva samples from OSCC patients with primary tumors. This suggests that the detection of lncRNAs in saliva can be used for clinical non-invasive and rapid diagnosis of OSCC (Tang *et al.* 2013).

MALAT1 is highly conserved in mammals and is also referred to as noncoding nuclear-enriched abundant transcript 2 (NEAT2). It is involved in

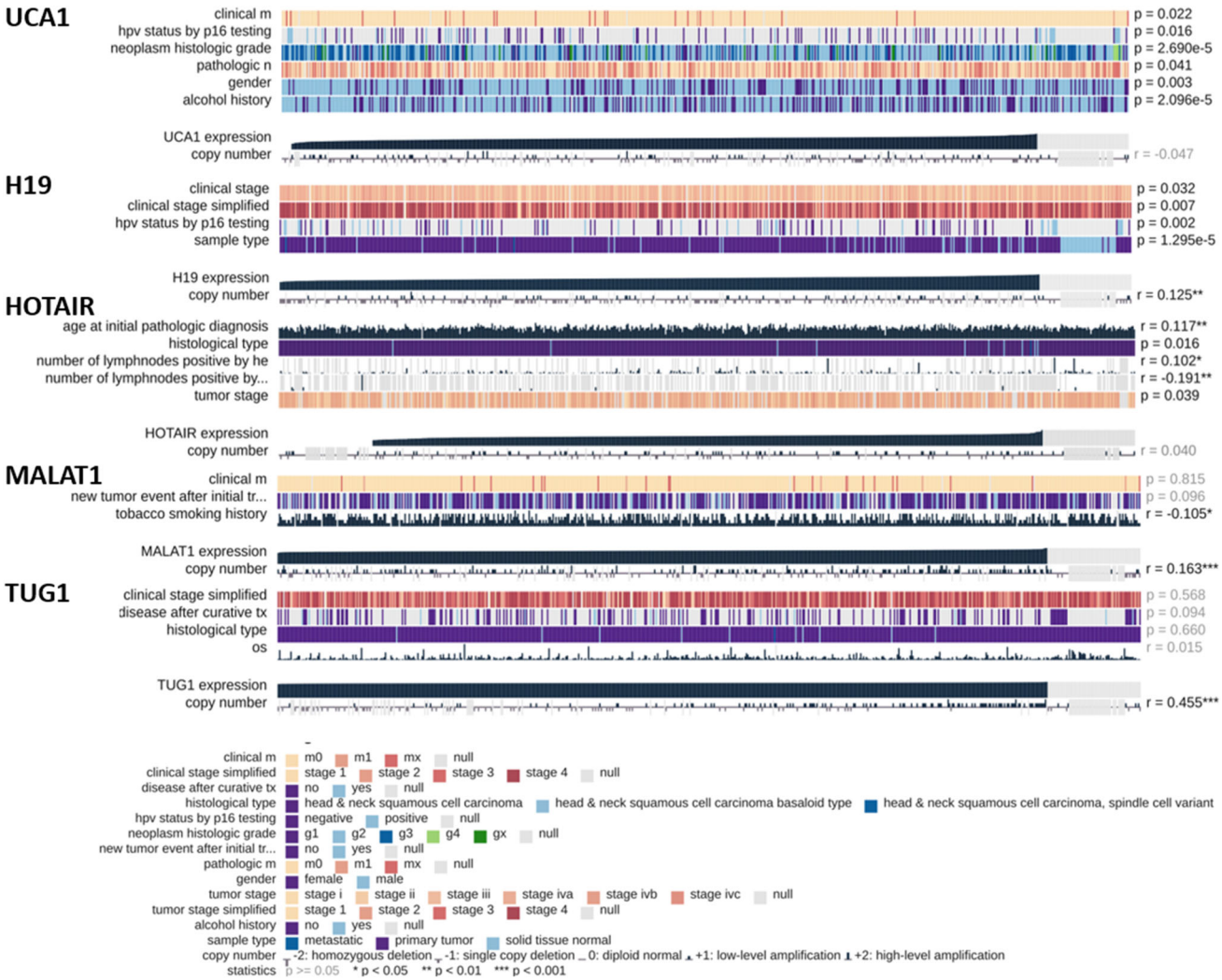


**Figure 8.** Correlation of RNA expression of UCA1, H19, HOTAIR, MALAT1 and TUG1 lncRNA with clinico-pathological features of HNSCC patients (n = 612) (**A–E**) RNA expression with different clinico-pathological features was analyzed in the HNSCC (normal tissues and primary tumors) from the publicly available MEXPRESS database.

