Comprehensive Analysis of Sinonasal Inverted Papilloma Expression Profiles Identified Long Noncoding RNA AKTIP as a Potential Biomarker

Hanyi He
Affiliated Hangzhou First People's Hospital Zhejiang University School of Medicine: Hangzhou First People's Hospital

Xinlu Wang
Zhejiang Chinese Medical University

Yueyue Lu
Zhejiang Chinese Medical University

Xiaojiang Lin
Kaihua People's Hospitial

Yuandong Li
Hangzhou Children's Hospital

Yong Li
Affiliated Hangzhou First People's Hospital Zhejiang University School of Medicine: Hangzhou First People's Hospital

Zhihong Lin
Zhejiang University School of Medicine

Zhiqi Ma
Hangzhou First People's Hospital

Xiaolin Cao
Hangzhou First People's Hospital

Yaoshu Teng (✉ tengyaoshu@163.com )
Hangzhou First People's Hospital  https://orcid.org/0000-0003-2335-1594

Primary research

Keywords: Inverted papilloma, Long noncoding RNA, Expression profiles, Bioinformatics, Biomarker

Posted Date: October 21st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-962679/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

**Background:** Long noncoding RNAs (lncRNAs) are a novel class of potential biomarkers and therapeutic targets for the treatment of neoplasms. The purpose of this study was to explore the expression profile, potential functions, and diagnostic and clinicopathological significance of lncRNAs in sinonasal inverted papilloma (SNIP).

**Methods** The expression profiles of lncRNAs and mRNAs were analyzed using a microarray. The potential functions and clinical implications of specific lncRNAs were further analyzed by bioinformatics and statistical methods.

**Results** Microarray analysis identified 1,668 significantly upregulated and 1,767 downregulated lncRNAs in SNIP. Several mRNAs coexpressed with lncRNAs were enriched in some biological processes and cellular signaling pathways related to tumorigenesis. Lnc-AKTIP might interact with a variety of tumor-associated proteins and transcription factors, such as PCBP2, IRF-1, and p53. Receiver operating characteristic curve analysis for lnc-AKTIP showed an area under the curve of 0.939. Notably, its expression level was significantly decreased in SNIP tissues versus normal tissues and was associated with SNIP staging.

**Conclusion** Lnc-AKTIP may serve as a valuable diagnostic biomarker and a therapeutic target for SNIP.

Introduction

Sinonasal inverted papilloma (SNIP) is a challenging benign tumor arising from the Schneiderian mucosa of the nasal cavity and paranasal sinuses, accounting for approximately 0.5% ~ 4% of all sinonasal neoplasms [1]. It has the biological characteristics of local invasiveness, a high recurrence rate and malignant potential [2]. Malignant transformation has been found in 5–15% of inverted papilloma lesions [3]. Human papillomavirus (HPV) is considered to be closely related to the pathogenesis of SNIP. The other considered risk factors for SNIP development include inflammatory infiltration, welding fumes and organic solvents [4–6]. However, little is known about the underlying molecular genetic alterations, specific pathologic mechanism and diagnostic biomarkers of this clinical entity.

Long noncoding RNAs (lncRNAs) are defined as noncoding RNAs (ncRNAs) that are greater than 200 nucleotides in length [7]. Compared with other ncRNAs (e.g., miRNAs), lncRNAs have a longer primary structure, can integrate with DNAs and RNAs, and can form a complex and diverse secondary spatial structure to interact with proteins [8]. Although lncRNAs do not encode proteins, some studies have shown that lncRNAs play crucial roles in governing a wide range of fundamental biological processes, including genomic imprinting, chromosome inactivation, differentiation and carcinogenesis, at both the transcriptional and post-transcriptional levels. Aberrant expression of lncRNAs is associated with several human diseases, such as various types of tumors, cardiovascular diseases, and neurological diseases [7, 9, 10]. Moreover, these abnormal lncRNAs are also found in circulating blood and/or urine [11, 12]. LncRNAs are a novel class of potential biomarkers and therapeutic targets for the treatment of neoplasms [13–15].

In our study, differentially expressed lncRNAs and mRNAs were comprehensively identified by detecting lncRNA and mRNA profiles in SNIP tissues. The functional enrichment analysis of mRNAs coexpressed with these lncRNAs was performed via bioinformatics methods. These findings were combined with the clinicopathological features of patients with SNIP to explore the clinical significance of specific lncRNAs in SNIP. This study aimed to provide novel information for further research on the pathogenesis of SNIP and to identify candidate diagnostic biomarkers and therapeutic targets.

Methods

**Patients & sample collection**

SNIP tissue samples were obtained from 41 patients with SNIP (29 males and 12 females; mean age 57.6 years; range: 32–85 years). The diagnosis of SNIP was confirmed by histopathological examination. The clinicopathological characteristics (age, Gender, smoking status, tumor staging and recurrence) of each SNIP patient were recorded (Table 1). A total of 12 patients with only nasal septum deviation were selected to provide nasal mucosal tissue samples (nine males and three females; mean age 51.2 years; range: 28–64 years as the control group). Patients were admitted to the Department of Otolaryngology, Affiliated Hangzhou
First People's Hospital, Zhejiang University School of Medicine and Department of Otolaryngology, Second Affiliated Hospital, School of Medicine, Zhejiang University between 2012 and 2020. All tissue samples were immediately preserved in RNA later Solution (Ambion, TX, USA) within 15 minutes after resection and then stored at -20°C until use. This study complied with the Declaration of Helsinki and was approved by the local ethics committee. Written informed consent was obtained at enrollment from each participating subject.

