Received: 24 May 2017 | Revised: 30 November 2017 | Accepted: 1 February 2018

DOI: 10.1002/hed.25136

ORIGINAL ARTICLE

WILEY

Differential expression and prognostic value of long non-coding RNA in HPV-negative head and neck squamous cell carcinoma

Sulsal-Ul Haque MD[1] | Liang Niu PhD[2] | Damaris Kuhnell MS[2] | Jacob Hendershot BS[2] | Jacek Biesiada PhD[2] | Wen Niu MS[2] | Matthew C. Hagan MD, PhD[3] | Karl T. Kelsey MD[4,5] | Keith A. Casper MD[6] | Trisha M. Wise-Draper MD, PhD[1,7] | Mario Medvedovic PhD[2] | Scott M. Langevin PhD[2]

[1] Department of Internal Medicine, Division of Hematology/Oncology, University of Cincinnati College of Medicine, Cincinnati, Ohio

[2] Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, Ohio

[3] Department of Pathology and Laboratory Medicine, University of Cincinnati College of Medicine, Cincinnati, Ohio

[4] Department of Epidemiology, Brown University, Providence, Rhode Island

[5] Department of Pathology and Laboratory Medicine, Brown University, Providence, Rhode Island

[6] Department of Otolaryngology, University of Michigan, Ann Arbor, Michigan

[7] Department of Cancer Biology, University of Cincinnati College of Medicine, Cincinnati, Ohio

Correspondence
Scott M. Langevin, 160 Panzeca Way, ML0056, Cincinnati, OH 45267.
Email: langevst@ucmail.uc.edu

Funding information
This work was supported by the National Cancer Institute (K22CA172358 to S.M. L.) and the National Institute for Environmental Health Sciences through the University of Cincinnati Center for Environmental Genetics (2P30ES006096)

Abstract

Background: Long non-coding RNA (lncRNA) has emerged as a new avenue of interest due to its various biological functions in cancer. Abnormal expression of lncRNA has been reported in other malignancies but has been understudied in head and neck squamous cell carcinoma (HNSCC).

Methods: The lncRNA expression was interrogated via quantitative real-time polymerase chain reaction (qRT-PCR) array for 19 human papillomavirus (HPV)-negative HNSCC tumor-normal pairs. The Cancer Genome Atlas (TCGA) was used to validate these results. The association between differentially expressed lncRNA and survival outcomes was analyzed.

Results: Differential expression was validated for 5 lncRNA (*SPRY4-IT1*, *HEIH*, *LUCAT1*, *LINC00152*, and *HAND2-AS1*). There was also an inverse association between *MEG3* expression (not significantly differentially expressed in TCGA tumors but highly variable expression) and 3-year recurrence-free survival (RFS).

Conclusion: We identified and validated differential expression of 5 lncRNA in HPV-negative HNSCC. Low *MEG3* expression was associated with favorable 3-year RFS, although the significance of this finding remains unclear.

KEYWORDS
head and neck cancer, head and neck squamous cell carcinoma (HNSCC), long noncoding RNA (lncRNA), noncoding RNA (ncRNA), survival, The Cancer Genome Atlas (TCGA)

1 | INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common cancer type worldwide and accounts for approximately 350 000 deaths per year.[1,2] Risk factors, such as tobacco, alcohol use, and, more recently, human papillomavirus (HPV) have been identified as etiological factors in the occurrence of HNSCC. Despite advances in the treatment

of localized HNSCC, approximately half of the patients will develop recurrent disease,[3] which is a major contributor to patient mortality. In addition, HPV-positive and HPV-negative HNSCC are biologically distinct with the latter being associated with poorer prognosis. Unlike other malignancies, there are no tools in widespread use that can identify early disease and there is no systematic approach that has proven effective in monitoring for early evidence of recurrence. Hence, novel markers are needed that appropriately characterize those patients with early-stage disease as well as identify and characterize the response of individual patient's treatment.

Although only 1.2% of our DNA sequence encodes proteins, approximately 75% of the human genome is capable of being transcribed into RNA,[4] and it has become increasingly apparent that RNA plays a diverse and important role in genome integrity through production of both proteins as well as non-coding RNA (ncRNA), including long non-coding RNA (lncRNA). Whereas the microRNA class of ncRNA has been widely studied in the context of cancer, lncRNA, which are larger (>200 bases and approximately 1-2 exons in length)[5] and have a more complex secondary/tertiary structure, have recently begun to garner increased attention.[6] This is largely due to their diverse biological functions,[7] which can include inhibition of target gene transcription, initiation of alternative splicing, generation of protein scaffolding and chromatin organization, and alteration of transcription factor activity.

Even though 18 000 human long non-coding human transcripts have been catalogued in GENCODE version 7,[4] the lncRNA Database (http://www.lncrnadb.org), a database of functionally annotated eukaryotic lncRNA, only contains information for 127 human lncRNA, highlighting the gap in our knowledge of lncRNA biology. Identification of lncRNA transcripts that are associated with human diseases and the corresponding pathobiology (e.g. aggressiveness or responsiveness to treatment) would, therefore, provide a welcome means for prioritizing functional studies.

Several lncRNA have recently been reported to be differentially expressed in various cancers and to play a role in cancer growth, invasion, epithelial-to-mesenchymal transition, motility, and metastatic potential[8,9] and, as such, are increasingly recognized as having a very strong potential as cancer biomarkers.[10] Several recent studies have reported that higher expression of lncRNA, such as *HOX antisense intergenic RNA* (*HOTAIR*) and *metastasis associated lung adenocarcinoma transcript 1* (*Malat-1*), is associated with invasion and metastasis in various epithelial cancers,[11–17] although the prognostic significance of lncRNA in HNSCC has been understudied to date. The purpose of this study was to identify differentially expressed lncRNA in HPV-negative HNSCC and assess its impact on outcomes in these groups of patients.

