Screening and identification of potential target genes in head and neck cancer using bioinformatics analysis

FUHAI CHEN, ANYUAN ZHENG, FEN LI, SILU WEN, SHIMING CHEN and ZEZHANG TAO

Department of Otolaryngology-Head and Neck Surgery, Renmin Hospital of Wuhan University, Wuhan, Hubei 430060, P.R. China

Received September 20, 2018; Accepted June 14, 2019

DOI: 10.3892/ol.2019.10616

Abstract. Head and neck cancer (HNC) is the sixth most common cancer worldwide. Recent studies on the pathogenesis of HNC have identified some biochemical associations of this disease, but the molecular mechanisms are not clear. To explore the genetic alterations in head and neck tumors, to identify new high-specificity and high-sensitivity tumor markers, and to investigate potentially effective therapeutic targets, in silico methods were used to study HNC. The GSE58911 microarray dataset was downloaded from the Gene Expression Omnibus online database to identify potential target genes in the carcinogenesis and progression of HNC. Differentially expressed genes (DEGs) were identified and functional enrichment analysis was performed. In addition, a protein-protein interaction network was also constructed, and gene analysis was undertaken using Search Tool for the Retrieval of Interacting Genes and Cytoscape. A total of 648 differentially expressed genes were identified. Kyoto Encyclopedia of Genes and Genomes pathway and Gene Ontology functional enrichment analysis of DEGs included muscle system process, extracellular matrix organization, actin binding, structural molecule activity, structural constituent of muscle, extracellular region part, ECM-receptor interaction, amoebiasis, focal adhesion, drug metabolism-cytochrome P450, and chemical carcinogenesis. There were 26 hub genes identified and biological process analysis revealed that these genes were mainly enriched in extracellular matrix organization, serine-type endopeptidase activity, extracellular matrix, and complement and coagulation cascades. Survival analysis revealed that interleukin (IL)-8 (C-X-C motif chemokine ligand 8), IL1B, and serpin family A member 1 may be involved in the carcinogenesis of HNC. In summary, the DEGs and hub genes identified in the present study may increase understanding of the molecular mechanisms of development of HNC and provide potential target genes for clinical diagnosis and targeted therapy.

Introduction

Head and neck cancer (HNC) is the sixth most common cancer worldwide (1). Annually, about 650,000 new cases and 350,000 deaths are reported worldwide, accounting for 6% of all cases (1,2). High-risk regions for lip and oral cavity cancers include Melanesia, South-Central Asia, and Central and Eastern Europe (2). High-risk regions for laryngeal cancer include Southern and Eastern Europe and Western Asia (2). Indonesia, Singapore, and Malaysia, which are inhabited to a large extent by Malay and Chinese individuals, have the highest incidence of nasopharyngeal carcinoma (3). Mounting evidence suggests that genetic variations or abnormal expression of keratinocyte differentiation associated protein, heme oxygenase 1 (HMOX1), Rac family small GTPase 1 (Rac1), and desmocollin 1 (DSC1) may be associated with the carcinogenesis and progression of head and neck tumors. Studies have also found that HMOX1 and keratin-associated proteins are associated with human papillomavirus (HPV)-related HNC (4,5). Inhibition of Rac1 activity may help to overcome primary or secondary chemo-radio-resistance in HNC (6). Occurrence and clinical prognosis of HNC are associated with overexpression of DSC1 (7). Early diagnosis and early treatment are keys to successful treatment of HNC, but no tumor markers with high specificity and sensitivity, or an effective therapeutic target, have been identified. Therefore, the survival rate and quality of life of patients with HNC is poor. As such, it is necessary to characterize the molecular mechanisms involved in the carcinogenesis of head and neck tumors. Better understanding of these mechanisms will allow for improved guidelines for diagnosis and treatment of head and neck tumors.

To explore the genetic alterations in head and neck tumors, identify new high-specificity and high-sensitivity tumor markers, and identify potentially effective therapeutic targets, in silico methods were used to study HNC. In the present study, GSE58911 was downloaded and analyzed from the Gene Expression Omnibus (GEO) database to obtain differentially expressed genes (DEGs) between HNC tissues...
and non-cancerous tissues. Subsequently, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, Gene Ontology (GO) enrichment analysis, and protein-protein interaction (PPI) network analysis was performed to characterize the molecular mechanisms underlying carcinogenesis and progression of HNC. A total of 648 differentially expressed genes and 26 hub genes were identified, which may be potential targets for clinical diagnosis and therapy of HNC.

Materials and methods

Microarray data. Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) is a public functional genomics data repository of high throughput gene expression data, chips, and microarrays (8,9). Platform (GPL) and Series (GSE) constitute the data from GEO. The gene expression dataset (GSE58911) (10) was downloaded from GEO (Affymetrix GPL6244, Affymetrix Human Gene 1.0 ST Array) and contains 15 HNC samples and 15 normal tissues distant to the HNC sample.

