Integrated analysis revealing genome-wide chromosomal copy number variation in supraglottic laryngeal squamous cell carcinoma

DONGJIE LI1, XIANFU WANG2, SHUNFEI LU3, PING WANG1, XIN WANG1, WANZHONG YIN1, WEI ZHU1 and SHIBO LI2

1Department of Otorhinolaryngology, Head and Neck Surgery, The First Hospital of Jilin University, Changchun, Jilin 130021, P.R. China; 2Department of Pediatrics, Genetics Laboratory, University of Oklahoma Health Sciences Center, Oklahoma, OK 73104, USA; 3Department of Clinical Medicine, Lishui College of Medicine, Lishui, Zhejiang 323000, P.R. China

Received October 24, 2019; Accepted April 27, 2020

DOI: 10.3892/ol.2020.11653

Abstract. Laryngeal squamous cell carcinoma (LSCC) is a genetically complex tumor type and one of the leading causes of cancer-associated disability and mortality. Genetic instability, such as chromosomal instability, is associated with the tumorigenesis of LSCC. Copy number variations (CNVs) have been demonstrated to contribute to the genetic diversity of tumor pathogenesis. Comparative genomic hybridization (CGH) has emerged as a high-throughput genomic technology that facilitates the aggregation of high-resolution data of cancer-associated genomic imbalances. In the present study, a total of 38 primary supraglottic LSCC cases were analyzed by high-resolution array-based CGH (aCGH) to improve the understanding of the genetic alterations in LSCC. Additionally, integration with bioinformatic analysis of microarray expression profiling data from the Gene Expression Omnibus (GEO) database provided a fundamental method for the identification of putative target genes. Genomic CNVs were detected in all cases. The size of net genomic imbalances per case ranged between a loss of 682.3 Mb (~24% of the genome) and a gain of 1,958.6 Mb (~69% of the genome). Recurrent gains included 2pter-q22.1, 3q26.1-qter, 5pter-p12, 7p22.3p14.1, 8p12p11.22, 8q24.13q24.3, 11q13.2q13.4, 12pter-p12.2, 18pter-p11.31 and 20p13p12.1, whereas recurrent losses included 3pter-p21.32, 4q28.1-q35.2, 5q13.2-qter, 9pter-p21.3 and monosomy 13. Gains of 3q26.1-qter were associated with tumor stage, poor differentiation and smoking history. Additionally, through integration with bioinformatic analysis of data from the GEO database, putative target oncogenes, including sex-determining region Y-box 2, eukaryotic translation initiation factor 4 gamma 1, fragile X-related gene 1, disheveled segment polarity protein 3, defective n cullin neddylation 1 domain containing 1, insulin like growth factor 2 mRNA binding protein 2 and CCDC26 long non-coding RNA, and tumor suppressor genes, such as CUB and sushi multiple domains 1, cyclin dependent kinase inhibitor 2A, protocadherin 20, serine peptidase inhibitor Kazal type 5 and Nei like DNA glycosylase 3, were identified in supraglottic LSCC. Supraglottic LSCC is a genetically complex tumor type and aCGH was demonstrated to be effective in the determination of molecular profiles with higher resolution. The present results enable the identification of putative target oncogenes and tumor suppressor gene mapping in supraglottic LSCC.

Introduction

Laryngeal squamous cell carcinoma (LSCC) was the second most predominant type of head and neck squamous cell cancer (HNSCC), and the sixth most common type of carcinoma worldwide in 2011 (1). Laryngeal cancer accounts for 2-5% of all malignancies with ~200,000 deaths annually (2); however, it has a major impact on phonation, swallowing, respiration and quality of life (3). Histologically, >95% of laryngeal carcinoma cases are identified as LSCC (4). For clinical and staging purposes, laryngeal cancer is currently clustered in supraglottic, glottic and subglottic regions, among which supraglottic LSCC is the second most common type of laryngeal cancer (5) and presents different biological properties, such as being more likely to recur and metastasize (6). Notably, supraglottic LSCC is more prevalent in some developing countries where alcohol and smoking are common risk factors (7).
The occurrence and development of LSCC are synergistically caused by extrinsic and inheritable multi-etiological factors and their interactions. In addition to environmental carcinogens, such as exposure to tobacco, alcohol and radioactive rays, and infectious agents, including viral infections, molecular and chromosomal alterations are the most common inheritable etiological factors (8). Copy number variations (CNVs) are the prevalent structural variations in the genome and have been demonstrated to contribute to the genetic diversity of tumor pathogenesis, being increasingly accepted as a major source of inheritable variations (9). CNVs have been demonstrated to result in loss-of-function variants of tumor suppressor genes and/or gain-of-function variants of oncogenes (10). Notably, high levels of DNA CNVs in tumors are frequently limited to certain chromosomal segments with renowned oncogenes that are overexpressed or triggered (11).