2 | MATERIALS AND METHODS

2.1 | Tumor samples

Initial identification of differentially expressed lncRNA in HPV-negative HNSCC was conducted using a quantitative real-time polymerase chain reaction (qRT-PCR) array that interrogates 84 lncRNA with reported involvement in various cancers using paired archival formalin-fixed paraffin-embedded tumor tissue and paired adjacent normal squamous tissue from 19 patients treated for incident HNSCC at the University of Cincinnati Cancer Institute (UCCI). Raw RNA-sequencing (RNA-seq) reads for all available HPV-negative head and neck cancer tumors (n = 444) and all available adjacent normal tissue (n = 44) was downloaded from The Cancer Genome Atlas (TCGA) for independent validation of significant results.

2.2 | RT2 Profiler long non-coding RNA quantitative polymerase chain reaction cancer array

Total RNA was extracted from each sample using the RNeasy formalin-fixed paraffin-embedded kit (Qiagen, Valencia, CA), and converted to cDNA using the RT2 First Strand kit (Qiagen), according to the manufacturer's respective suggested protocols. The RT2 Profiler lncRNA quantitative polymerase chain reaction (PCR) Cancer PathwayFinder-Array (96-well format) was used to profile expression of 84 lncRNA transcripts that have been associated with various cancer types. The array also contains 3 reverse transcription controls, 3 positive PCR controls, a probe for human gDNA contamination, and 5 housekeeping genes for normalization: *beta actin* (*ACTB*); *β-2 microglobulin* (*B2M*); *ribosomal protein, large, P0* (*RPLP0*); *7SK small nuclear RNA* (*RN7SK*); and *small nucleolar RNA*, *H/ACA box 73A* (*SNORA73A*). The PCR mastermix was prepared with 250 ng of total cDNA and dispensed in 24 μl aliquots into each RT2 lncRNA PCR array well. The qRT-PCR was performed using an Applied Biosystems StepOnePlus Real-Time PCR System (ThermoFisher Scientific, Waltham, MA) under the following conditions: 10 minutes at 95°C using HotStart Taq Polymerase followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. A complete list of the 84 lncRNA included on the array can be found in Supporting Information Table S1.

2.3 | Differential expression

Expression was normalized to the geometric mean of 5 housekeeping control genes: *ACTB*; *B2M*; *RPLP0*; *RN7SK*; and *SNORA73A*.

Differential expression of each lncRNA was described in terms of fold-change for tumors relative to adjacent normal

TABLE 1 Clinicodemographic characteristics of HPV-negative head and neck squamous cell carcinoma cases from the University of Cincinnati Cancer Institute and The Cancer Genome Atlas validation set

	University of Cincinnati Cancer Institute (n = 19)	TCGA validation set (n = 444)	P difference
Age			
Years, median (range)	64 (51-86)	61.5 (19-90)	0.28[a]
Sex			0.83[b]
Male	14 (74%)	317 (71%)	
Female	5 (26%)	127 (29%)	
Race			0.44[b]
White	16 (84%)	355 (83%)	
Black	3 (16%)	44 (10%)	
Other	0 (0%)	30 (7%)	
Smoking status			0.01[b]
Never	1 (5%)	93 (21%)	
Former	15 (79%)	181 (42%)	
Current	3 (16%)	160 (37%)	
Primary tumor site			0.75[b]
Oral cavity	12 (63%)	294 (66%)	
Pharynx	1 (5%)	41 (9%)	
Larynx	6 (32%)	109 (25%)	
AJCC stage group			0.68[b]
I	0 (0%)	23 (5%)	
II	2 (11%)	74 (17%)	
III	5 (28%)	82 (18%)	
IV	11 (61%)	265 (60%)	
Tumor grade			0.25[b]
Well differentiated	6 (32%)	59 (13%)	
Moderately differentiated	9 (47%)	276 (63%)	
Poorly differentiated	4 (21%)	96 (22%)	
Undifferentiated	0 (0%)	2 (<1%)	
Undetermined	0 (0%)	8 (2%)	

Abbreviation: AJCC, American Joint Committee on Cancer; TCGA, The Cancer Genome Atlas.
[a]Wilcoxon rank sum test.
[b]Fisher's exact test.

tissue based on a 1-sample *t* test, adjusted for multiple testing by false discovery rate estimation and Q values using the methods proposed by Benjamini and Hochberg.[18] Expression was considered significantly differential where $Q \leq 0.1$.

