Original Article
Long noncoding RNA, LINC00460, as a prognostic biomarker in head and neck squamous cell carcinoma (HNSCC)

Ritu Chaudhary¹, Xuefeng Wang², Biwei Cao², Janis De La Iglesia³, Jude Masannah¹, Feifei Song¹, Juan C Hernandez-Prera³, Nicholas T Gimbrone³, Robbert JC Slebos¹, Christine H Chung¹

Departments of ¹Head and Neck-Endocrine Oncology, ²Biostatistics and Bioinformatics, ³Pathology, Moffitt Cancer Center, Tampa, FL, USA

Received August 20, 2019; Accepted January 17, 2020; Epub February 15, 2020; Published February 28, 2020

Abstract: Head and neck squamous cell carcinoma (HNSCC) is an aggressive epithelial malignancy characterized by frequent mutations and metastasis. Long noncoding RNAs (IncRNAs) have been implicated in tumorigenesis and serve as novel prognostic biomarkers in different cancers. To enhance our understanding of IncRNAs that may have biological significance in HNSCC and may serve as prognostic biomarkers, we globally profiled IncRNAs in HNSCC by analyzing the RNA-seq data from The Atlas of Noncoding RNAs in Cancer (TANRIC) database. Of 3576 IncRNAs, we identified 926 (higher-688, lower-238) IncRNAs with a 2-fold abundance difference among the forty HNSCC and paired adjacent normal tissue. We investigated differential abundance of IncRNAs based on TP53 mutation and p16 status. We found 133 IncRNAs to have differential abundance by 2-fold among the mutant vs wild-type TP53 samples, whereas among p16-negative vs positive samples, we identified 710 IncRNAs with the same criteria. Meanwhile, analysis of the 15 most abundant IncRNAs in the tumor samples identified five IncRNAs whose higher abundance was associated with poor overall patient survival. Among these five, higher abundance of LINC00460 associated with poor patient survival in an independent cohort of 82 HNSCC patients. To further evaluate the potential function of LINC00460, we performed IncRNA-mRNAs co-expression analysis and found that higher abundance of LINC00460 associated with cancer-related biological pathways including EMT and other inflammatory response pathways. In summary, we report LINC00460 is more abundant in tumors compared to adjacent normal tissue and that it may serve as a potential prognostic biomarker in HNSCC.

Keywords: Head and neck squamous cell carcinoma (HNSCC), long noncoding RNAs (IncRNA), RNA-seq, prognostic biomarker, human papillomavirus (HPV)

Introduction
Head and neck cancer (HNC) represents the sixth most common cancer worldwide and accounts for 4% of all cancers in the United States (https://www.cancer.net/). In 2019, an estimated 65,410 new cases of HNC will develop in US alone, leading to an estimated 14,620 deaths [1]. HNC comprise a broad range of tumors including tumors of the oral cavity, pharynx, larynx, nasal cavity, paranasal sinuses and salivary glands [2]. More than 90% of HNC originate from mucosal epithelium and are referred to as head and neck squamous cell carcinoma (HNSCC) [3]. The risk factors for developing HNSCC include smoking, excessive alcohol consumption, and human papillomavirus (HPV) infection [4, 5]. HPV status is an independent prognostic factor for overall survival in HNSCC [6]. Despite considerable advancements in the treatment regimen and increased understanding of the associated risk factors, the clinical outcome for HNSCC remains poor due to tumor recurrence and metastasis [7]. Therefore, a better understanding of molecular alteration and mechanisms underlying HNSCC, as well as identification of prognostic biomarkers, would facilitate improved diagnostic and therapeutic approaches for HNSCC treatment.