Comparative genomic hybridization (CGH) has emerged as a high-throughput genomic technology to facilitate the aggregation of high-resolution data of cancer-associated genomic imbalances. In the present study, high-resolution genomic microarray CGH was applied to detect regions of gain or amplification and losses in supraglottic LSCC. By integration of CNVs with bioinformatic analysis, the possible candidate target genes with strong clinical significance were identified, providing novel insights into the molecular pathological mechanisms underlying supraglottic LSCC. Additionally, associations between genetic alternations and lymph node metastasis, tumor stage and distant metastasis were statistically analyzed, and differentially expressed genes (DEGs) between samples from patients with HNSCC and normal controls were identified using microarray data from the Gene Expression Omnibus (GEO) database to determine the optimal gene combination with diagnostic value for HNSCC. The findings of the present study may elucidate the mechanism underlying the development of supraglottic LSCC and uncover potential diagnostic and therapeutic biomarkers for supraglottic LSCC.

Materials and methods

Collection of tumor tissues. Tissues from 38 patients with primary supraglottic LSCC were collected during excision surgery between November 2011 and October 2016 at the Department of Otorhinolaryngology, Head and Neck Surgery, The First Hospital of Jilin University (Changchun, China). All patients had histopathologically confirmed LSCC and TNM stage was determined according to American Joint Committee on Cancer staging system (12). The present study included 31 male and 7 female patients aged between 52 and 75 years (median age, 62 years). The clinicopathological features and risk factors of the patients with supraglottic LSCC are summarized in Table I. All patients had negative histories of exposure to either chemotherapy or radiotherapy before surgery, and no other cancers were detected.

Written informed consent was provided by all patients enrolled in the present study with the authorization of the Ethics Committee of The First Hospital of Jilin University. Tumor tissues and adjacent non-tumor tissues (>1 cm from tumor) were surgically obtained at the Department of Otorhinolaryngology, Head and Neck Surgery. The primary tumor tissues were obtained from the center of the tumor lesion, confirmed by an independent pathologist after surgical removal, immediately snap-frozen in liquid nitrogen and stored at -80°C until further use. Samples were first digested using proteinase K, followed by the phenol-chloroform method for DNA isolation according to standard protocols. Briefly, frozen tumor samples were grinded to small pellets and treated with SE buffer (75 mM NaCl, 25 mM EDTA and 10N NaOH pH 8.0), 20% SDS and proteinase K. After incubation overnight at 50°C, DNA was extracted by phenol-chloroform extraction and ethanol precipitation. The DNA was washed with 70% ethanol and air dried. The DNA was dissolved in 50 µl TE buffer (10 mM Tris HCl pH 7.4 and 1 mM EDTA) and the DNA concentration was measured by Nanodrop 2000 (Thermo Fisher Scientific, Inc.) and run on a 0.8% agarose gel.

Array-based CGH (aCGH) assay. aCGH was conducted on a SurePrint 2x400 k oligomer CGH chip (Agilent Technologies, Inc.) according to the manufacturer's protocol. The reference DNA used was included in the aforementioned labeling kit to identify CNVs (male reference DNA, cat. no. 5190-3796; and female reference DNA, cat. no. 5190-3797; Agilent Technologies, Inc.). The DNA from patients and the reference DNA were labeled with different color fluorescent dyes via PCR for 40 h at 48°C, namely with Cyanine 5 and Cyanine 3, respectively, and then equal labeling products were mixed together for competitive co-hybridization onto a chip by incubation in a hybridization oven (Agilent Technologies, Inc.) for 40 h at 48°C. The chip was washed using the washing buffer in the kit and scanned with an MS200 laser scanner using the NimbleGen system (Roche Applied Science).

Array data analysis. Images were analyzed using the Cytogenomics v4.0 software (Agilent Technologies, Inc.). The cut-off threshold of positive CNVs was determined using a log2 ratio with ±0.15 in ≥10 neighboring probes coordinating gain or loss, due to complicated DNA components from tumor tissues, such as tumor cell diversity and mosaic variation. Amplifications were specified as those with a smoothed log R ratio >0.5, while homozygous deletions were specified as those with a smoothed log R ratio <0.5.