2.4 | Replication using The Cancer Genome Atlas RNA-sequencing data

The RNA-seq data in the form of bam files of aligned reads was obtained for head and neck cancers in TCGA project were downloaded from the Cancer Genomics Hub. Reads not mapped to the human genome were aligned to HPV16 (NC_001526) E6 and E7 viral oncoprotein transcript sequences using *bowtie* aligner.[19] HPV status was inferred by designating the sample to be from an HPV-positive tumor if >1000 reads mapped to HPV oncoproteins and an HPV-negative tumor otherwise. The harmonized read counts for TCGA head and neck samples aligned to lncRNA defined in the GDC.h38 GENCODE version 22 GTF file were downloaded from the National Cancer Institute (NCI) Genomic Data Commons[19] using the *TCGAbiolinks* Bioconductor package.[20]

Normalized counts (count per million) for each of the 7 lncRNA in each sample were calculated using the *cpm* function in the Bioconductor package *edgeR*.[21] Since the distributions of lncRNA expression were nonlinear, differential expression was assessed nonparametrically using the

TABLE 2 Differentially expressed long non-coding RNA in HPV-negative head and neck squamous cell carcinoma

	Human lncRNA qPCR Cancer Array ($n_{T\text{-}N\ pairs}$ = 19)				TCGA RNA-seq replication set (n = 444)	
lncRNA	**Ensemble ID**	**Median fold-change**	***P* value**	**Q value**	**Median fold-change**	***P* value**
SPRY4-IT1	ENSG00000281881	2.19	.006	0.07	3.52	< .0001
HEIH	ENSG00000278970	0.56	.0001	0.01	0.64	< .0001
LUCAT1	ENSG00000248323	2.14	.002	0.04	1.76	< .0001
LINC00152	ENSG00000222041	5.11	.003	0.06	1.56	< .0001
HAND2-AS1	ENSG00000237125	0.47	.005	0.06	0.13	< .0001
MEG3	ENSG00000214548	0.16	.0003	0.01	0.33	.50
TERC	ENSG00000270141	2.32	.0007	0.02	0.64	.12

Abbreviations: lncRNA, long non-coding RNA; qPCR, quantitative polymerase chain reaction; RNA-seq, RNA-sequencing; TCGA, The Cancer Genome Atlas.
Q value significant at $Q < 0.01$.

Wilcoxon rank sum test, comparing tumor expression (n = 444) to that of all available normal samples (n = 44), and was considered differential where $P \leq .05$. Median fold-change was determined for each tumor by comparing its expression to the median expression of the normal samples. A description of the distribution of expression values for the 7 lncRNA in the 44 adjacent normal samples is provided in Supporting Information Table S2.

2.5 | Survival analysis using The Cancer Genome Atlas samples

To visualize the 5-year overall survival (OS) and 3-year recurrence-free survival (RFS), univariate Kaplan-Meier and cumulative incidence (to account for death as a competing risk[22,23]) functions were generated, respectively, comparing curves for low, normal, and high expression levels. Discrete multivariable Cox proportional hazards and cumulative incidence models were fit to assess 5-year OS and 3-year RFS, respectively, for each of the significant lncRNA, adjusted for age, sex, race, smoking status, tumor site, and stage at diagnosis, as established *a priori*. Missing values of model covariates were imputed (m = 20) using multivariate normal regression, based on age, sex, stage, and primary tumor site. Log-log plots (ie, -log(-log (S(t))) vs log(t)) were generated for each model to verify that the proportional hazards assumption was met. Statistical analyses were conducted using Stata 13 (Stata Corp, College Station, TX). All statistical tests were 2-sided, and significance was considered when the unadjusted $P \leq .05$.

3 | RESULTS

The median age for the UCCI cases (n = 19 tumor-normal pairs) was 64 years, 74% of which were men. The UCCI and TCGA sets differed in terms of smoking status ($P = .01$), with fewer nonsmokers in the UCCI set but were comparable in terms of age, sex, race, primary tumor site, American Joint Committee on Cancer (AJCC) stage group, and tumor grade (Table 1). The majority of tumors in both sets originated in the oral cavity (63% and 66%, respectively), and presented at an advanced stage (III or IV).

Eleven of the lncRNA included on the array were not detected in any of the tumor or normal samples (Supporting Information Table S1). Twenty lncRNA were significantly differentially expressed at a nominal P value $\leq .05$ (8 upregulated and 12 downregulated). After false discovery rate adjustment, 7 lncRNA remained significantly differential ($Q \leq 0.1$: *SPRY4-IT1*, *HEIH*, *LUCAT1*, *LINC00152*, and *HAND2-AS1*, *MEG3*, and *TERC*; 4 upregulated and 3 downregulated; Table 2). Expression of these 7 lncRNA was also significantly differential for 5 of the 7 lncRNA in the TCGA validation set (*SPRY4-IT1*, *HEIH*, *LUCAT1*, *LINC00152*, and *HAND2-AS1*). It should be noted, however, that although *MEG3* and *TERC* were not significantly differentially expressed in the TCGA validation set, there was wide variability in terms of expression in both directions (i.e., upregulation and downregulation) for individual tumors relative adjacent normal samples (Figure 1B). Interestingly, black patients in the TCGA dataset were significantly more likely to exhibit high (>2-fold) and low (<0.5-fold) *MEG3* expression (Table 3).

The relationship between each of the 7 lncRNA and 5-year OS and 3-year RFS was assessed using the TCGA dataset. Among the 7 lncRNA identified, only differing levels of *MEG3* had an impact on 3-year RFS. Patients with low *MEG3* expression (<0.5 fold change) were found to have a significantly lower 3-year RFS, whereas higher *MEG3* expression (>2-fold change) seemed to have better 3-year RFS, although this did not reach statistical significance