To understand the biology and mechanism of HNSCC progression, research has focused on the genetic alterations of protein coding genes including TP53, CDKN2A/B, MYC, EGFR,
LncRNAs in HNSCC

PIK3CA and PTEN [8-10]. In recent years, micro-RNAs were shown to have a vital role in HNSCC progression [11, 12]. However, the role of an evolving class of noncoding RNAs known as long noncoding RNAs (lncRNAs) in HNSCC initiation and progression is poorly understood. LncRNAs are transcripts >200 nucleotides long and lack a functional open reading frame. Recent studies have shown that aberrant expression of lncRNAs is associated with different cancer types. Several lncRNAs that have been functionally characterized play key roles in cancer related pathways including proliferation, migration, and genome stability [13, 14]. Functional variability and diverse expression pattern make lncRNAs suitable targets for diagnostic and therapeutic purposes [15, 16]. In HNSCC, biological function and molecular mechanism of only a handful of lncRNAs have been identified, including MALAT1, HOTAIR, FOXCUT, H19, CCAT1, LINC00312, NEAT1, TUG1, UCA1, SOX21-AS1 and NKILA, and these lncRNAs have been implicated in cancer related pathways including migration, invasion, and metastasis [17-27].

In this study, we investigated IncRNAs abundance changes in HNSCC by utilizing RNA-seq data from The Atlas of Noncoding RNAs in Cancer (TANRIC) database, an open-access resource that contains lncRNAs expression data in patient cohort from 20 different cancer types, including TCGA, Broad institute cancer cell line encyclopedia, and others. For our analysis, we utilized TANRIC dataset consisting of lncRNA expression data for about 12,727 genes from 40 tumor-adjacent normal pair and 386 unpaired tumors from the TCGA HNSCC cohort of 426 cases [28]. We included only lncRNAs with the 25th-percentile RPKM value higher than 0 and whose 90th-percentile RPKM value was greater than 0.1, resulting in a lncRNA set of 3576 lncRNAs.

Clinical data for 426 TCGA HNSCC samples were downloaded from the TCGA Data Portal. Patient TP53 status was obtained from cBioPortal (http://www.cbioportal.org/). Patient smoking status was defined as follows: Smokers (smoked more than 100 cigarettes in their lifetime and are currently smoking) and non-smokers (smoked less than 100 cigarettes in lifetime).

Independent cohort clinical information and RNA-seq

HNSCC patients were identified through three IRB approved studies (between 2006 and 2016) at Moffitt Cancer Center: Total Cancer Care (MCC#14690) with a tissue diagnosis of HNSCC, Epidemiology of Head and Neck Cancer Study (MCC#17041) and Evaluation of The Tumor and Its Microenvironment in Head and Neck Cancer Patients (MCC#18754). From these studies, a cohort of 82 cases with p16 immunohistochemical staining (IHC)-negative squamous cell carcinoma of the oral cavity, pharynx and larynx were included (Table 1). Clinical data was extracted from medical records and included patient demographics, smoking history, treatment history, and disease outcome. Patient smoking status was defined as mentioned above.

RNA-seq on all tumors was performed at the HudsonAlpha Institute for Biotechnology. Total RNA was extracted using Qiagen RNAeasy plus mini kit. RNA from all formalin-fixed paraffin-embedded (FFPE) samples was extracted using the Qiagen All prep FFPE DNA/RNA kit. RNA libraries were prepared using the Illumina TruSeq RNA Exome protocol and kit reagents.

Material and methods

TANRIC database and clinical information of HNSCC patients

For lncRNA abundance, we downloaded the HNSCC RNA-seq data from The Atlas of Noncoding RNAs in Cancer (TANRIC, https://ibl.mdanderson.org/tanric/_design/basic/index.html), an open-access resource that contains lncRNAs expression data in patient cohort from 20 different cancer types, including TCGA, Broad institute cancer cell line encyclopedia, and others. For our analysis, we utilized TANRIC dataset consisting of lncRNA expression data for about 12,727 genes from 40 tumor-adjacent normal pair and 386 unpaired tumors from the TCGA HNSCC cohort of 426 cases [28]. We included only lncRNAs with the 25th-percentile RPKM value higher than 0 and whose 90th-percentile RPKM value was greater than 0.1, resulting in a lncRNA set of 3576 lncRNAs. Clinical data for 426 TCGA HNSCC samples were downloaded from the TCGA Data Portal. Patient TP53 status was obtained from cBioPortal (http://www.cbioportal.org/). Patient smoking status was defined as follows: Smokers (smoked more than 100 cigarettes in their lifetime and are currently smoking) and non-smokers (smoked less than 100 cigarettes in lifetime).