transcriptional and post-transcriptional regulation (Hutchinson *et al.* 2007). In our study, we found a significant down-regulation of MALAT1 in OSCC tumors compared to adjacent normal tissue. Consistent with our data clinical correlation data with OSCC patients, TCGA analysis on publicly available data showed that HOTAIR expression was associated with tumor stage whereas MALAT1 was negatively correlated with smoking status of the HNSCC patients. The role of MALAT1 in cancer progression and metastasis is controversial. In the early studies, MALAT1 has been shown to promote cellular proliferation by regulating the activity of the E2F1 transcription factor, and then thereby enhances tumorigenesis (Tripathi *et al.* 2013). Exogenous expression of MALAT1 has also been shown to induce STAT3 expression in OSCC cells (Henson *et al.* 2009). On the other hand, several studies showed that the expression of MALAT1 is down-regulated in glioma (Cao *et al.* 2016), colorectal cancer (Barbagallo *et al.* 2018) and breast cancer (Xu *et al.* 2015; Kim *et al.* 2018; Kwok *et al.* 2018; Sun and Ma 2019). Importantly, inactivation of the *Malat1* gene in a transgenic mouse model of breast cancer, without altering the expression of its adjacent genes, promotes lung metastasis (Kim *et al.* 2018). The TCGA analysis showed increased expression of MALAT1 in the HNSCC tumors and it was also accompanied by gene amplification in 4% of HNSCC patients. Consistently, gene amplification in MALAT1 has been reported in esophageal cancer (Hu *et al.* 2015). Mechanistically, the MALAT1 lncRNA binds and inactivates the pro-metastatic transcription factor TEAD, preventing TEAD from associating with its co-activator YAP and target gene promoters (Kim *et al.* 2018). Furthermore,

MALAT1 levels inversely correlate with breast cancer progression and metastatic ability (Hsu *et al.* 2016; Kim *et al.* 2018). Down-regulation of MALAT1 induces EMT via the PI3K-AKT pathway in breast cancer cells (Xu *et al.* 2015).

H19 lncRNA is an oncofetal, a maternally expressed and paternally imprinted 2.7 kb gene (Raveh *et al.* 2015). Its role in tumor initiation and progression has long been a subject of controversy, it is actively involved in all stages of tumorigenesis and is expressed in almost every human cancer (Raveh *et al.* 2015). However, this is the first report showing the expression of H19 lncRNA in OSCC. Here, we have shown that expression of *H19* was down-regulated in tumor tissues of 84.3% of OSCC patients, which is in accordance with our TCGA dataset analysis. H19 is also found to be down-regulated in papillary thyroid carcinoma (PTC) tissues as well as in PTC cell lines compared to controls. Decreased H19 expression was correlated with lymph node metastasis. Furthermore, H19 overexpression reduced PTC cell proliferation and migration. It also inhibited the expression of tumor necrosis factor receptor 2 (Lan *et al.* 2018). The TCGA analysis across pan-cancer showed that the expression of H19 is down-regulated in several cancer types (figure 7). Down-regulation of H19 in OSCC tumors may not be due to alteration in the methylation status of its promoter, rather it could be due to other epigenetic mechanisms (Nordin *et al.* 2014). Two SNPs, rs217727 and rs2839701 in H19 were found to be associated with the risk of OSCC. Further, rs2839701 C>G inhibited the transcription activity and was correlated with the decreased expression of downstream gene MRPL23-AS1 that was down-regulated in OSCC (Yuan *et al.* 2018). However, the mechanism by which the down-regulation of H19 promotes oral carcinogenesis is not known.