Identification of DEGs. DEGs between cancerous and non-cancerous tissues were screened using GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r). To identify genes that are differentially expressed across experimental conditions, GEO2R, an interactive web tool, allows users to compare >2 groups of samples in a GEO Series. Results are presented as a table of genes ordered by significance. Log fold-change (FC) ≥1 or ≤−1 and adjusted P-value <0.05 were considered to be statistically different.

KEGG and GO enrichment analysis of DEGs. A comprehensive set of functional annotation tools were provided by the Database for Annotation, Visualization, and Integrated Discovery (DAVID; http://david.ncifcrf.gov) (version 6.8). DAVID is an online biological information database for investigators to understand biological significance underlying a large number of genes (11). KEGG (http://www.kegg.jp), an integrated database resource, is used for the biological interpretation of genome sequences and other high-throughput data (12). The GO (www.geneontology.org) project is a major bioinformatics tool and represents the most comprehensive resource currently available for computable knowledge regarding the functions of genes and gene products (13). Enrichment analysis from GO and KEGG pathways for differentially expressed genes was obtained using DAVID. P<0.05 was considered to indicate a statistically significant difference.

PPI network construction and analysis. A PPI network of DEGs was constructed using the Search Tool for the Retrieval of Interacting Genes (STRING) online database (version 10.5; http://string-db.org) (14). Through the STRING database, DEGs with a combined score ≥0.4 were chosen to construct a PPI network which could be visualized using Cytoscape software (version 3.4.0; www.cytoscape.org) (15). The functional modules of the PPI network were then identified using the Molecular Complex Detection (MCODE) (version 1.4.2) plug-in of Cytoscape (16). The criteria for selection were as follows: Max depth, 100; degree cut-off, 2; k-score, 2 and node score cut-off, 0.2.

Hub gene selection and analysis. Hub genes were selected using Cytoscape software. A network of hub genes and their co-expressed genes was analyzed using the cBioPortal for Cancer Genomics (http://www.cbioportal.org) (17,18), which allows for visualization, analysis, and download of large-scale cancer genomics data sets. Hierarchical clusters of hub genes were constructed using the next generation University of California Santa Cruz (UCSC) Cancer Browser: UCSC Xena (http://xena.ucsc.edu) (19). The sample source ‘The Cancer Genome Atlas Head-Neck Squamous Cell Carcinoma (HNSC)’ was selected for these 26 hub genetic analyses and 604 samples were selected for analysis. The overall survival and disease-free survival rate analyses of hub genes was performed by constructing Kaplan-Meier curves using the cBioPortal online platform (statistical analysis performed is a log-rank test). Furthermore, the relationship between expression patterns, tumor grades, and HPV infection status was analyzed using Oncomine (https://www.oncomine.org) (20-29).

Results

Identification and PPI network construction of DEGs in HNC. After the standardization of the microarray results, 648 differentially expressed genes were identified between HNC tissues and normal tissues. The results from the GSE58911 dataset are represented as a volcano plot (Fig. 1A). The PPI network of DEGs was constructed (Fig. 1B). There were 554 nodes and 1574 edges in the PPI network, and the average node score was 5.68 (Fig. 1B).

GO and KEGG pathway enrichment analyses of DEGs. To analyze the biological classification of DEGs, GO and KEGG pathway enrichment analyses were performed using DAVID (Table I). The results of GO analysis showed that changes in biological processes of DEGs were mainly enriched in ‘enriched in muscle system process’, ‘extracellular matrix organization’, ‘muscle contraction’, ‘extracellular structure organization’, and ‘muscle filament sliding’. Molecular function DEGs included ‘actin binding’, ‘structural constituent of muscle’, ‘cytoskeletal protein binding’, ‘structural molecule activity’ and ‘actinin binding’. Cell component DEGs included ‘extracellular region part’, ‘contractile fiber’, ‘extracellular region’, ‘sarcomere’, and ‘myofibril’. KEGG pathway analysis showed that the DEGs were mainly enriched in ‘extracellular matrix (ECM)-receptor interaction’, ‘focal adhesion’, ‘amoebiasis’, ‘drug metabolism-cytochrome P450’, ‘chemical carcinogenesis’, ‘dilated cardiomyopathy’, ‘small cell lung cancer’, ‘hypertrophic cardiomyopathy’, and ‘retinol metabolism’.