| Characteristic         | n  |
|------------------------|----|
| Age                    |    |
| <50 years              | 14 |
| ≥50 years              | 27 |
| Gender                 |    |
| Male                   | 29 |
| Female                 | 12 |
| Smoking status         |    |
| Yes                    | 26 |
| No                     | 15 |
| Tumor staging<sup>a</sup> |    |
| I                      | 8  |
| II                     | 20 |
| III                    | 11 |
| IV                     | 2  |
| V                      | 28 |
| V+II                   | 13 |
| Recurrence             |    |
| Yes                    | 6  |
| No                     | 35 |

<sup>a</sup> Tumor staging is defined according to a staging system for inverted papilloma described by Krouse JH (2000).

**RNA extraction**

Total RNA was isolated from nasal mucosal tissues using TRIzol Reagent (Life Technologies, Carlsbad, CA, US) following the manufacturer's instructions and then quantified using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). RNA integrity was inspected by an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, US), and RNA samples with an RNA integrity number (RIN) values ≥6.0 and 28S/18S values >0.7 were deemed acceptable for microarray and reverse transcription (RT) experiments.

**Microarray assay**

A human IncRNA microarray (4x180K; v 6.0) was manufactured at Shanghai Biotechnology Corporation (Shanghai, China), which contains 95,956 capture probes for 77,103 IncRNAs and 18,853 RNAs based on the most authoritative databases, such as
GENCODE v21, Ensembl, LNCipedia v3.1, Lncrnadb, Noncode v4 and UCSC. Microarray assays were performed according to the manufacturer's instructions. Briefly, total RNA was amplified and labeled by the Low Input Quick Amp WT Labeling Kit (Agilent Technologies, Santa Clara, CA, US). Labeled cRNA was purified with the RNeasy mini kit (QIAGEN, GmBH, Germany). Each slide was hybridized with 1.65 µg Cy3-labeled cRNA using a Gene Expression Hybridization Kit (Agilent Technologies, Santa Clara, CA, US). After 17 hours of hybridization, the slides were washed with the Gene Expression Wash Buffer Kit (Agilent Technologies, Santa Clara, CA, US). These slides were scanned by an Agilent Microarray Scanner (Agilent Technologies, Santa Clara, CA, US). Data were extracted with Feature Extraction software10.7 (Agilent Technologies, Santa Clara, CA, US). Raw data were normalized by the quantile algorithm, Gene Spring Software 11.0 (Agilent Technologies, Santa Clara, CA, US). Differentially expressed IncRNAs and mRNAs with statistical significance between two groups were identified through volcano plot filtering (fold change ≥2.0 and P<0.05). The Gene Cluster (v 3.0) and Java TreeView software programs were used to perform hierarchical cluster analysis of these differentially expressed IncRNAs and mRNAs.

**Quantitative real-time reverse transcription PCR (qRT-PCR)**

Total RNA was reverse transcribed to cDNA using Prime Script RT Master Mix (TaKaRa, Dalian, China) following the manufacturer’s protocols. qRT-PCR was performed by using SYBR Premix Ex Taq II (TaKaRa, Dalian, China) on the 7900 HT Sequence Detection System (ABI, USA). The primer sequences are listed in Table 2 and were synthesized by Invitrogen (Shanghai, China). Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as an endogenous control, and gene expression was compared by the threshold cycle (2^{-ΔΔCt}) method.

| IncRNA/mRNA       | Forward primer (5’ to 3’)          | Reverse primer (5’ to 3’)          | Product(bp) |
|-------------------|------------------------------------|------------------------------------|-------------|
| Inc-SERPINB3-4:1  | GACATACAGAGTGGGTTGCG               | GCTCTGAGGTACAGTGTGCTGA             | 91          |
| Inc-AZIN1-1:5     | ACGGCGAATTTTCTGACCA                | CCCTGGAAGTAGAACAGGGA               | 163         |
| NR_029957         | ATGGGCCGTCTTACCAGACAT              | GCGGATGGACGGTTTTACCA               | 71          |
| Inc-GNG5P2-2:2    | CTGCTTTTGTCTTGATGTGCT              | GCATGTTGCTTTGGGTAACAA              | 112         |
| Inc-AKTIP-5:1     | AGGAACACAGGAAAAAGTGCG              | GCCTATCACAGACTGCTTCA               | 100         |
| NR_024061         | TGGGAGTTTACGACAGGACG               | GCGCCTAGCAAAAAAGGTTTC             | 158         |
| Inc-MUTED-2:4     | TGGACACATCATGTCATATGGCA            | CGGCTTCAGCAGTGGTGGTGC             | 187         |
| Inc-CRLF1-1:1     | GAGTGGATTTTTCGCGGCG                | GCACCGAGTAAGGAGGATT                | 90          |
| COX6B2            | AGCCACGAAATAGCCACCA                | CCTGGGGTCACTGAGTTGCC              | 160         |
| COL12A1           | ATCCAGGTTCCGGGTACAC                | TCTTTGTAGTGGGTAACCG               | 168         |
| RARRES2           | TGCCCCATAGAGACCCAGA                | GAAGTGAAGCTGTTGGGGT               | 102         |
| GAPDH             | GAATGGGCACAGGTTAGGG                | AAAACATCACCAGGAGG                 | 134         |

**Construction of IncRNA-mRNA coexpression network**

Pearson correlation coefficient (PCC) R values were calculated to evaluate the correlation between the differentially expressed IncRNAs and mRNAs. Fisher's exact test implemented in the WGCNA library of R was adopted to estimate p-values for each correction pair. The p-value was further adjusted to the false discovery rate (FDR) by Bonferroni multiple test correction. The IncRNA-mRNA correlations with an R value ≥0.8 and FDR <0.005 were considered statistically significant correlated pairs. The coexpression network showing the significant pairs was visualized by using Cytoscape software (The Cytoscape Consortium, San Diego, CA, USA).