The Indian subcontinent represents one of the major contributors to oral cancer cases; however, only a few studies have compared the expression of lncRNAs with oral cancer in the Indian subcontinent. In a study, Linc-RoR has been shown to be overexpressed in undifferentiated oral tumors and showed a strong association with tumor recurrence and poor therapeutic response (Arunkumar *et al.* 2017). The same group has also shown that OIP5-AS1 lncRNA is overexpressed in oral tumors and in tumors of epithelial origin from the TCGA database (Arunkumar *et al.* 2018). As only limited data is available in the International Consortium for lncRNA expression in OSCC from India patients, further studies are warranted to ascertain their prognostic role in OSCC.

## Acknowledgements

This work was supported by Science and Engineering Research Board (SERB), Govt. of India's Scheme for Young Scientist research grant (SB/YS/LS-184/2013) and extramural research grant (EMR/2016/005009) by SERB to AK and ICMR senior research fellowship (ICMR/2019-6836) to SV. RP is recipient of DBT fellowship DBT/2018/AIIMS-B/1153.

## References

Arunkumar G, Deva Magendhra Rao AK, Manikandan M, Arun K, Vinothkumar V, Revathidevi S, Rajkumar KS, Rajaraman R and Munirajan AK 2017 Expression profiling of long non-coding RNA identifies linc-RoR as a prognostic biomarker in oral cancer. *Tumor Biol.* **39** 1010428317698366

Arunkumar G, Anand S, Raksha P, Dhamodharan S, Prasanna Srinivasa Rao H, Subbiah S, Murugan AK and Munirajan AK 2018 LncRNA OIP5-AS1 is overexpressed in undifferentiated oral tumors and integrated analysis identifies as a downstream effector of stemness-associated transcription factors. *Sci. Rep.* **8** 7018.

Barbagallo C, Brex D, Caponnetto A, Cirnigliaro M, Scalia M, Magnano A, Caltabiano R, Barbagallo D, Biondi A, Cappellani A, Basile F DPC and Purrello M RM 2018 LncRNA UCA1, Up-regulated in CRC Biopsies and Down-regulated in Serum Exosomes, Controls mRNA Expression by RNA-RNA Interactions. *Mol. Ther. Nucleic Acids* **12** 229–241

Brannan C, Dees E, Ingram R and Tilghman S 1990 The product of the H19 gene may function as an RNA. *Mol. Cell Biol.* **10** 28–36

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A 2018 Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J. Clin.* **68** 394–424

Cao S, Wang Y, Li J, Lv M, Niu H and Tian Y 2016 Tumor-suppressive function of long noncoding RNA MALAT1 in glioma cells by suppressing miR-155 expression and activating FBXW7 function. *Am. J. Cancer Res.* **6** 2561–2574

Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S 2017 UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* **19** 649–658

Fang Z, Wu L, Wang L, Yang Y, Meng Y and Yang H 2014 Increased expression of the long non-coding RNA UCA1 in tongue squamous cell carcinomas: A possible correlation with cancer metastasis. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **117** 89–95

Fang Z, Zhao J, Xie W, Sun Q, Wang H and Qiao B 2017 LncRNA UCA1 promotes proliferation and cisplatin resistance of oral squamous cell carcinoma by sunppressing miR-184 expression. *Cancer Med.* **6** 2897–2908

Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E *et al.* 2013 Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6** pl1

Ghafouri-Fard S, Esmaeili M and Taheri M 2020 H19 lncRNA: Roles in tumorigenesis. *Biomed. Pharmacother.* **123** 109774

Gupta R, Shah N, Wang K, Kim J, Horlings H, Wong D, Tsai M, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB van de VM and Sukumar S CH 2010 Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464** 1071–1076

Gyorffy B, Surowiak P, Budczies J and Lánczky A 2013 Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS ONE* **8** e82241