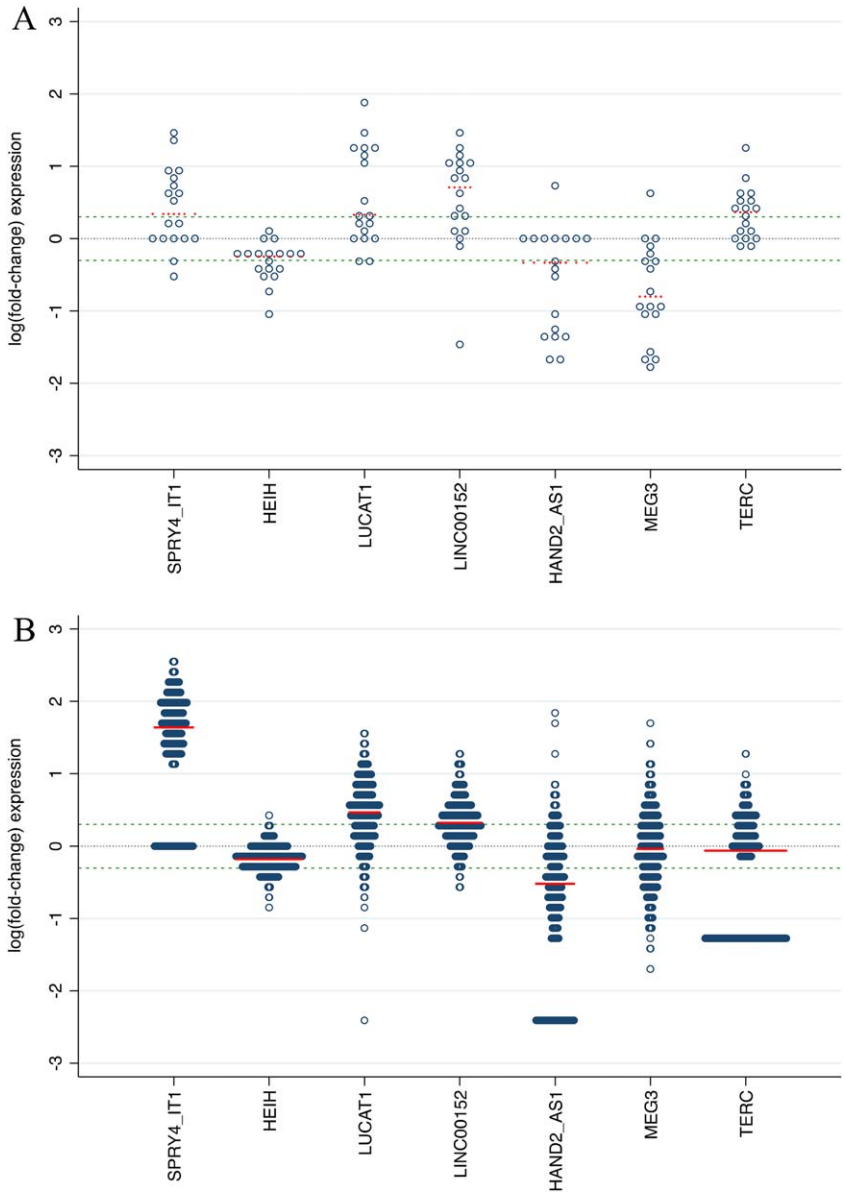

FIGURE 1 Distribution of log (fold-change) in HPV-negative head and neck squamous cell carcinoma tumors relative to adjacent non-malignant squamous tissue for A, University of Cincinnati patients (n = 19 tumor-normal pairs) and B, HPV-negative tumors (n = 444) and adjacent normal tissue (n = 44) with RNA-sequencing data through The Cancer Genome Atlas (TCGA). The black dashed center line at 0 corresponds to neutral expression and the blue dashed lines above and below the center line correspond to respective fold changes of 2.0 and 0.5, respectively; mean log (fold-change) for each long non-coding RNA is denoted by a red dashed line. Fold-change for the 19 tumors samples in A was calculated by comparing the expression of each tumor to its paired adjacent normal sample; fold-change for the 444 TCGA tumors in B was calculated by comparing the expression to the median expression of the 44 adjacent normal TCGA samples [Color figure can be viewed at wileyonlinelibrary.com]

(Figure 2 and Table 4). We also analyzed the associations between the 7 lncRNA and clinical characteristics of HPV-negative HNSCC in TCGA and found that low *MEG3* expression was associated with locally advanced disease and low expression of *HANDS-2AS1* correlated with more locally advanced cancer, although this did not seem to have an impact on survival outcomes. Restriction of the differential expression analysis to the 44 tumor-normal pairs in the TCGA validation set yielded similar results (data not shown).

4 | DISCUSSION

There is increasing evidence that aberrant expression of lncRNA plays a role in the genesis and progression of HNSCC.[24,25] Through our present study, we were able to identify and validate 5 differentially expressed lncRNA in HPV-negative HNSCC (*SPRY4-IT1*, *HEIH*, *LUCAT1*, *LINC00152*, and *HAND2-AS1*). Further, we found that low expression of *MEG3* was associated with more favorable 3-

TABLE 3 Association between expression of significantly differential long non-coding RNA and patients characteristics and clinical features in The Cancer Genome Atlas human papillomavirus-negative head and neck squamous cell carcinoma samples (n = 444)