**Gene function analysis**
The coexpressed mRNAs were imported into the Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8 (http://david.abcc.ncifcrf.gov), which utilized Gene Ontology (GO) and pathway analysis to identify the enriched GO themes and cell signaling pathways of these coexpressed mRNAs. The thresholds were set as $P<0.05$ and FDR<0.05.

**Bioinformatics analysis of specific lncRNAs**

The genomic locations of candidate lncRNAs were confirmed by using UCSC Genome Browser (http://genomeasia.ucsc.edu/index.html). Secondary structures were shown via RNAfold minimum free energy estimations based on the RNAfold web server (http://ma.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). catRAPID analysis (http://service.tartaglialab.com/) was performed to predict the potential interacting proteins of lncRNAs. Moreover, TRANSFAC (http://www.gene-regulation.com/index2.html) was used to predict the potential transcription factors of lncRNAs.

**Statistical analysis**

Statistical analyses were performed using SPSS software for Windows (version 16.0; SPSS, Inc., Chicago, IL, USA). The differences between the two groups were determined using a two-tailed Student's $t$-test. Pearson correlation analysis was performed to investigate the linear relationship between the microarray data and qRT-PCR results. A receiver operating characteristic (ROC) curve was established to evaluate the diagnostic accuracy of lncRNAs as biomarkers of SNIP. The chi-square test was used to determine the relationships between clinicopathological characteristics and altered lncRNA expression. $P<0.05$ was considered statistically significant.

**Results**

**Overview of the expression profiles of lncRNAs and mRNAs in SNIP tissues**

To identify the differentially expressed lncRNAs and mRNAs potentially involved in SNIP, we first examined the expression patterns of lncRNAs and mRNAs in the SNIP tissues using a microarray assay. Our data showed that a total of 3,435 lncRNAs (1,668 upregulated and 1,767 downregulated) and 4,696 mRNAs (2,424 upregulated and 2,272 downregulated) were significantly differentially expressed in the SNIP tissues compared with the corresponding nontumorous tissues from the control group (fold change>2; $P<0.05$), as indicated by the violin plots and heat maps (Fig. 1). The top 20 differentially expressed lncRNAs and mRNAs are listed in Tables 3 and 4. Among these lncRNAs, Inc-SPRR1B-1:1 (log$_2$ fold change: 8.074459) and Inc-PRH1-1:13 (log$_2$ fold change: -8.09616) were the most upregulated and downregulated lncRNAs, respectively. In addition, A2ML1 (log$_2$ fold change: 8.517957) and DMBT1 (log$_2$ fold change: -12.915) were the most upregulated and downregulated mRNAs, respectively. Our results indicate that these 4 aberrantly expressed RNAs may play critical roles in the development and progression of SNIP.
Table 3

Top 20 differentially expressed lncRNAs in SNIP tissues detected by microarray assay

| Up-regulated lncRNAs | Down-regulated lncRNAs |
|----------------------|------------------------|
|                      |                        |                      |                        |
| lncRNA               | Chromosome             | \( \log_2 (\text{Fold change}) \) | P value | lncRNA               | Chromosome             | \( \log_2 (\text{Fold change}) \) | P value |
|                      |                        |                        |         |                      |                        |                        |         |
| Inc-SPRR1B-1:1       | chr1                   | 8.074459               | 0.000516 | Inc-PRH1-1:13        | chr12                  | -8.09616               | 0.006248 |
| NR_003062            | chr1                   | 7.344342               | 0.000166 | Inc-PRH1-1:13        | chr12                  | -8.09616               | 0.006248 |
| ENST00000411759      | chr17                  | 6.782256               | 0.001133 | ENST00000414404      | chr2                   | -6.66018               | 2.70E-05 |
| Inc-GNRRH-5:2        | chr4                   | 6.097429               | 0.0053   | NR_036489            | chr18                  | -5.97944               | 4.70E-05 |
| NR_126404            | chr2                   | 5.899363               | 0.005206 | Inc-ZKSCAN1-3:2      | chr7                   | -5.77709               | 0.000537 |
| Inc-PROM2-1:2        | chr2                   | 5.808771               | 0.001056 | NR_027622            | chr3                   | -5.68696               | 8.08E-06 |
| NR_073414            | chr17                  | 5.606244               | 0.000893 | Inc-RP11-497E19.2.1-3:3 | chr14                  | -5.523                 | 0.00147   |
| NR_104048            | chr4                   | 5.282333               | 0.001278 | ENST00000624094      | chr11                  | -5.48935               | 1.14E-05 |
| NR_073414            | chr17                  | 5.234484               | 0.001084 | Inc-BHLHA15-1:1      | chr7                   | -5.37011               | 0.000427 |
| NR_027054            | chr9                   | 5.038543               | 2.05E-06 | Inc-LRRC10-1:2       | chr12                  | -5.08804               | 6.40E-05 |
| Inc-IL20RB-2:1       | chr3                   | 5.021085               | 0.002429 | ENST00000623553      | chr16                  | -5.02176               | 8.71E-06 |
| NR_120497            | chr10                  | 4.946139               | 0.00024  | Inc-MYL9-1:1         | chr20                  | -5.00485               | 0.000938 |
| Inc-SERPINEB3-4:1    | chr18                  | 4.919234               | 0.0121   | Inc-GMPPA-2:1        | chr2                   | -5.00169               | 0.000102 |
| Inc-APLP2-5:2        | chr11                  | 4.889049               | 0.000642 | Inc-ANKRD22-1:2      | chr10                  | -4.91236               | 0.0004   |
| ENST00000420269      | chr22                  | 4.862116               | 0.000494 | ENST00000612804      | chr19                  | -4.85893               | 0.019692 |
| ENST00000596379      | chr19                  | 4.839326               | 0.000487 | Inc-RP11-497E19.2.1-3:3 | chr14                  | -4.85061               | 0.00182   |
| NR_027054            | chr9                   | 4.814526               | 3.70E-05 | ENST00000602964      | chr5                   | -4.84958               | 0.000416 |
| NR_125989            | chr1                   | 4.749302               | 0.00103  | Inc-LUC7L-2:1        | chr16                  | -4.74143               | 0.010327 |
| ENST00000504297      | chr5                   | 4.692076               | 0.001742 | Inc-CXCL12-4:1       | chr10                  | -4.69443               | 5.81E-05 |
| ENST00000566876      | chr16                  | 4.588224               | 3.87E-05 | Inc-PTPLB-3:2        | chr3                   | -4.66377               | 2.11E-05 |
Validation of the candidate lncRNAs and mRNAs by qRT-PCR