Independent cohort clinical information and RNA-seq

HNSCC patients were identified through three IRB approved studies (between 2006 and 2016) at Moffitt Cancer Center: Total Cancer Care (MCC#14690) with a tissue diagnosis of HNSCC, Epidemiology of Head and Neck Cancer Study (MCC#17041) and Evaluation of The Tumor and Its Microenvironment in Head and Neck Cancer Patients (MCC#18754). From these studies, a cohort of 82 cases with p16 immunohistochemical staining (IHC)-negative squamous cell carcinoma of the oral cavity, pharynx and larynx were included (Table 1). Clinical data was extracted from medical records and included patient demographics, smoking history, treatment history, and disease outcome. Patient smoking status was defined as mentioned above.

RNA-seq on all tumors was performed at the HudsonAlpha Institute for Biotechnology. Total RNA was extracted using Qiagen RNAeasy plus mini kit. RNA from all formalin-fixed paraffin-embedded (FFPE) samples was extracted using the Qiagen All prep FFPE DNA/RNA kit. RNA libraries were prepared using the Illumina TruSeq RNA Exome protocol and kit reagents.
and further sequenced on an Illumina HiSeq 4000 with 100 million total reads per sample (50 million paired reads). Raw count data was normalized and further analyzed using DESeq2 package [29].

Statistical analysis

Principle component analysis (PCA) was performed to assess whether the grouping of samples based on lncRNA gene expression profiles reflects tissue types or whether it reveals potential patient-level batch effects. The input values for PCA consisted log2 transformed RPKM expression values of all lncRNAs retained in the analysis. We extracted the first principal component (PC1) and the second principal component (PC2) from the PCA and then visualized the sample clusters using scatter plot. To identify any significant differentially expressed IncRNA genes (DEGs) between patient groups (defined by TP53 and p16 status), differential expression analysis of IncRNAs was carried out using lmFit function from the limma R package (version 3.36.5). Differentially expressed genes were selected using fold-change above 2 and a FDR value of <0.05 was used as the cutoff values.

In order to further explore the co-expression relationships between the IncRNA and pathways defined by protein-coding genes, we selected protein-code genes that are correlated with IncRNA biomarkers (based on Spearman’s rank correlation analysis). On this list of ranked genes, we performed gene set enrichment analysis (GSEA) using Java GSEA software version 3.0 (http://software.broadinstitute.org/gsea) with default settings and selecting the phenotype label as the permutation type. The Molecular Signatures Database (MSigDB) version 6.2 Hallmark (H) and oncogenic signatures (C6) gene set collections were evaluated in the GSEA. Finally, Kaplan-Meier estimator and log-rank test were used to assess the survival differences between patient groups defined by different IncRNA expression signatures. All statistical analyses were performed using R version 3.5.1.

Results

Identification of differentially regulated IncRNAs in HNSCC

We evaluated the differential expression profile of the long non-coding RNAs (IncRNA) between HNSCC tumors and adjacent matched normal tissue by performing principal component analysis (PCA). PCA showed that the expression levels of IncRNAs in tumors are considerably different from that of adjacent normal tissue (Figure 1A). To identify deregulated IncRNAs in HNSCC, we explored the IncRNA abundance using RNA-seq data from 40 tumor-normal paired samples in the TANRIC database. Statistical analysis using limma t-test revealed a total of 926 differentially expressed IncRNAs (FDR ≤0.05) between HNSCC and adjacent normal tissue (Figure 1B). Among these, 688 IncRNAs were significantly higher by at least 2-fold with LINCO1614 (AC093850, ENSG00000230838.1) to be the most abundant IncRNA (5-fold) and 238 IncRNAs were significantly lower by at least 2-fold with LINCO2487 (ENSG00000203688.4) to be the least abundant IncRNA (5-fold) (Figure 1B).
Next, we evaluated the expression of the top ten and bottom ten most abundant lncRNAs from the tumor-normal paired analysis in all tumors (n=426) versus normal tissue (n=42) using the nonparametric Mann-Whitney rank test. Our analysis confirmed that all ten lncRNAs with higher abundance and ten lncRNAs with lower abundance from the tumor-normal paired analysis were also differentially expressed in all tumors compared to adjacent normal tissue (Figure 1C and 1D).