	SPRY4-IT1		*HEIH*		*LUCAT1*		*LINC00152*		*HAND2-AS1*		*MEG3*		*TERC*	
N = 444	**Low (n = 89)**	**High (n = 0)**	**Low (n = 92)**	**High (n = 2)**	**Low (n = 19)**	**High (n = 291)**	**Low (n = 6)**	**High (n = 291)**	**Low (n = 275)**	**High (n = 43)**	**Low (n = 134)**	**High (n = 107)**	**Low (n = 189)**	**High (n = 99)**
Age, per decade	1.08 (0.87-1.34)	—	**0.75 (0.60-0.94)**	—	1.11 (0.65-1.88)	0.99 (0.82-1.21)	—	1.09 (0.83-1.43)	0.87 (0.70-1.08)	**0.67 (0.48-0.92)**	0.84 (0.68-1.04)	0.93 (0.74-1.16)	1.05 (0.86-1.28)	**1.40 (1.08-1.81)**
Sex														
Male	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***
Female	0.72 (0.41-1.26)	—	0.56 (0.30-1.06)	—	0.60 (0.15-2.44)	1.07 (0.64-1.78)	—	0.90 (0.46-1.77)	0.60 (0.35-1.01)	0.62 (0.26-1.47)	1.45 (0.84-2.50)	1.00 (0.55-1.81)	1.36 (0.81-2.30)	0.65 (0.33-1.28)
Race														
White	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***
Black	1.26 (0.49-3.23)	—	1.15 (0.51-2.58)	—	0.49 (0.06-4.40)	0.74 (0.36-1.50)	—	0.94 (0.35-2.54)	1.16 (0.51-2.64)	1.57 (0.47-5.30)	**3.34 (1.43-7.80)**	**2.85 (1.18-6.90)**	0.90 (0.41-1.99)	1.05 (0.43-2.57)
Other	0.97 (0.37-2.58)	—	**3.50 (1.47-8.35)**	—	0.88 (0.09-8.95)	1.41 (0.54-3.68)	—	0.81 (0.28-2.38)	0.90 (0.37-2.23)	—	0.74 (0.29-1.86)	0.41 (0.13-1.30)	0.85 (0.35-2.09)	1.00 (0.34-2.98)
Smoking status														
Never	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***
Former	1.27 (0.65-2.45)	—	1.12 (0.53-2.38)	—	0.73 (0.19-2.83)	0.67 (0.36-1.26)	—	0.60 (0.25-1.43)	0.59 (0.45-1.72)	0.54 (0.20-1.47)	1.64 (0.84-3.22)	0.68 (0.35-1.34)	1.57 (0.84-2.90)	1.59 (0.73-3.50)
Current	1.59 (0.77-3.30)	—	1.37 (0.65-2.91)	—	0.39 (0.07-2.20)	0.79 (0.40-1.54)	—	0.78 (0.30-2.00)	0.59 (0.29-1.20)	0.45 (0.16-1.24)	0.90 (0.44-1.85)	0.53 (0.26-1.08)	1.48 (0.77-2.84)	1.30 (0.56-3.03)
Primary tumor site														
Oral cavity	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***
Pharynx	1.34 (0.52-3.48)	—	0.53 (0.19-1.47)	—	3.20 (0.83-12.40)	0.61 (0.28-1.29)	—	**0.36 (0.14-0.92)**	1.08 (0.49-2.40)	0.52 (0.13-2.09)	1.29 (0.58-2.90)	0.92 (0.36-2.32)	1.15 (0.50-2.64)	1.73 (0.67-4.49)
Larynx	0.82 (0.43-1.54)	—	0.93 (0.50-1.74)	—	1.25 (0.31-5.08)	0.96 (0.55-1.66)	—	0.53 (0.26-1.06)	1.33 (0.74-2.41)	0.48 (0.16-1.47)	0.71 (0.39-1.30)	0.86 (0.46-1.62)	0.56 (0.31-1.01)	1.49 (0.77-2.88)
AJCC Stage Group														
Early (I or II)	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***	***Reference***
Late (III or IV)	1.09 (0.52-2.28)	—	0.62 (0.28-1.37)	—	2.06 (0.42-10.13)	1.48 (0.78-2.82)	—	1.49 (0.65-3.42)	**2.86 (0.74-2.41)**	0.75 (0.20-2.81)	1.11 (0.56-2.18)	1.01 (0.48-2.14)	1.84 (0.94-3.60)	**0.39 (0.17-0.91)**

(Continues)

TABLE 3 (Continued)

	SPRY4-IT1		*HEIH*		*LUCAT1*		*LINC00152*		*HAND2-AS1*		*MEG3*		*TERC*	
N = 444	**Low (n = 89)**	**High (n = 0)**	**Low (n = 92)**	**High (n = 2)**	**Low (n = 19)**	**High (n = 291)**	**Low (n = 6)**	**High (n = 291)**	**Low (n = 275)**	**High (n = 43)**	**Low (n = 134)**	**High (n = 107)**	**Low (n = 189)**	**High (n = 99)**
N classification														
N0	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
N1-3	0.88 (0.47-1.65)	—	1.65 (0.86-3.19)	—	0.47 (0.13-1.71)	0.82 (0.47-1.44)	—	1.25 (0.59-2.63)	**0.28 (0.15-0.54)**	1.31 (0.39-4.44)	**0.49 (0.28-0.88)**	0.72 (0.38-1.34)	0.97 (0.56-1.69)	**2.68 (1.26-5.69)**
Tumor grade														
Low	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Moderate	1.39 (0.66-2.92)	—	1.22 (0.56-2.64)	—	1.87 (0.20-17.04)	1.38 (0.73-2.58)	—	1.34 (0.58-3.13)	1.07 (0.52-2.19)	0.74 (0.25-2.19)	0.83 (0.41-1.66)	1.01 (0.48-2.19)	0.94 (0.48-1.85)	1.17 (0.49-2.81)
High/ undifferentiated	0.64 (0.28-1.46)	—	0.67 (0.26-1.72)	—	7013 (0.78-64.74)	1.22 (0.58-2.58)	—	1.21 (0.45-3.28)	0.87 (0.38-1.96)	0.94 (0.28-3.21)	0.65 (0.29-1.47)	1.33 (0.31-1.85)	0.62 (0.28-1.35)	0.71 (0.26-1.90)

Abbreviation: AJCC, American Joint Committee on Cancer.