Affymetrix microarray data. The microarray gene expression profiling dataset GSE6631 for HNSCC was downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS2520). The HNSCC-associated dataset GSE6631 was from the GPL8300 Affymetrix Human Genome U95 Version 2 Array platform, and included 44 HNSCC tumors and 44 matching normal mucosa samples (13). GSE6631 series matrix .txt files and GPL8300 platform .txt files were extracted for data processing.

Data processing and identification of DEGs. The gene probe IDs in the matrix files were converted to the gene symbols in the platform files using a Perl script to generate a matrix file with the global standard gene name. Each dataset was then standardized using the normalizeBetweenArrays function of the limma R package v3.10.1 (http://www.bioconductor.org/packages/release/bioc/html/limma.html). All gene expression
data were transformed to a log₂ scale. DEGs between HNSCC tumor tissues and control samples were determined using the limma R package. P-values were adjusted using the Benjamini-Hochberg false discovery rate method, and adjusted P<0.05 and |log fold-change (FC)|>1 were used as the cut-off criteria. Heatmaps and volcano maps were visualized using the ggplot2 R package v3.0.0.

Statistical analysis. Data are presented as the case number and all experiments were performed once. Significant associations between genomic aberrations and clinicopathological factors were determined using the \( \chi^2 \) test in SPSS v20.0 (IBM Corp.). P<0.05 was considered to indicate a statistically significant difference.

Results

Identification of recurrent CNVs in supraglottic LSCC via aCGH. Genomic CNVs (gains, losses, amplifications and homozygous deletions) were detected via aCGH in all 38 samples of supraglottic LSCC. Genomic imbalance profiling is presented in Fig. 1. Net gains (21 cases) were more frequent compared with net losses (17 cases). As shown in Fig. 2, the sizes of net genomic imbalances per case ranged between a loss of 849.4 Mb (~29.9% of the genome) and a gain of 1,958.6 Mb (~69.0% of genome). The average number of gains per case was 15.7, ranging between 2 and 37, whereas the average number of losses per case was 9.2, ranging between 0 and 24 (data not shown). The gain sizes ranged between 252 kb and 204 Mb, whereas the loss sizes ranged between 174 kb and 198 Mb. Overall, 41/946 (4.3%) of the total genomic imbalances were <1 Mb, among which 31/41 (75.6%) were gains and 10/41 (24.4%) were losses (data not shown). The most frequent genomic imbalances were considered as those detected in ≥10/38 (26.3%) samples of supraglottic LSCC and included 10 gains (2pter‑q22.1, 3q26.1‑qter, 5pter‑p12, 7p22.3p14.1, 8p12p11.22, 8q24.13q24.3, 11q13.2q13.4, 12pter‑p12.2, 18pter‑p11.31 and 20p13p12.1), and 5 losses in various chromosome regions (3pter‑p21.32, 4q28.1‑q35.2, 5q13.2‑qter, 9pter‑p21.3 and monosomy 13) (Table II). As presented in Fig. 3, nearly half of the cases had between 10 and 29 genetic alterations.

The high-level copy number gains were analyzed in the amplifications with a log2 ratio >0.5. A total of 41 amplified chromosome segmental regions were detected and are summarized in Table III. Among these regions, the 3q26.32q27.2 region was amplified in 8 cases and gained in 23 cases, in which the size of the smallest region of overlap (SRO) was ~9.48 Mb, including sex-determining region Y-box 2 (SOX2), eukaryotic translation initiation factor 4 gamma 1 (EIF4G1), fragile X-related gene 1 (FXR1), disheveled segment polarity protein 3 (DVL3), defective n cullin neddylation 1 domain containing 1 (DCUN1D1) and insulin like growth factor 2 mRNA binding protein 2 (IGF2BP2) genes. The 8q24.21 region was amplified in 7 cases and gained in 14 cases, with an SRO of ~672 kb in size, including the CCDC26 long non-coding RNA (CCDC26) gene.

Four interesting potential homozygous losses with a log2 ratio >-0.4 were identified (Table IV), harboring putative tumor suppressor genes, such as Nei-like DNA glycosylase 3 (NEIL3), CUB and Sushi multiple domains 1 (CSMD1), cyclin dependent kinase inhibitor 2A (CDKN2A) and protocadherin 20 (PCDH20).