**LncRNAs associated with HPV, smoking, and TP53 mutation**

High-risk HPV is associated with a subset of HNSCC [30, 31]. Presence of HPV in the HN-SCCs was associated with significantly improved response to given treatments and overall patient survival [32]. In HNSCC, the role of HPV on lncRNAs expression is poorly understood; therefore, we evaluated the abundance of lncRNAs with presence or absence of HPV. Using immunohistochemical detection of p16 as a surrogate marker for HPV status obtained from the HNSC TCGA, we performed limma t-test on TANRIC lncRNA expression data consisting of p16-positive (n=24) and p16-negative (n=65) patient samples (Table 1). Using a 2-fold cutoff value and FDR ≤0.05, we identified 710 lncRNAs to be differentially regulated among p16-negative and positive samples (Figure 2A). Among these, 201 lncRNAs had significantly higher and 509 lncRNAs had significantly lower
Figure 2. Differential abundance of lncRNAs by p16 expression and TP53 mutational status. A. Volcano plot of differentially expressed lncRNAs comparing p16-negative and p16-positive HNSCCs. The bar chart represents the top ten most differentially expressed lncRNAs (both increased and decreased in p16-positive tumors). Red: most abundant in p16-negative tumors. Blue: most abundant in p16-positive tumors. B. Volcano plot of differentially expressed lncRNAs comparing mutant-TP53 and wild-type-TP53 HNSCCs. The bar chart represents top ten most differential lncRNAs in mutant versus wild-type p53 (in both directions). Red: most abundant in mutant-TP53 tumors. Blue: most abundant in wild type-TP53 tumors. C. Table representing number of common lncRNAs classified on the basis of higher or lower abundance in tumor samples, p16-negative samples and mutant-TP53 status.
LncRNAs in HNSCC

Table 2. LncRNAs commonly deregulated in the context of tumorigenesis, HPV and TP53 mutation

<table>
<thead>
<tr>
<th>61 lncRNAs with higher abundance in tumor, p16-negative and mutant-TP53</th>
<th>5 lncRNAs with lower abundance in tumor, p16-negative and mutant-TP53</th>
</tr>
</thead>
</table>

abundance in p16-negative tumors compared to p16-positive tumors (Figure 2A).

Mutations in the tumor suppressor gene, TP53, are the most frequent of all somatic genomic alterations in HNSCC, and these mutations are associated with poor survival and resistance to chemotherapy in HNSCC patients [9, 10, 33]. To identify lncRNAs associated with TP53 mutation, we performed statistical analysis using limma t-test on the TANRIC lncRNA expression data consisting of wild-type (n=102) and mutant-TP53 (n=281) patient samples. Overall, we identified 133 lncRNAs to be differentially regulated among mutant vs wild-type-TP53 samples with significance (fold change 2-fold and FDR ≤0.05) (Figure 2B). Among these, 104 lncRNAs had a significantly higher abundance and 29 lncRNAs were of lower abundance in the mutant-TP53 patient samples (Figure 2B). Our results indicate that mutations in TP53 affect the abundance of these lncRNAs and that these lncRNAs may affect the downstream regulatory TP53-pathways.

Smoking increases the risk of HNSCC and is a prognostic factor in HNSCC. The role of smoking on lncRNA expression in HNSCC is not well studied. Hence, we analyzed the differential expression profile between smokers and non-smokers in the HNSCC tumors in TANRIC database. Overall, we identified four lncRNAs that had higher and 19 lncRNAs that had lower abundance in the non-smokers with a ≥2 fold-change at FDR ≤0.05 compared to smokers. The small number of differentially expressed lncRNAs suggests that the abundance of lncRNAs may not be significantly affected by the smoking status.