Figures in boldface indicate statistical significance.

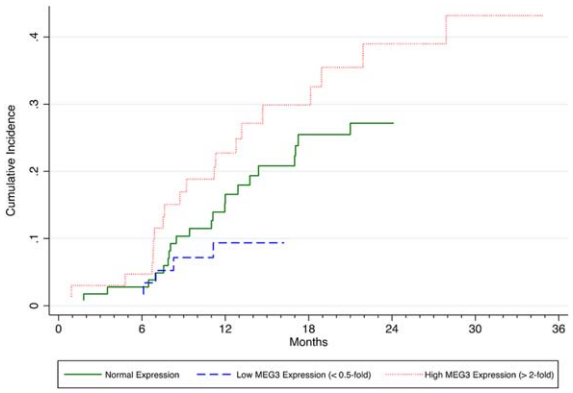

FIGURE 2 Cumulative incidence function for 3-year relapse-free survival of HPV-negative The Cancer Genome Atlas head and neck squamous cell carcinoma tumors according to *MEG3* expression level [Color figure can be viewed at wileyonlinelibrary.com]

year RFS, although the significance of this finding is unclear.

MEG3 is a lncRNA located at the *DLK-MEG3* locus on chromosome 14q32.3 and is reported to be a tumor suppressor gene, exerting its effect in part through interaction with tumor suppressor and master regulator p53.[26,27] *MEG3* has been reported to be downregulated in multiple solid tumor types,[28] which is consistent with what we observed in our 19 HPV-negative HNSCC tumor-adjacent normal pairs. However, our observed association of low *MEG3* expression with better 3-year RFS is contrary to what has been reported for other solid tumor types[29] and, therefore, should be interpreted with caution. In particular, Jia and colleagues[30] found that low expression of *MEG3* correlated with poorer outcomes in squamous cell carcinoma (SCC) of the anterior tongue, and overexpression of *MEG3* inhibited cell proliferation and cell cycle progression in SCC-15 and Cal 27 tongue SCC lines.

Although the other differentially expressed lncRNA transcripts in our study were not associated with OS or RFS in the TCGA head and neck tumors, they have been identified as potential markers for poor prognosis in other cancer types. *SPRY4-IT1*, which is located on the *SPRY4* gene, has been implicated in cell growth invasion and increased apoptosis,[31] and elevated expression of *SRPY4-IT1* has been associated with poorer outcomes in esophageal SCC, which, like HPV-negative HNSCC, is also strongly associated with tobacco and alcohol.[32] Overexpression of *HEIH* has been reported in hepatocellular carcinoma, in which it is an independent predictor for recurrence and survival and interacts with the lysine methyltransferase and Polycomb Repressor Complex 2 member *Enhancer of Zeste Homolog 2* (*EZH2*).[33] Upregulation of both *LUCAT1* and *LINC00152* have been associated with poorer outcomes for non-small cell lung cancer,[34,35] and *HAND2-AS1* has been reported to be upregulated in

TABLE 4 Crude and adjusted hazard ratios and subhazard ratios for the association between expression of differentially expressed long non-coding RNA and 5-year overall survival and 3-year recurrence-free survival, respectively, in patients with HPV-negative head and neck squamous cell carcinoma

	5-year Overall Survival (OS) (N = 439)			3-year Relapse-Free Survival (RFS) (N = 272)		
lncRNA expression	**n**	**Crude HR**	**Adjusted HR[a]**	**n**	**Crude SHR**	**Adjusted SHR[a]**
SPRY4-IT1						
Low (<0.5-fold)	0	—	—	0	—	—
Normal[b]	87	1.00 (reference)	1.00 (reference)	54	1.00 (reference)	1.00 (reference)
High (>2-fold)	352	0.93 (0.62-1.38)	0.89 (0.60-1.34)	218	1.40 (0.63-3.12)	1.30 (0.54-3.13)
HEIH						
Low (<0.5-fold)	92	0.82 (0.54-1.28)	0.88 (0.56-1.36)	59	0.91 (0.45-1.84)	1.14 (0.54-2.40)
Normal[b]	345	1.00 (reference)	1.00 (reference)	212	1.00 (reference)	1.00 (reference)
High (>2-fold)	2	—	—	1	—	—
LUCAT1						
Low (<0.5-fold)	18	1.49 (0.67-3.32)	1.61 (0.71-3.68)	12	3.07 (0.84-11.19)	2.58 (0.71-9.33)
Normal[b]	131	1.00 (reference)	1.00 (reference)	80	1.00 (reference)	1.00 (reference)
High (>2-fold)	290	0.92 (0.64-1.31)	0.85 (0.60-1.23)	180	1.50 (0.74-3.02)	1.54 (0.70-3.41)
LINC00152						
Low (<0.5-fold)	8	—	—	7	—	—
Normal[b]	68	1.00 (reference)	1.00 (reference)	41	1.00 (reference)	1.00 (reference)
High (>2-fold)	363	0.84 (0.55-1.28)	0.74 (0.48-1.14)	224	1.32 (0.52-3.34)	1.11 (0.41-3.04)
HAND2-AS1						
Low (<0.5-fold)	271	0.92 (0.64-1.31)	0.96 (0.66-1.39)	167	0.87 (0.47-1.64)	0.83 (0.43-1.60)
Normal[b]	126	1.00 (reference)	1.00 (reference)	80	1.00 (reference)	1.00 (reference)
High (>2-fold)	42	0.63 (0.30-1.34)	0.71 (0.33-1.51)	25	0.94 (0.31-2.83)	0.82 (0.25-2.65)
MEG3						
Low (<0.5-fold)	133	1.09 (0.75-1.57)	1.11 (0.76-1.62)	71	**0.38 (0.14-1.00)**	**0.28 (0.10-0.78)**
Normal[b]	202	1.00 (reference)	1.00 (reference)	127	1.00 (reference)	1.00 (reference)
High (>2-fold)	104	0.88 (0.56-1.40)	0.84 (0.53-1.34)	74	1.63 (0.89-2.98)	1.43 (0.75-2.74)
TERC						
Low (<0.5-fold)	187	1.01 (0.69-1.47)	0.90 (0.61-1.33)	115	1.25 (0.62-2.51)	1.30 (0.61-2.79)
Normal[b]	155	1.00 (reference)	1.00 (reference)	97	1.00 (reference)	1.00 (reference)
High (>2-fold)	97	0.98 (0.62-1.54)	0.97 (0.60-1.55)	60	1.75 (0.82-3.72)	2.24 (0.95-5.25)