To validate the microarray data, we performed qRT-PCR to confirm the expression levels of 8 lncRNAs and 3 mRNAs that were randomly selected from the differentially expressed RNAs detected by the microarray experiment. qRT-PCR was performed in 2 extended panels of the SNIP group (n=41) and control group (n=12). The qRT-PCR results for 6 lncRNAs (lnc-SERPINB3-4:1, NR_029957, lnc-GNG5P2-2:2, lnc-AKTIP-5:1, lnc-MUTED-2:4 and lnc-CRLF1-1:1) and 2 mRNAs (COX6B2 and COL12A1) were consistent with those from the microarray study, while the results for lnc-AZIN1-1:5, NR_024061 and RARRES2 were inconsistent, resulting in a concordance rate of 72.7% (8/11) (Fig. 2A). Pearson correlation analysis further showed a positive correlation between the microarray data and the qRT-PCR results (r=0.8208, P=0.002) (Fig. 2B).

LncRNA-mRNA coexpression network

To explore the potential interaction between the lncRNAs and mRNAs in the SNIP tissues, we analyzed the correlation between the top 400 differentially expressed lncRNAs and mRNAs from our microarray data by calculating R and FDR values. Based on an R value ≥ 0.8 and an FDR < 0.005, the lncRNA-mRNA coexpression network was constructed and visualized by using Cytoscape software. The network contained 305 network nodes, including 155 lncRNAs and 150 mRNAs, in which 670 significant correlation

| Up-regulated mRNAs | Down-regulated mRNAs |
|--------------------|-----------------------|
| **mRNAs** | **Chromosome** | **log₂ (Fold change)** | **P value** | **mRNAs** | **Chromosome** | **log₂ (Fold change)** | **P value** |
| A2ML1 | chr12 | 8.517957 | 0.00045 | DMBT1 | chr10 | -12.915 | 9.21E-10 |
| KRT6A | chr12 | 8.29795 | 0.00339 | LTF | chr3 | -9.64493 | 0.000496 |
| SPRR1B | chr1 | 8.009959 | 4.16E-05 | PRR4 | chr12 | -9.34335 | 0.003043 |
| KRT13 | chr17 | 7.873107 | 1.42E-05 | STATH | chr4 | -9.20928 | 0.001623 |
| CPA4 | chr7 | 7.634678 | 0.002453 | PRR4 | chr12 | -8.69048 | 0.001741 |
| KRT13 | chr17 | 7.544229 | 0.007242 | PRR4 | chr12 | -8.63995 | 0.001877 |
| KRT13 | chr17 | 7.361434 | 0.000629 | SCGB3A1 | chr5 | -8.34723 | 4.29E-06 |
| FAM83A | chr8 | 7.334105 | 0.000539 | HP | chr16 | -7.92955 | 4.30E-05 |
| SPRR2B | chr1 | 7.214323 | 0.001015 | PIP | chr7 | -7.83139 | 5.26E-05 |
| SPRR3 | chr1 | 7.085979 | 0.000475 | HPR | chr16 | -7.39776 | 1.51E-05 |
| S100A8 | chr1 | 6.956209 | 1.50E-05 | PI16 | chr6 | -7.12393 | 0.000269 |
| SPRR2D | chr1 | 6.954412 | 0.000155 | PLA2G2A | chr1 | -7.12383 | 0.000317 |
| CLCA4 | chr1 | 6.864341 | 0.000155 | CHRM3 | chr1 | -7.04896 | 1.40E-05 |
| CLCA2 | chr1 | 6.751895 | 0.0005067 | AZGP1 | chr7 | -6.87092 | 8.10E-06 |
| PADI1 | chr1 | 6.635014 | 0.012041 | MYH11 | chr16 | -6.79036 | 0.000277 |
| S100A9 | chr1 | 6.557596 | 0.000457 | PLA2G2A | chr1 | -6.70155 | 0.00044 |
| SPRR2A | chr1 | 6.434732 | 0.000212 | CNN1 | chr19 | -6.68817 | 7.24E-05 |
| CRNN | chr1 | 6.350062 | 0.006995 | PPP1R1B | chr17 | -6.5309 | 1.43E-05 |
| RHCG | chr15 | 6.340818 | 0.004233 | LYZ | chr12 | -6.39861 | 3.31E-05 |
| LYPD3 | chr19 | 6.27005 | 0.000446 | PTGER3 | chr1 | -6.36724 | 8.18E-05 |
pairs were positive and 99 pairs were negative. This network also showed that a single IncRNA could regulate the mRNA expression of multiple coding genes and that some IncRNAs could coregulate the expression of the same gene (Fig. 3).