Genomic imbalances are associated with lymph node metastasis, tumor stage, tumor differentiation and smoking history. Patients with LSCC were divided into different groups according to their clinicopathological characteristics as follows: i) With or without lymph node metastasis; ii) tumor stages I-II or III-IV; iii) well, moderate or poor differentiation;
Figure 1. Summary of the array-based comparative genomic hybridization results from 38 samples of supraglottic laryngeal squamous cell carcinoma. DNA gains are presented as green vertical lines on the right of the chromosome idiograms, whereas DNA losses are indicated as red vertical lines on the left of the chromosome idiograms.

Figure 2. Net genomic imbalances in 38 samples of supraglottic laryngeal squamous cell carcinoma. Red columns indicate genomic losses and green columns represent genomic gains, while blue columns represent net genomic imbalances. Net genomic gains were more common than net genomic losses.
and iv) with or without a smoking history. Differences in frequencies (>10/38) of genomic aberrations among the different groups of patients with supraglottic LSCC were analyzed using the \( \chi^2 \) test, revealing a significant association between 3q26.1-qter gain and tumor stage, poor differentiation and smoking history (Table V).

Identification of DEGs between patients with HNSCC and normal controls. The expression levels in the HNSCC chip dataset GSE6631 were normalized (Fig. 4). A total of 140 DEGs were detected between 44 HNSCC tumors and their paired normal mucosa samples using adjusted P<0.05 and |log FC|>1 as cut-off criteria, among which 89 DEGs were downregulated and 51 were upregulated. The volcano plot and heatmap of DEGs are shown in Figs. 5 and 6, respectively. The chromosome regions of upregulated and downregulated DEGs are listed in Table VI.

Candidate target genes of gains and losses in supraglottic LSCC. The Affymetrix microarray expression profiling data of HNSCC and the gene expression profiling of supraglottic LSCC samples detected by aCGH were integratively analyzed to identify the potential target genes of downregulated serine peptidase...
inhibitor Kazal type 5 (SPINK5), which is closely associated with squamous cell carcinoma. Analysis of candidate target genes of gains in 3q26.32q27.2 and 8q24.21 revealed upregulation of SOX2, EIF4G1, FXR1, DVL3, DCUN1D1, IGF2BP2, and CCDC26 (Table III), and downregulation of CDKN2A, SPINK5, PCDH20, CSMD1, and NEIL3 in supraglottic LSCC (Table IV).

**Discussion**

The present study evaluated genetic alternations in LSCC using aCGH and bioinformatics analysis of microarray expression profiling. Primary LSCC accounts for 20% of all HNSCC and is associated with a high risk of developing metastatic...
Table IV. Potential homozygous segments and target genes in supraglottic laryngeal squamous cell carcinoma samples.

<table>
<thead>
<tr>
<th>Chromosome region</th>
<th>Log&lt;sub&gt;2&lt;/sub&gt; ratio</th>
<th>Genomic coordinates, bp</th>
<th>Size, Mb</th>
<th>Target gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>4q34.3</td>
<td>0.6</td>
<td>177027640-178012479</td>
<td>0.94</td>
<td>NEIL3</td>
</tr>
<tr>
<td>8p23.3p23.2</td>
<td>0.8</td>
<td>0-4845472</td>
<td>4.84</td>
<td>CSMD1</td>
</tr>
<tr>
<td>9p21.3</td>
<td>0.5</td>
<td>21952870-25578273</td>
<td>3.62</td>
<td>CDKN2A</td>
</tr>
<tr>
<td>13q21.2</td>
<td>0.6</td>
<td>61391312-614188052</td>
<td>1.10</td>
<td>PCDH20</td>
</tr>
</tbody>
</table>

Figure 4. GSE6631 dataset normalized for gene expression. (A) Gene expression before normalization. (B) Gene expression after normalization.