[a]Adjusted for age, sex, race/ethnicity, smoking status, primary tumor site, and AJCC stage group.
[b]Tumor expression >0.5-fold and < 2-fold was considered to be within normal range.
Abbreviations: HR, hazard ratio; lncRNA, long non-coding RNA; SHR, subhazard ratio.
Figures in boldface indicate statistical significance.

stage IV-S neuroblastoma[36] but downregulated in metastatic hepatocellular carcinoma,[37] with the latter being more in-line with our narrative of lower expression of *HAND2-AS1* in HPV-negative HNSCC.

Strengths of our study include the comprehensive assessment of 84 cancer-associated lncRNA and our ability to access raw RNA-seq data from TCGA. This allowed the alignment of our data with annotated lncRNA sequences to validate our findings in an independent dataset and assess the impact of significantly differential lncRNA on OS and RFS, and to infer HPV status by aligning to HPV16 E6/E7 viral mRNA transcripts. Our study also has several limitations, including the modest sample size of our initial discovery set, which may have adversely impacted our power to detect smaller effect sizes, increasing the risk of false-negative results. However, our use of a stringent false discovery rate-control and validation in an independent set of tumors using TCGA data yields high confidence in our significant results. Furthermore, because no adjustments were made for multiple comparisons in the survival analyses, we cannot rule out the

spurious nature of the observed association between *MEG3* and RFS. Additionally, use of archival formalin-fixed paraffin-embedded tissue for our discovery set likely attenuated the lncRNA expression levels, which could reduce our sensitivity for detection of signal or more subtle differences in expression. Our use of the RNeasy formalin-fixed paraffin-embedded kit, which is specifically engineered to maximize the integrity and downstream results for RNA extracted from formalin-fixed paraffin-embedded,[38] helps mitigate this issue; it is also notable that 5 of the 7 differentially expressed lncRNA identified with the array were replicated using RNA-seq data from fresh tissue, supporting the validity of our findings.

5 | CONCLUSION

Expression of lncRNA is dysregulated in HPV-negative HNSCC. Specifically, we have identified and validated 5 differentially expressed lncRNA: *SPRY4-IT1*, *HEIH*, *LUCAT1*, *LINC00152*, and *HAND2-AS1*. Additional studies are needed to confirm the potential prognostic value of *MEG3* expression in HNSCC.

ORCID

Sulsal-Ul Haque http://orcid.org/0000-0002-1145-621X

REFERENCES

[1] Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin.* 2005;55(2):74-108.

[2] Parkin DM, Muir CS. Cancer incidence in five continents. Comparability and quality of data. *IARC Sci Publ.* 1992;(120):45-173.

[3] Brockstein B, Haraf DJ, Rademaker AW, et al. Patterns of failure, prognostic factors and survival in locoregionally advanced head and neck cancer treated with concomitant chemoradiotherapy: a 9-year, 337-patient, multi-institutional experience. *Ann Oncol.* 2004;15(8):1179-1186.

[4] Djebali S, Davis CA, Merkel A, et al. Landscape of transcription in human cells. *Nature.* 2012;489(7414):101-108.

[5] Pauli A, Rinn JL, Schier AF. Non-coding RNAs as regulators of embryogenesis. *Nat Rev Genet.* 2011;12(2):136-149.

[6] Chandra Gupta S, Nandan Tripathi Y. Potential of long non-coding RNAs in cancer patients: from biomarkers to therapeutic targets. *Int J Cancer.* 2017;140(9):1955-1967.

[7] Cech TR, Steitz JA. The noncoding RNA revolution-trashing old rules to forge new ones. *Cell.* 2014;157(1):77-94.

[8] Yang G, Lu X, Yuan L. LncRNA: a link between RNA and cancer. *Biochim Biophys Acta.* 2014;1839(11):1097-1109.

[9] Serviss JT, Johnsson P, Grandér D. An emerging role for long non-coding RNAs in cancer metastasis. *Front Genet.* 2014;5:234.

[10] Qiu MT, Hu JW, Yin R, Xu L. Long noncoding RNA: an emerging paradigm of cancer research. *Tumour Biol.* 2013;34(2):613-620.

[11] Gutschner T, Hämmerle M, Eissmann M, et al. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res.* 2013;73(3):1180-1189.

[12] Schmidt LH, Spieker T, Koschmieder S, et al. The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth. *J Thorac Oncol.* 2011;6(12):1984-1992.