Han Y, Liu Y, Gui Y and Cai Z 2013 Long intergenic non-coding RNA TUG1 is overexpressed in urothelial carcinoma of the bladder. *J. Surg. Onco.* **107** 555–559

Henson B, Bhattacharjee S, O'Dee D, Feingold E and Gollin S 2009 Decreased expression of miR-125b and miR-100 in oral cancer cells contributes to malignancy. *Genes Chromosomes Cancer* **48** 569–582

Hsu C, Lin P, Lin H, Lai C, Yang C, Lin S and Yang M 2016 Altered expression of imprinted genes in squamous cell carcinoma of the head and neck. *Anticancer Res.* **36** 2251–2258

Hu L, Wu Y, Tan D, Meng H, Wang K, Bai Y and Yang K 2015 Up-regulation of long noncoding RNA MALAT1 contributes to proliferation and metastasis in esophageal squamous cell carcinoma. *J. Exp. Clin. Cancer Res.* **34** 7

Hutchinson J, Ensminger A, Clemson C, Lynch C, Lawrence J and Chess A 2007 A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8** 39

Jain D, Dravid C, Singla A, Kumari S and Grover R 2019 Comparison of the Seventh and Eighth Editions of the American Joint Committee on Cancer pT and pN Classifications as predictors of survival in patients with oral squamous cell carcinoma. *Am. J. Clin. Pathol.* **151** 292–301

Kim J, Piao H, Kim B, *et al.* 2018 Long noncoding RNA MALAT1 suppresses breast cancer metastasis. *Nat. Genet.* **50** 1705–1715

Koch A, De Meyer T, Jeschke J and Van Criekinge W 2015 MEXPRESS: Visualizing expression, DNA methylation and clinical TCGA data. *BMC Genomics* **16** https://doi.org/10.1186/s12864-015-1847-z

Kwok Z, Roche V, Chew X, Fadieieva A and Tay Y 2018 A non-canonical tumor suppressive role for the long non-coding RNA MALAT1 in colon and breast cancers. *Int. J. Cancer* **143** 668–678

Lan X, Sun W, Dong W, Wang Z, Zhang T, He L and Zhang H 2018 Down-regulation of long noncoding RNA H19 contributes to the proliferation and migration of papillary thyroid carcinoma. *Gene* **646** 98–105

Lv J, Ma L, Chen X, Huang X and Wang Q 2014 Down-regulation of LncRNAH19 and MiR-675 promotes migration and invasion of human hepatocellular carcinoma cells through AKT/GSK-3β/Cdc25A signaling pathway. *J. Huazhong Univ. Sci. Technol. Med. Sci.* **34** 363–369

Mallath MK, Taylor DG, Badwe RA, *et al.* 2014 The growing burden of cancer in India: Epidemiology and social context. *Lancet Oncol.* **15** e205–e212

Nordin M, Bergman D, Halje M, Engström W and Ward A 2014 Epigenetic regulation of the Igf2/H19 gene cluster. *Cell Prolif.* **47** 189–199

Qiu Y Le, Liu YH, Ban JD, Wang WJ, Han M, Kong P and Li BH 2019 Pathway analysis of a genome-wide association study on a long non-coding RNA expression profile in oral squamous cell carcinoma. *Oncol. Rep.* **41** 895–907

Raveh E, Matouk IJ, Gilon M and Hochberg A 2015 The H19 long non-coding RNA in cancer initiation, progression and metastasis - a proposed unifying theory. *Mol. Cancer* **14** 1–14

Sun Y and Ma L 2019 New Insights into long non-coding RNA MALAT1 in cancer and metastasis. *Cancers* **11** pii: E216

Sun S, Wu Y, Guo W, *et al.* 2018 Stat3/hotair signaling axis regulates hnscc growth in an ezh2-dependent manner. *Clin. Cancer Res.* **24** 2665–2677

Tang H, Wu Z, Zhang J and Su B 2013 Salivary lncRNA as a potential marker for oral squamous cell carcinoma diagnosis. *Mol. Med. Rep.* **7** 761–766