Figure 5. Volcano plot of the differentially expressed genes between supraglottic laryngeal squamous cell carcinoma samples and paired controls in the GSE6631 dataset. Red dots represent upregulated and downregulated genes based on an adj.P<0.05 and |logFC|>1, and black dots represent genes with no significant difference in expression. adj.P.Val, adjusted P-value; FC, fold-change.
Table V. Genomic aberrations associated with clinicopathological characteristics of the supraglottic laryngeal squamous cell carcinoma samples.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Gain/loss</th>
<th>pN status, n (%)</th>
<th>Stage, n (%)</th>
<th>Differentiation, n (%)</th>
<th>Smoking history, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive</td>
<td>Negative</td>
<td>c2</td>
<td>P-value</td>
</tr>
<tr>
<td>3q26.1-qter</td>
<td>gain</td>
<td>Yes</td>
<td>15</td>
<td>16</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>3</td>
<td>4</td>
<td>(31) (48.39)</td>
</tr>
<tr>
<td>5pter-p12</td>
<td>gain</td>
<td>Yes</td>
<td>12</td>
<td>14</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>6</td>
<td>6</td>
<td>(26) (46.15)</td>
</tr>
<tr>
<td>18pter-p11.31</td>
<td>gain</td>
<td>Yes</td>
<td>7</td>
<td>13</td>
<td>2.591</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>11</td>
<td>11</td>
<td>(20) (35.00)</td>
</tr>
<tr>
<td>8q24.13-q23.3</td>
<td>gain</td>
<td>Yes</td>
<td>10</td>
<td>9</td>
<td>0.422</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>8</td>
<td>11</td>
<td>(19) (52.63)</td>
</tr>
<tr>
<td>3pter-p21.31</td>
<td>loss</td>
<td>Yes</td>
<td>9</td>
<td>11</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>9</td>
<td>9</td>
<td>(20) (45.00)</td>
</tr>
<tr>
<td>5q13.2-qter</td>
<td>loss</td>
<td>Yes</td>
<td>10</td>
<td>8</td>
<td>0.920</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>8</td>
<td>12</td>
<td>(18) (55.56)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(20) (40.00)</td>
<td>0.095</td>
<td>0.758</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(20) (40.00)</td>
<td>0.095</td>
<td>0.758</td>
<td>8</td>
</tr>
</tbody>
</table>
squamous cell carcinoma (15). Supraglottic LSCC is emerging as the second most common type of laryngeal malignancy (5). In the present study, 38 primary supraglottic LSCC cases and paired controls were screened via aCGH and the high-density Affymetrix Genome-wide Human Genome U95 Version 2 Array from the GEO database to analyze CNVs (gains, losses, amplifications and homozygous deletions) and their association with clinicopathological characteristics. Important roles of certain chromosomal regions and gene loci were identified for the biology and prognosis of supraglottic LSCC.

Chromosomal instability is a numerical and structural chromosomal abnormality commonly resulting in DNA CNVs and genetic heterogeneity that may trigger cancer development processes, including LSCC tumorigenesis (16). Compared with the detection of expression profiles using conventional karyotyping methods, aCGH has been demonstrated to be more effective in the determination of molecular profiles with a 5-10 times higher resolution (17,18). In the present study, a total of 598 gains and amplifications, and 348 losses were identified using aGCH. The frequency of gains and amplifications was 1.72 times higher than that of losses. A total of 21/38 cases with genomic imbalances had net genomic gains (range, 54.86-1,958.63 Mb) and 17/38 cases had net genomic losses (range, 22.02-849.44 Mb), demonstrating a wider range in the net genomic gains than in the losses. The most frequent genomic imbalances detected in the present study included gains such as 3q26.1-qter (31/38), 5pter-p12 (26/38) and 18pter-p11.31 (20/38), and losses such as 3pter-p21.32 (20/38), 5q13.2-qter (18/38) and 9pter-p21.3 (16/38). Previous studies have reported that the gain of genomic material of 3q has a high prevalence in squamous cell carcinomas originating from the head and neck region (19,20), and that it is associated with disease-specific survival time (21), consistent with...
previous studies in LSCC (22-24). The chromosomal region of 3q, particularly 3q26-29, was demonstrated to be a defining feature of squamous cell carcinomas, with an ~75% predicted occurrence in HNSCC (25). Furthermore, gains of 1q13 and loss of 18q23 were positively associated with a poor outcome in patients with LSCC (23,26,27).