[13] Liu XH, Liu ZL, Sun M, Liu J, Wang ZX, De W. The long non-coding RNA HOTAIR indicates a poor prognosis and promotes metastasis in non-small cell lung cancer. *BMC Cancer.* 2013;13:464.

[14] Nakagawa T, Endo H, Yokoyama M, et al. Large noncoding RNA HOTAIR enhances aggressive biological behavior and is associated with short disease-free survival in human non-small cell lung cancer. *Biochem Biophys Res Commun.* 2013;436(2):319-324.

[15] Zhao W, An Y, Liang Y, Xie XW. Role of HOTAIR long noncoding RNA in metastatic progression of lung cancer. *Eur Rev Med Pharmacol Sci.* 2014;18(13):1930-1936.

[16] Zhang J, Zhang P, Wang L, Piao HL, Ma L. Long non-coding RNA HOTAIR in carcinogenesis and metastasis. *Acta Biochim Biophys Sin (Shanghai).* 2014;46(1):1-5.

[17] Gutschner T, Hämmerle M, Diederichs S. MALAT1 – a paradigm for long noncoding RNA function in cancer. *J Mol Med (Berl).* 2013;91(7):791-801.

[18] Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met* 1995;57(1):289-300.

[19] Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10(3):R25.

[20] Grossman RL, Heath AP, Ferretti V, et al. Toward a shared vision for cancer genomic data. *N Engl J Med.* 2016;375(12):1109-1112.

[21] Colaprico A, Silva TC, Olsen C, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 2016;44(8):e71.

[22] Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26(1):139-140.

[23] Kim HT. Cumulative incidence in competing risks data and competing risks regression analysis. *Clin Cancer Res.* 2007;13(2 Pt 1):559-565.

[24] Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. *Br J Cancer.* 2004;91(7):1229-1235.

[25] Nohata N, Abba MC, Gutkind JS. Unraveling the oral cancer lncRNAome: identification of novel lncRNAs associated with malignant progression and HPV infection. *Oral Oncol.* 2016;59:58-66.

[26] Zou AE, Zheng H, Saad MA, et al. The non-coding landscape of head and neck squamous cell carcinoma. *Oncotarget.* 2016;7(32):51211-51222.

[27] Tang W, Dong K, Li K, Dong R, Zheng S. MEG3, HCN3 and linc01105 influence the proliferation and apoptosis of neuroblastoma cells via the HIF-1α and p53 pathways. *Sci Rep.* 2016;6:36268.

[28] Zhou Y, Zhong Y, Wang Y, et al. Activation of p53 by MEG3 non-coding RNA. *J Biol Chem.* 2007;282(34):24731-24742.

[29] Zhou Y, Zhang X, Klibanski A. MEG3 noncoding RNA: a tumor suppressor. *J Mol Endocrinol.* 2012;48(3):R45-R53.

[30] Cui X, Jing X, Long C, Tian J, Zhu J. Long noncoding RNA MEG3, a potential novel biomarker to predict the clinical outcome of cancer patients: a meta-analysis. *Oncotarget.* 2017;8(12):19049-19056.

[31] Jia LF, Wei SB, Gan YH, et al. Expression, regulation and roles of miR-26a and MEG3 in tongue squamous cell carcinoma. *Int J Cancer.* 2014;135(10):2282-2293.

[32] Zou Y, Jiang Z, Yu X, et al. Upregulation of long noncoding RNA SPRY4-IT1 modulates proliferation, migration, apoptosis, and network formation in trophoblast cells HTR-8SV/neo. *PLoS One.* 2013;8(11):e79598.

[33] Zhang CY, Li RK, Qi Y, et al. Upregulation of long noncoding RNA SPRY4-IT1 promotes metastasis of esophageal squamous cell carcinoma via induction of epithelial-mesenchymal transition. *Cell Biol Toxicol.* 2016;32(5):391-401.

[34] Yang F, Zhang L, Huo XS, et al. Long noncoding RNA high expression in hepatocellular carcinoma facilitates tumor growth through enhancer of zeste homolog 2 in humans. *Hepatology.* 2011;54(5):1679-1689.

[35] Renhua G, Yue S, Shidai J, Jing F, Xiyi L. 165P: Long noncoding RNA LUCAT1 is associated with poor prognosis in human non-small cell lung cancer and affects cell proliferation via regulating p21 and p57 expression. *J Thoracic Oncol.* 2016;11(4 Suppl):S129.

[36] Chen QN, Chen X, Chen ZY, et al. Long intergenic non-coding RNA 00152 promotes lung adenocarcinoma proliferation via interacting with EZH2 and repressing IL24 expression. *Mol Cancer.* 2017;16(1):17.

[37] Voth H, Oberthuer A, Simon T, Kahlert Y, Berthold F, Fischer M. Identification of DEIN, a novel gene with high expression levels in stage IVS neuroblastoma. *Mol Cancer Res.* 2007;5(12):1276-1284.

[38] Yang Y, Chen L, Gu J, et al. Recurrently deregulated lncRNAs in hepatocellular carcinoma. *Nat Commun.* 2017;8:14421.

[39] Belder N, Coskun Ö, Doganay Erdogan B, et al. From RNA isolation to microarray analysis: comparison of methods in FFPE tissues. *Pathol Res Pract.* 2016;212(8):678-685.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Haque S-U, Niu L, Kuhnell D, et al. Differential expression and prognostic value of long non-coding RNA in HPV-negative head and neck squamous cell carcinoma. *Head & Neck.* 2018;40:1555–1564. https://doi.org/10.1002/hed.25136