**Enrichment analysis of mRNAs coexpressed with IncRNAs**

To investigate the potential functions of the differentially expressed IncRNAs in the progression of SNIP, we analyzed the functional (GO and pathway) enrichment of the candidate mRNAs in the IncRNA-mRNA coexpression network. The GO analytical data showed that several significantly overrepresented GO terms were included in the biological process, molecular function and cellular component categories. These mRNAs were enriched in multiple biological processes, such as natural killer cell-mediated cytotoxicity (GO: 0042267), negative regulation of proteolysis (GO: 0045861) and negative regulation of protein processing (GO: 0010955). Pathway enrichment analysis indicated that this subset of differentially expressed mRNAs was involved in the PPAR signaling pathway (ID: hsa03320), Jak-STAT signaling pathway (ID: hsa04630) and insulin signaling pathway (ID: hsa04910) (Fig. 4).

**Bioinformatics analysis of Inc-AKTIP**

Lnc-AKTIP was predicted to be found on chromosome 16q12.2 and assumed to be unable to encode genes (Fig. 5A). The optimal secondary structure for Lnc-AKTIP had several hairpin loops with a minimum free energy (MFE) of -152.30 kcal/mol (Fig. 5B). There were strong interactions of Lnc-AKTIP with several proteins. Among these proteins, ELAVL3, ELAVL2 and PCBP2 proteins were found to frequently bind with Lnc-AKTIP (Fig. 5C). Prediction of potential transcription factors showed that MSTRG.93125.4 can combine with 46 transcription factors, such as IRF-1, p53, GATA-2, Elk-1, and HNF-1A (Fig. 5D). The coexpressed mRNAs of Lnc-AKTIP were used for further pathway enrichment analysis, and multiple tumor-related signaling pathways, including chemical carcinogenesis, the p53 signaling pathway, and viral protein interactions with cytokines and cytokine receptors, were found to be enriched (Fig. 5E).

**Prediction efficiency of Lnc-AKTIP as a potential biomarker of SNIP**

To confirm whether Lnc-AKTIP could act as a diagnostic biomarker of SNIP, we performed qRT-PCR in two extended panels of the SNIP group (n=41) and the control group (n=12) and established ROC curves to determine its diagnostic contribution in SNIP. ROC curve analysis revealed that Lnc-AKTIP had a high accuracy in distinguishing SNIP patients from controls [AUC=0.939; 95% CI: 0.859–1.019; sensitivity=95.1%, specificity=81.6%] (Fig. 6).

**Relationships between clinicopathological features and Lnc-AKTIP expression**

To determine the potential clinicopathological implications of altered Lnc-AKTIP expression, the correlations between Lnc-AKTIP expression and clinicopathological features from 41 patients with SNIP were analyzed using the chi-square test. Lnc-AKTIP was found to be significantly related to tumor stage (P<0.007) (Table 5).
### Table 5

Relationships between clinicopathological characteristics and Inc-AKTIP expression in patients with SNIP

| Parameter        | Inc-AKTIP |  |  |
|------------------|-----------|---|---|
|                  | Up\(^{a}\) regulation (%;N=22) | Down\(^{b}\) regulation (%;N=19) | \(P\) Value |
| Age              | 0.994     |  |  |
| <50 years        | 8(36.3)   | 6(31.6)  |  |  |
| \(\geq 50\) years | 14(63.6)  | 13(68.4) |  |  |
| Gender           | 0.699     |  |  |
| Male             | 15(68.2)  | 14(73.7) |  |  |
| Female           | 7(31.8)   | 5(26.3)  |  |  |
| Smoking status   | 0.921     |  |  |
| Yes              | 13(59.1)  | 13(68.4) |  |  |
| No               | 9(40.9)   | 6(31.6)  |  |  |
| Tumor staging\(^{c}\) | 0.007     |  |  |
| I+II             | 19(86.4)  | 9(47.4)  |  |  |
| I+III            | 3(13.6)   | 10(52.6) |  |  |
| Recurrence       | 0.803     |  |  |
| Yes              | 3(13.6)   | 3(15.8)  |  |  |
| No               | 19(86.4)  | 16(84.2) |  |  |

\(^{a}\) Greater than or equal to 0.182 (Average relative expression level of Inc-AKTIP in 41 patients with SNIP)

\(^{b}\) less than 0.182

\(^{c}\) Tumor staging is defined according to a staging system for inverted papilloma described by Krouse JH (2000).