As our analysis indicates that HPV and TP53 mutation are major factors contributing to differential expression of selected lncRNAs, we next investigated if any of the 926 tumor-associated lncRNAs in Figure 1B overlap with the 710 HPV associated and/or 133 TP53 mutation associated lncRNAs. Interestingly, we found that about 9% (61/688) of the lncRNAs with increased abundance in tumors also have increased abundance associated with p16-negative and mutant-TP53 tumors (Figure 2C; Table 2). Although it's a small percentage of lncRNAs, as mentioned before, HPV and TP53 are prognostic factors contributing to HNC progression, and our findings suggests that some of the lncRNAs of high abundance in tumor samples may have an important regulatory role in p16-negative and mutant-TP53 tumor progression.

Association of patient survival with lncRNAs

In order to determine which lncRNAs could be potentially significant in the pathogenesis of HNSCC, we evaluated the association between lncRNA abundance and overall survival using the clinical data from TCGA. We analyzed the abundance of the top 15-upregulated lncRNAs in 426 HNSCC patients and divided them into two groups based on the median abundance value. We found that high abundance of five lncRNAs LINC01614, CASC9, LINC02081, LINC00460 and LINC01980 were associated with poor patient survival under univariate Kaplan-Meier analysis (Figure 3A).

It is well-established that HPV status can affect patient prognosis in HNSCC. Therefore, we evaluated the effect of lncRNA regulation on patient survival in p16-positive and p16-negative cohorts respectively, using log-rank test. In the p16-negative patient cohort (n=65), among these five lncRNAs only the high abundance of LINC00460 was associated with poor overall...
Figure 3. Kaplan-Meier curves showing correlation of five tumor associated lncRNAs with overall patient survival. A. Higher abundance of lncRNAs including LINC01614, LINC00460, LINC02081, LINC01980, and CASC9 was significantly associated with poor overall survival in TCGA dataset (n=426). B. Higher abundance of LINC00460 in the p16-negative patient cohort was significantly associated with poor overall survival (n=65). C. Higher abundance of LINC00460 was significantly associated with poor overall survival in an independent cohort consisting of 82 HNSCC cases.
survival (Figure 3B). Whereas in the p16-positive patient cohort (n=24), none of the lncRNAs were associated with patient survival, although this comparison had low statistical power due to the small sample size (data not shown).

To further examine the prognostic value associated with lncRNAs, we performed Kaplan-Meier analysis in an independent cohort consisting of 82 p16-negative HNSCC obtained from the Moffitt Cancer Center (Table 1). Survival analysis showed that higher abundance of LINC00460 in this independent cohort was also associated in poor overall survival (Figure 3C), thus indicating that LINC00460 may be a prognostic factor for patient survival in p16-negative HNSCC.

LncRNA-protein co-expression analysis

As our analysis suggests that higher abundance of LINC00460 is associated with poor patient survival, we became interested to delineate its biological function. To do so, we performed LINC00460-mRNAs co-expression analysis by performing Spearman’s rank correlation analysis to select protein-encoding genes associated with this lncRNA. Using gene expression data from TCGA, we compared the ranked gene sets with Gene Set Enrichment Analyses (GSEA) against the Hallmark of Cancer and Oncogenic gene signatures [34]. We found that higher expression of genes correlated with LINC00460 were associated with biological pathways including epithelial-mesenchymal transition (EMT), interferon alpha response (INF-α), angiogenesis, inflammatory response, KRAS signaling and TGF-β signaling (Figure 4A). Among the oncogenic signature pathways, we found that the LINC00460-related genes were associated with upregulation of TBK1 and STK33 and downregulation of RB. More importantly, we were able to validate some of these findings in the independent patient cohort (n=82) using gene expression data from the Moffitt Cancer Center (Figure 4B). In this independent data set, higher abundance of LINC00460-related genes were associated with increased expression of EMT, STK33 and TBK1 similar to what was observed in the TCGA data. However, the list of pathways did not completely overlap, probably due to the heterogeneity in the tumors and technical variations in these two datasets.