Tripathi V, Shen Z, Chakraborty A, *et al.* 2013 Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. *PLoS Genet.* **9** e1003368

Vishwakarma S, Agarwal R, Goel SK, Panday RK, Singh R, Sukumaran R, Khare S and Kumar A 2017 Altered expression of sphingosine-1-phosphate metabolizing enzymes in oral cancer correlate with clinicopathological attributes. *Cancer Invest.* **35** 139–141

Xu S, Sui S, Zhang J, *et al.* 2015 Down-regulation of long noncoding RNA MALAT1 induces epithelial-to-mesenchymal transition via the PI3K-AKT pathway in breast cancer. *Int. J. Clin. Exp. Pathol.* **8** 4881–4891

Xu Y, Yao Y, Leng K, Li Z, Qin W, Zhong X, Kang P, Wan M, Jiang X and Cui Y 2017 Long non-coding RNA UCA1 indicates an unfavorable prognosis and promotes tumorigenesis via regulating AKT/GSK-3β signaling pathway in cholangiocarcinoma. *Oncotarget* **8** 96203–96214

Yan G, Wang X, Yang M, Lu L and Zhou Q 2017 Long non-coding RNA TUG1 promotes progression of oral

squamous cell carcinoma through up-regulating FMNL2 by sponging miR-219. *Am. J. Cancer Res.* **7** 1899–1912

Yan J, Chen D, Chen X, Sun X, Dong Q, Hu C, Zhou F and Chen W 2019 Down-regulation of lncRNA CCDC26 contributes to imatinib resistance in human gastrointestinal stromal tumors through IGF-1R up-regulation. *Braz. J. Med. Biol. Res.* **52** e8399

Yang YT, Wang YF, Lai JY, Shen SY, Wang F, Kong J, Zhang W and Yang HY 2016 Long non-coding RNA UCA1 contributes to the progression of oral squamous cell carcinoma by regulating the WNT/β-catenin signaling pathway. *Cancer Sci.* **107** 1581–1589

Yao F, Wang Q and Wu. Q 2019 The prognostic value and mechanisms of lncRNA UCA1 in human cancer. *Cancer Manag. Res.* **11** 7685–7696

Yu X and Li Z 2015 Long non-coding RNA HOTAIR: A novel oncogene (Review). *Mol. Med. Rep.* **12** 5611–5618

Yuan Z, Yu Y, Zhang B, *et al.* 2018 Genetic variants in lncRNA H19 are associated with the risk of oral squamous cell carcinoma in a Chinese population. *Oncotarget* **9** 23915–23922

Zhang EB, Yin DD, Sun M, *et al.* 2014 P53-regulated long non-coding RNA TUG1 affects cell proliferation in human non-small cell lung cancer, partly through epigenetically regulating HOXB7 expression. *Cell Death Dis.* **5** 1–12

Zhang DM, Lin ZY, Yang ZH, Wang YY, Wan D, Zhong JL, Zhuang PL, Huang ZQ, Zhou B and Chen WL 2017 lncRNA H19 promotes tongue squamous cell carcinoma progression through β-catenin/GSK3β/EMT signaling via association with EZH2. *Am. J. Transl. Res.* **9** 3474–3486

Zhang L, Meng X, Zhu X, Yang D, Chen R, Jiang Y and Xu T 2019a Long non-coding RNAs in Oral squamous cell carcinoma: biologic function, mechanisms and clinical implications. *Mol. Cancer* **18** 102

Zhang T, Liang L, Liu X, Wu J, Su K, Chen J and Zheng Q 2019b LncRNA UCA1/miR-124 axis modulates TGFβ1-induced epithelial-mesenchymal transition and invasion of tongue cancer cells through JAG1/Notch signaling. *J. Cell Biochem.* **120** 10495–10504

Zhu M, Chen Q, Liu X, *et al.* 2014 LncRNA H19/miR-675 axis represses prostate cancer metastasis by targeting TGFBI. *FEBS J.* **281** 3766–3775

Corresponding editor: Rita Mulherkar