### Discussion

The etiology and pathogenesis of SNIP are still unclear and may be related to HPV infection. p53, gelsolin and cathepsin S have been found to be abnormally expressed in SNIP tissue, some of which are possibly associated with the occurrence and development of SNIP [16–18]. However, there are few studies on how to regulate the expression levels of these genes at the posttranscriptional level. Our previous microarray study found that 58 miRNAs, as ncRNAs, were significantly differentially expressed in SNIP tissues, and the expression level of miRNA-214-3p was correlated with SNIP tumor stage and recurrence [19]. Kakizaki et al [20] also found that miR-296-3p might play a critical role in the malignant transformation of SNIP via the regulation of PTEN and the subsequent inhibition of the PI3K/Akt signaling pathway. To explore the expression levels and potential molecular mechanism of IncRNAs in SNIP, we detected the expression profile of IncRNAs in SNIP tissues by a microarray assay. Our data revealed that a total of 3,435 IncRNAs (1,668 upregulated and 1,767 downregulated) were significantly differentially expressed in the SNIP tissues compared with the corresponding nontumorous tissues. These epigenetic studies suggested that noncoding RNAs likely participate in the occurrence and development of SNIP by regulating the expression of tumor-related genes.

The biological functions of IncRNAs are very complex and have not yet been fully elucidated. It is generally accepted that IncRNAs can regulate the expression level of mRNAs through a variety of molecular mechanisms, such as interfering with the transcription
of the promoter regions of protein-encoding genes, inhibiting RNA polymerase II or mediating chromatin remodeling and histone modification to regulate the expression of downstream genes, blocking mRNA cleavage by complementary double chain structure, and combining with miRNA response elements (MREs) to interfere with the expression of miRNA target genes [21–23]. Thus, IncRNAs can directly bind to transcription factors or change chromatin structure at the transcriptional level. They can also be involved in regulating mRNA processing and translation at the posttranscriptional level. Our study showed that there was a complex coexpression network between the IncRNAs and mRNAs in SNIP. Furthermore, the mRNAs coexpressed with IncRNA AKTIP were enriched in tumor-related biological processes and signaling pathways, such as chemical carcinogenesis, the p53 signaling pathway, and viral protein interactions with cytokines and cytokine receptors. These results revealed that Inc-AKTIP on chromosome 16q12.2 might regulate the pathogenesis of SNIP through mRNAs coexpressed with IncRNAs and a variety of potential molecular mechanisms.

In addition, our bioinformatics analysis indicated that Inc-AKTIP potentially interacted with Inc-AKTIP and PCBP2 protein. PCBP2 is one of the major cellular poly(rC)-binding proteins. Together with PCBP1, this protein also functions as a translational coactivator of poliovirus RNA via a sequence-specific interaction with stem-loop IV of the internal ribosome entry site (IRES), promoting poliovirus RNA replication by binding to its 5’-terminal cloverleaf structure. It has also been implicated in translational control of 15-lipoxygenase mRNA, HPV mRNA, and hepatitis A virus RNA [24, 25]. Wen D et al [26] reported that LINC02535 functions with PCBP2 to facilitate the repair of DNA damage and then to promote cervical cancer progression by stabilizing RRM1 mRNA. Our study also found that IncRNA AKTIP might also bind to multiple tumor-related transcription factors, such as IRF-1, p53 and GATA-2. Thus, IncRNA AKTIP may also regulate the occurrence and development of SNIP via some potential interaction mechanisms between ncRNAs and proteins.

A growing number of studies have confirmed that IncRNAs can be used as molecular biomarkers for the diagnosis and prognosis evaluation of many diseases. For instance, serum IncRNA LOC284454 has been shown to be a good clinical diagnostic biomarker in nasopharyngeal carcinoma, oral cancer, and thyroid cancer [27]. Kopczynska M et al [28] reported that in head and neck squamous cell carcinoma (HNSCC) patients, HPV-positive patients with high IncRNA PRINS expression demonstrated significantly better overall survival and disease-free survival than those with low expression. A majority of HPV-positive patients with high PRINS expression demonstrated a high number of immune cells within tumors. It is likely that IncRNA PRINS could be used as a potential prognostic biomarker for HNSCC patients. Along with these clues, we investigated which of the differentially expressed IncRNAs in SNIP tissue could be used as diagnostic biomarkers. Our study demonstrated that Inc-AKTIP could yield a ROC curve area of 0.939 with 95.1% sensitivity and 81.6% specificity in discriminating SNIP patients from controls. Therefore, Inc-AKTIP probably provides great potential as a novel biomarker in the molecular pathological diagnosis of SNIP. However, its specificity needs to be analyzed in depth between SNIP and other sinonasal tumors. In addition, it remains to be further explored whether plasma/serum Inc-AKTIP can serve as a noninvasive diagnostic biomarker for the early detection of SNIP.

The expression levels of IncRNAs are closely related to the pathological characteristics of many tumors. Wang P et al [29] demonstrated that IncRNA NEAT1 might act as an oncogene. Its increased expression was correlated with T grade, neck nodal metastasis, clinical staging, drinking history, and smoking history in laryngeal squamous cell carcinoma patients. We also found that the expression of Inc-AKTIP was significantly associated with tumor staging of SNIP, which provided a new clue for us to clarify the epigenetic mechanism of SNIP.

In conclusion, our investigation, although preliminary, has revealed the expression profiles of IncRNAs and their potential biological functions in SNIP. We have further demonstrated the potential interaction between Inc-AKTIP and multiple tumor-related transcription factors and the significant correlation between the downregulation of Inc-AKTIP and SNIP tumor stage. Lnc-AKTIP might act as a putative biomarker in SNIP. Such information would be helpful in further investigating the pathogenesis of SNIP and identifying novel therapeutic targets for the treatment of SNIP patients.