Discussion

Emerging evidence suggests that lncRNAs play a vital role in cancer related pathways and can mediate oncogenic functions by interacting with proteins, DNA, or RNAs, and regulating gene expression transcriptionally or post-translationally [13, 14]. In recent years, lncRNAs have been identified as targets for novel diagnostic and therapeutic strategies [15, 16]. LncRNAs have been implicated in tumor progression, invasion, and metastasis of HNSCCs, indicating that they might function as novel biomarkers and/or as therapeutic targets for patient treatment [35-37]. However, in HNSCC there is only a small fraction of lncRNAs with known biological function while the vast majority of lncRNAs are of unknown function. To further enrich our understanding of lncRNA functioning in HNSCC, we analyzed lncRNA expression data from the TANRIC HNSCC database and here report that novel lncRNAs can be potential biomarkers in HNSCC prognosis.

We investigated differential expression profile of lncRNAs in tumor compared to adjacent, matched normal HNSCC from TCGA. Interestingly, some of the well-studied lncRNAs that have been previously reported to be differentially regulated in HNSCC (for example MALAT1, UCA1, NEAT1, MEG3 and Gas5) were either not found in our analysis or their expression level was not significantly different between the tumor and normal tissues. Similarly, there were other research groups that reported inconsistent findings [38-40]. These discrepancies may arise due to different sequencing approaches being used for analyzing the global lncRNA expression changes, tissue heterogeneity, and/or heterogeneity and inadequate sample size of different patient cohorts. Moreover, lncRNA expression may vary among tissue locations because many lncRNAs are known to be expressed in a tissue-type specific manner [41]. On the other hand, several lncRNAs that have been previously reported to be of direct relevance in HNSCC were identified in our analysis. For instance, HOTAIR a very well-studied lncRNA in different cancers was more abundant in HNSCC tumor samples than in normal tissue, a finding that was consistent with reports in literature for the different HNSCC subtypes. In laryngeal squamous cell carcinoma, HOTAIR is overexpressed and regulates
PTEN methylation [42]. Similarly, in oral squamous cell carcinoma, HOTAIR promotes tumor cell invasion and metastasis by recruiting EZH2 and repressing E-cadherin [18].

Human papillomavirus is reported as an independent prognostic factor in HNSCC and is associated with an improved prognosis in a subset of patients with HNSCC [6, 43]. Here, we report on lncRNAs with differential expression between p16-negative and p16-positive patient samples. Interestingly, several of the lncRNAs with high expression in the p16-negative samples are associated with several dis-
LncRNAs in HNSCC

eases including cancer. LINC00162 (also known as PICSAR) was among the top 20 upregulated lncRNAs in p16-negative samples and it has been previously shown to promote the progression of cutaneous squamous cell carcinoma by activation of the ERK1/2 signaling pathway and downregulating DUSP6 expression [44]. Similarly, the lncRNAs that were downregulated in p16-negative patients were associated with tumor progression. For instance, among the 20 downregulated lncRNAs, FEZF1-AS1 facilitates tumorigenesis and progression in colorectal, partly through FEZF1 induction [45] while PART1 lncRNA has been shown to function as competitive endogenous RNA by binding to miR-129 and thus facilitates Bcl-2 expression in esophageal squamous cell carcinomas [46]. High levels of serum lncRNA PART1 in esophageal squamous cell carcinomas are associated with poor patient response to gefitinib treatment. Additionally, IncRNA CDKN2B-AS1 (a.k.a. ANRIL) promotes tumorigenesis in HNSCC [47]. LncRNA CDKN2B-AS1 is antisense to CDKN2B and antisense downstream to CDKN2A (p16) gene, which has a tumor suppressor function. In the future, it will be interesting to investigate if lncRNA CDKN2B-AS1 interferes with biological function of p16 in HNSCC.