In the present study, amplifications were detected in 41 segmental regions, among which 3q26.32q27.2 and 8q24.21 were the most repeatedly involved regions. Amplification of 3q26.32q27.2 was one of the most relevant findings of the present study. In 31/38 cases with copy number gains of 3q26.1-qter, 23 cases exhibited gains and 8 cases amplifications. Amplification of the 3q26q27 region is one of the most recurrent genetic alterations in HNSCC and other carcinomas, and has been demonstrated to be associated with tumor progression and a poor prognosis (20,28). In the present study, the SRO located in locus 3q26q27 harbored several oncogenes, including SOX2, EIF4G1, FXR1, DVL3, DCUN1D1 and IGF2BP2. SOX2 has been demonstrated to be a candidate tumor driver gene in squamous cell carcinoma of the lung and esophagus (29-31). EIF4G1 is an RNA-binding protein that serves an important role in the occurrence and development of breast cancer, squamous cell lung carcinoma and other types of cancer (32-33). Jin et al (34) reported an upregulation of FXR1 expression in colorectal cancer, in which it acted as an oncogene. DVL3 is a family member of the Wnt receptor complex and the most critical regulator of the Wnt signaling pathway, and is involved in cervical, breast, liver, prostate, lung and esophageal cancer (35). DCUN1D1 expression is upregulated in human laryngeal carcinoma tissues (36). IGF2BP2, also known as IMP2, has been reported to be dysregulated in several types of cancer, including colorectal cancer (37), colon cancer (38), esophageal adenocarcinoma (39) and breast cancer (40).

Notably, the reciprocal loss of 3p and gain of 3q were detected in 20/38 cases in the present study, together with 18/38 cases of reciprocal gain of 5p and loss of 5q. Frequent occurrence of the reciprocal loss and gain of chromosomes 3 and 5 were observed in various epithelial tumors. In particular, the isochromosome 3q has been detected in lung cancer and HNSCC (41,42). Therefore, formation of isochromosome 3q may be an intermediate mechanism of somatic chromosomal aberrations, leading to a reciprocal change during epithelial cell carcinogenesis.

In addition, 3 cases of amplification and 14 of gain in the 7p11.2 chromosomal fragment harboring the oncogene epidermal growth factor receptor (EGFR) were detected in the present study. As a tyrosine kinase receptor, EGFR is widely expressed in human epithelial cell membranes. Amplification and upregulation of EGFR have been observed in patients with LSCC and esophageal squamous cell carcinoma (ESCC) with a poor prognosis, serving a crucial role in ESCC progression (43). The association analysis between genomic aberrations and clinicopathological features in the present study revealed that gains of 3q26.1-qter were associated with tumor stage, poor differentiation and smoking history.

Four homozygous segments possibly containing putative tumor suppressor genes were discovered in the present study, including NEIL3, CSMD1, CDKN2A and PCDH20. CDKN2A is considered to be the most commonly affected gene after TP53 in terms of homozygous deletion, promoter hypermethylation, loss of heterozygosity and point mutations in various types of cancer, including LSCC (44,45). CSMD1 has been demonstrated to act as a tumor suppressor in human breast, colorectal and head and neck cancer (46-48), whereas PCDH20 has been proposed as a tumor suppressor gene in hepatic carcinoma and lung cancer (49,50). However, the role of NEIL3 as a tumor suppressor gene remains unclear. Wang et al (51) proposed SPINK5 as a novel tumor suppressor targeting the Wnt/β-catenin signaling pathway in esophageal cancer.

In conclusion, using integration analysis of CNVs from the GEO database and the gene expression profiling of supraglottic LSCC samples detected via aCGH, potential candidate target genes with strong clinical significance were identified, such as the upregulation of SOX2, EIF4G1, FXR1, DVL3, DCUN1D1, IGF2BP2 and CCDC26 expression, and downregulation of CDKN2A, SPINK5, PCDH20, CSMD1 and NEIL3 expression, providing novel insights into the molecular pathological mechanisms underlying supraglottic LSCC. However, the findings of the present study were limited by the relatively small sample size. Future studies should involve the cooperation of researchers from different countries and regions to further verify the findings of the present study, and should investigate the mechanisms underlying the development of supraglottic LSCC, potentially leading to the identification of novel diagnostic and therapeutic biomarkers for supraglottic LSCC.

Acknowledgements
Not applicable.

Funding
No funding was received.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. The GSE6631 dataset is available from the GEO database (https://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS2520).

Authors’ contributions
WZ conceived and designed the study. WY, PW and XiaW acquired the data. DL and SLu analyzed and interpreted the data. DL drafted the manuscript. SLi and XinW critically revised the manuscript for important intellectual content and analyzed the data. WZ and SLi approved the final manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Written informed consent was provided by all patients enrolled in the present study with the authorization of the Ethics Committee of The First Hospital of Jilin University (approval no. 2017-415; Changchun, China).
Patient consent for publication

Written informed consent for publication was provided by all patients.

Competing interests

The authors declare that they have no competing interests.

References


This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.