Abbreviations

IncRNAs: Long noncoding RNAs; SNIP: Sinonasal inverted papilloma; HPV: Human papillomavirus; ncRNAs: noncoding RNAs; RIN: RNA integrity number; GAPDH: Glyceraldehyde 3-phosphate dehydrogenase; PCC: Pearson correlation coefficient; FDR: False discovery rate; RT: Reverse transcription; DAVID: Database for Annotation, Visualization, and Integrated Discovery; GO: Gene
Declarations

Acknowledgements

We particularly thank the patients, whose biospecimens were analyzed in this study. We are also thankful to Shanghai Biochip Co., Ltd for the construction of the database.

Authors’ contributions

HYH, YL, ZHL and YST conceived and designed the experiments. XLW, YL, ZHL and XLC have made substantial contribution to collected tissue samples. HYH, XLW, YYL, XJL, and ZQM performed the experiments and analyzed the data. YST and YDL prepared the figures and tables and drafted the manuscript. All authors read and approved the final manuscript.

Availability of data and materials

The datasets used and analyzed in the current study are available from the corresponding author on reasonable request.

Funding

This work was supported by grants from the Medical and Health Technology Program of Zhejiang (No: 2021RC101; No: 2019RC065), Science and Technology Development Project of Hangzhou (No: 20201203B199) and Medical and Health Technology Program of Hangzhou (No: A20200274).

Ethics approval and consent to participate

This study was approved by the Ethics Committee of Hangzhou First People’s Hospital and Second Affiliated Hospital, School of Medicine, Zhejiang University. All enrolled patients signed written informed consent prior to sample collection.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Vrabec DP. The inverted Schneiderian papilloma: a 25-year study. Laryngoscope. 1994;104(5 Pt 1):582–605.
2. Krouse JH. Development of a staging system for inverted papilloma. Laryngoscope. 2000;110(6):965–8.
3. Govindaraj S, Wang H. Does human papilloma virus play a role in sinonasal inverted papilloma? Curr Opin Otolaryngol Head Neck Surg. 2014;22(1):47–51.
4. Zydroń R, Marszalek A, Bodnar M, Kosikowski P, Greczka G, Wierzbicka M. The analysis of expression of p16 protein in group of 53 patients treated for sinonasal inverted papilloma. Braz J Otorhinolaryngol. 2018;84(3):338–43.

5. Fulla M, Szafarowski T, Frias-Gomez J, Quiros B, Clavero O, Goma M, et al. Human Papillomavirus and Factors Associated with Recurrence in Sinonasal Inverted Papillomas from Poland and Spain. Head Neck Pathol. 2020;14(3):758–67.

6. Papagiannopoulos P, Tong CL, Kuan EC, Tajudeen BA, Yver CM, Kohanski MA, et al. Inverted papilloma is associated with greater radiographic inflammatory disease than other sinonasal malignancy. Int Forum Allergy Rhinol. 2020;10(3):278–81.

7. Ranjbar R, Behjatfar M, Teimouri A, Aghaie FA, Maniati M, Taheri-Anghaneh M. Long non-coding RNAs and microorganism-associated cancers. Cell Biochem Funct. 2021;39(7):844–53.

8. Bracht JR, Wang X, Shetty K, Chen X, Uttarotai GJ, Callihan EC, et al. Chromosome fusions triggered by noncoding RNA. RNA Biol. 2017;14(5):620–31.

9. Santoro M, Nociti V, Lucchini M, Loiodice M, Centofanti F, Botta A, et al. A pilot study of IncRNAs expression profile in serum of progressive multiple sclerosis patients. Eur Rev Med Pharmacol Sci. 2020;24(6):3267–73.

10. Mei B, Liu H, Yang S, Liang MY, Yue Y, Huang SQ, et al. Long non-coding RNA expression profile in permanent atrial fibrillation patients with rheumatic heart disease. Eur Rev Med Pharmacol Sci. 2018;22(20):6940–7.

11. Shi T, Gao G, Cao Y. Long Noncoding RNAs as Novel Biomarkers Have a Promising Future in Cancer Diagnostics. Dis Markers. 2016;2016:9085195.

12. Brunner AL, Beck AH, Edris B, Sweeney RT, Zhu SX, Li R, et al. Transcriptional profiling of long non-coding RNAs and novel transcribed regions across a diverse panel of archived human cancers. Genome Biol. 2012;13(8):R75.

13. Khandelwal A, Bacolla A, Vasquez KM, Jain A. Long non-coding RNA: A new paradigm for lung cancer. Mol Carcinog. 2015;54(11):1235–51.

14. Chandra GS, Nandan TY. Potential of long non-coding RNAs in cancer patients: From biomarkers to therapeutic targets. Int J Cancer. 2017;140(9):1955–67.

15. Wu T, Du Y. LncRNAs. From Basic Research to Medical Application. Int J Biol Sci. 2017;13(3):295–307.

16. Altavilla G, Staffieri A, Busatto G, Canesso A, Giacomelli L, Marioni G. Expression of p53, p16INK4A, pRb, p21WAF1/CIP1, p27KIP1, cyclin D1, Ki-67 and HPV DNA in sinonasal endophytic Schneiderian (inverted) papilloma. Acta Otolaryngol. 2009;129(11):1242–9.