Mutations in the TP53 gene are very frequent in HNSCC [10], therefore we investigated the abundance of lncRNAs in the patients with mutant-TP53 versus wild-type-TP53. There were many lncRNAs with differential abundance with respect to TP53 mutational status, and several that had higher abundance in mutant-TP53 patients were reported to have functional significance cancer. For example, in gastric cancer, IncRNA LIN000668 is activated by E2F1 and acts as a regulator of cell cycle by associating with PRC2. This association results in epigenetic repression of cyclin-dependent protein kinase inhibitors [48]. It has been reported that the TP53 mutations are more frequently associated HPV-negative HNSCC [49]. Nohata et al. [39] reported differential IncRNAs expression in HPV-positive and wild-type-TP53 subsets from TANRIC. They identified 140-upregulated IncRNAs (>1.5-fold change and \( q < 0.001 \)) in HPV-positive tumors. Applying similar criteria to our dataset, we confirmed 91/140 IncRNAs. Similarly, they reported 30 IncRNAs to be upregulated (>1.5-fold change and \( q < 0.001 \)) in wild-type-TP53 samples. Applying similar criteria to our dataset, we confirmed 10/30 lncRNAs, indicating consistencies in the two studies.

We next focused on the lncRNAs that were commonly deregulated in the context of tumorigenesis, HPV, and TP53. Interestingly, we identified 61 lncRNAs with higher abundance in p16-negative and mutant-TP53 tumors. Among these, some of the lncRNAs have been functionally characterized in other cancers. For example, LINC00707 promotes cell proliferation and migration in lung adenocarcinoma by regulating Cdc42 [50]. LINC00704 (also called as MANCR) is upregulated in breast cancer and is functionally known to regulate genomic stability by regulating cell proliferation and viability [51]. LncRNA RP11-567G11.1 is up-regulated in pancreatic cancer tissues, and depletion of RP11-567G11.1 increases the sensitivity of pancreatic cancer cells to gemcitabine [52]. To determine if these lncRNAs can function as a potential biomarker in p16-negative HNSCC patients with TP53 mutation, we need to further investigate their mechanism of action in HNSCC.

A promising field of research has been on lncRNAs that may function as prognostic biomarkers in HNSCC progression. The main lncRNA with a prognostic value in both the TCGA and Moffitt patient cohort was LINC00460. In a previous study in HNSCC, LINC00460 was reported as a prognostic predictor of poor patient survival [53]. Our study confirms this role both in the TCGA and Moffitt HNSCC data sets. Therefore, our finding not only corroborates the previous study but further enhances our understanding of the prognostic value of lncRNAs in HNSCC progression. Furthermore, we performed lncRNA-mRNA co-expression analysis using TCGA dataset and found that higher LINC00460 is associated with many oncogenic and cancer related pathways including EMT, IFN-α, TGF-β, NF-κB, and other inflammatory pathways, and we were able to validate some of these pathways in an independent cohort. From this analysis, we speculate that LINC00460 might be interacting with the proteins that are involved in cell development, proliferation, adhesion, tumorigenesis, and other tumor promoting pathways. In esophageal squamous cell carcinoma, LINC00460 function as an oncogene and promotes cell growth and apoptosis [54]. Similarly, in lung cancer cells LINC00460 promotes cell migration and invasion and thus induces EMT.
In colorectal cancer cells, it affects cell proliferation and apoptosis by regulating the expression of Krüppel-like factor 2 and cullin 4A [56]. However, to elucidate the mechanism by which LINC00460 and other lncRNAs contribute to HNSCC pathogenesis and progression, further experimental validation will be needed.

In summary, we present a comprehensive analysis of lncRNAs abundance in HNSCC in context of p16 status and TP53 mutation status. Our analysis revealed LINC00460 as a prognostic marker of survival in patients with p16-negative HNSCC. Further studies are required to understand how LINC00460 plays a regulatory role in HNSCC contributing to worse prognosis.

Acknowledgements

This work is supported in part by the Moffitt’s Total Cancer Care Initiative, Collaborative Data Services, Biostatistics and Bioinformatics, and Tissue Core Facilities at the H. Lee Moffitt Cancer Center & Research Institute, an NCI-designated Comprehensive Cancer Center (P30-CA076292) and the James and Esther King Biomedical Research Grant (7JK02). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Grant Sponsor or the H. Lee Moffitt Cancer Center & Research Institute.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Christine H Chung, Department of Head and Neck-Endocrine Oncology, Moffitt Cancer Center, 12902 Magnolia Drive, Tampa, FL 33612, USA. E-mail: Christine.Chung@Moffitt.Org

References

LncRNAs in HNSCC