17. Cho JE, Park W, Kim DC, Kim HJ, Kim SW, Kang JM, et al. Down-regulation of gelsolin may play a role in the progression of inverted papilloma through an antiapoptotic mechanism. Am J Rhinol Allergy. 2012;26(3):177–82.

18. Huang CC, Lee TJ, Chang PH, Lee YS, Chuang CC, Jhang YJ, et al. Expression of cathepsin S and its inhibitor stefin A in sinonasal inverted papilloma. Rhinology. 2010;48(3):352–7.

19. Teng Y, Li Y, Lin Z, Gao Y, Cao X, Lou X, et al. Analysis of miRNA expression profiling identifies miR-214-3p as a novel biomarker in sinonasal inverted papilloma. Epigenomics-UK. 2018;10(12):1541–53.

20. Kakizaki T, Hatakeyama H, Nakamaru Y, Takagi D, Mizumachi T, Sakashita T, et al. Role of microRNA-296-3p in the malignant transformation of sinonasal inverted papilloma. Oncol Lett. 2017;14(1):987–92.

21. Sharma A, Kansara S, Mahajan M, Yadav B, Garg M, Pandey AK. Long non-coding RNAs orchestrate various molecular and cellular processes by modulating epithelial-mesenchymal transition in head and neck squamous cell carcinoma. Biochim Biophys Acta Mol Basis Dis. 2021;1867(11):166240.

22. Vance KW, Ponting CP. Transcriptional regulatory functions of nuclear long noncoding RNAs. Trends Genet. 2014;30(8):348–55.

23. Zhang K, Shi ZM, Chang YN, Hu ZM, Qi HX, Hong W. The ways of action of long non-coding RNAs in cytoplasm and nucleus. Gene. 2014;547(1):1–9.

24. Yanatori I, Richardson DR, Toyokuni S, Kishi F. The new role of poly (rC)-binding proteins as iron transport chaperones: Proteins that could couple with inter-organelle interactions to safely traffic iron. Biochim Biophys Acta Gen Subj. 2020;1864(11):129685.

25. Sean P, Nguyen JH, Semler BL. The linker domain of poly(rC) binding protein 2 is a major determinant in poliovirus cap-independent translation. Virology. 2008;378(2):243–53.
26. Wen D, Huang Z, Li Z, Tang X, Wen X, Liu J, et al. LINC02535 co-functions with PCBP2 to regulate DNA damage repair in cervical cancer by stabilizing RRM1 mRNA. J Cell Physiol. 2020;235(10):7592–603.

27. Fan C, Wang J, Tang Y, Zhang S, Xiong F, Guo C, et al. Upregulation of long non-coding RNA LOC284454 may serve as a new serum diagnostic biomarker for head and neck cancers. BMC Cancer. 2020;20(1):917.

28. Kopczynska M, Kolenda T, Guglas K, Sobocinska J, Teresiak A, Blizniak R, et al. PRINS IncRNA Is a New Biomarker Candidate for HPV Infection and Prognosis of Head and Neck Squamous Cell Carcinomas. Diagnostics (Basel). 2020;10(10):762.

29. Wang P, Li QY, Sun YN, Wang JT, Liu M. Long Non-coding RNA NEAT1: A Potential Biomarker in the Progression of Laryngeal Squamous Cell Carcinoma. ORL J Otorhinolaryngol Relat Spec. 2021:1–7.

Figures

![Figure 1](image)

**Figure 1**

Alterations in IncRNA and mRNA expression profiles between SNIP tissues and nontumorous tissues. A-B: The violin plot is a convenient way to quickly visualize the distributions of a dataset of IncRNA (A) and mRNA (B) profiles. C-D: Hierarchical clustering analysis of differentially expressed IncRNAs (C) and mRNAs (D) (fold change>2; P<0.05). Each row indicates one IncRNA or mRNA,
and each column indicates one sample. The lncRNA and mRNA expression levels are illustrated using a color key and histogram. The red and blue bars denote high and low relative expression, respectively.

**Figure 2**

qRT-PCR validation of eight differentially expressed lncRNAs and three mRNAs selected by the microarray assay. (A) Comparison of lncRNA or mRNA expression levels obtained by microarray and qRT-PCR analysis. Upregulated and downregulated lncRNAs or mRNAs are indicated by bars above and below the horizontal axis, respectively. (B) Pearson correlation analysis investigating the linear relationship between the microarray data and qRT-PCR results.
Figure 3
Prediction of the lncRNA-mRNA association network. The coexpression network was composed of 305 network nodes and 769 connections between 155 lncRNAs and 150 coding genes. The red and green circles denote high and low relative expression, respectively. The arrow represents positive regulation, and the flat-head line represents negative regulation.
Figure 4

An enrichment analysis of mRNAs coexpressed with IncRNAs. A: Top 30 GO enrichment terms. B: Top 30 KEGG pathway enrichment terms.
Figure 5

Bioinformatics analysis of Inc-AKTIP. A: Chromosome location of Inc-AKTIP. B: Optimal secondary structure for Inc-AKTIP. C: Interaction between Inc-AKTIP and RNA motif proteins. D: Interaction between Inc-AKTIP and transcription factors. Lnc-AKTIP is indicated by a red circle. The transcription factors are indicated by blue circles. E: Pathway annotation of the coexpressed mRNAs of Inc-AKTIP.
Figure 6

ROC curve analysis was used to evaluate the accuracy of Inc-AKTIP in discriminating SNIP patients.