Hub gene selection and analysis. Using the MCODE plug-in of Cytoscape, 26 genes were identified as hub genes. The results of GO and KEGG pathway analyses indicated that the hub genes were mainly enriched in ‘extracellular matrix organization’, ‘collagen catabolic process’, ‘extracellular structure organization’, ‘multicellular organism catabolic process’, ‘collagen metabolic process’, ‘serine-type endopeptidase activity’, ‘extracellular matrix’, ‘proteinaceous extracellular matrix’, ‘extracellular space’, ‘extracellular region part’, ‘extracellular region’, and ‘complement and coagulation cascades’
(Table II). The abbreviations, official full names, and synonyms for these hub genes are shown in Table III. A network of the hub genes and their co-expressed genes was analyzed using cBioPortal for Cancer Genomics (Fig. 2A). Hierarchical clustering revealed that the expression of hub genes could differentiate the HNC samples from normal samples (Fig. 2B). From figure 2B, it can be seen that 22 of the 26 hub genes were highly expressed in head and neck tumors compared with normal tissues, whereas expression of four genes (MMRN1/ECM1/EXCL12/CFD) was relatively high in the normal tissues. Furthermore, hierarchical clustering showed that HPV infection status determined by fluorescent in situ hybridization (FISH) testing (Fig. 2C) and P16 testing (Fig. 2D) was negatively associated with expression of the gene, although the mechanisms remains unknown. Overall survival rate analysis of the hub genes was performed using Kaplan-Meier
curves in the cBioPortal online platform. Patients with HNC and high expression of interleukin (IL)8, IL1B and serpin family A member 1 (SERPINA1) had worse overall survival and worse disease-free survival (Fig. 3A and B). Oncomine analysis of cancer vs. normal tissues indicated that IL8, IL1B, and SERPINA1 were over-expressed in HNC in the different datasets (Fig. 4A, B and C). Higher mRNA expression levels of IL8 was associated with tumor grade (P=0.001). However,
the mRNA expression levels of IL1B and SERPINA1 were not associated with tumor grade (P>0.05). Higher mRNA expression levels of IL8 (P=6.30x10^{-9}) and IL1B (P=3.48x10^{-6}) were associated with HPV infection status. The mRNA expression levels of SERPINA1 however, were not associated with HPV infection status (Fig. 5A-F).

**Discussion**

HNC is the sixth most common cancer worldwide and is associated with severe disease- and treatment-related morbidity, with a 5-year survival rate of <60% (1,2). The survival rate has not improved across more than two decades due to lack of early detection (30,31). There are two primary causes of HNC: Tobacco and alcohol use, and human papilloma-virus infection (32). Previous studies have shown that HPV infection plays a role in the pathogenesis of head and neck tumors (33-35). The oncomine analysis of cancer vs. normal tissue for IL8, IL1B and SERPINA1 demonstrated that the expression was compared with the normal tissues (Fig. 4). mRNA expressions of IL8, IL1B and SERPINA1 were higher in the HPV-negative group compared with that in the HPV-positive group (Fig. 5D-F). These seemingly contradictory results are understandable. There are many studies suggesting that HPV-positive head and neck tumors were associated with improved disease-free and overall survival (32,36). According to the present study, IL8, IL1B and SERPINA1 are highly expressed in the HNC, and the results in Fig. 3 indicate that the overall survival rate and disease-free survival rate of patients with high expression of these 3 genes are worse. Therefore, it is understandable that IL8, IL1B and SERPINA1 have higher expressions in HPV-negative tumor patients. However, the underlying molecular mechanisms of HNC remain unclear. Abnormal expression of transglutaminase 3, regenerating islet-derived protein 3, keratin 8, and phosphatase and tensin homolog is associated with HNC (37-40). In

---

**Table III. Abbreviations, official full names and synonyms for the 26 hub genes.**

| Number | Gene symbol | Official full name | Also known as |
|--------|-------------|-------------------|---------------|
| 1      | TNC         | Tenascin C        | GP; JI; TN; HXB; GMEM; TN-C; DFNA56; 150-225 |
| 2      | PLOD2       | Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2 | LH2; TLH; BRKS2 |
| 3      | COL3A1      | Collagen type III alpha 1 chain | ED54A; ED5VASC |
| 4      | IL1B        | Interleukin 1 beta | IL-1; ILF2; IL1-BETA |
| 5      | COL12A1     | Collagen type XII alpha 1 chain | UCMD2; BTHLM2; ED5MYP; COL12A1L; BA209D5;1; DJ234P15;1 |
| 6      | COL5A2      | Collagen type V alpha 2 chain | ED5C; ED5CL2 |
| 7      | MMRN1       | Multimerin 1      | ECM; MMRN; GPla; EMILIN4 |
| 8      | IL8         | C-X-C motif chemokine ligand 8 | NAF; GCP1; LECT; LUCT; NAP1; GCP-1; LYNAP; MDNCF; MONAP; NAP-1 |
| 9      | PLAU        | Plasminogen activator, urokinase | CD87; UPAR; URKR; U-PAR |
| 10     | PLAUR       | Plasminogen activator urokinase receptor | CD87; UPAR; URKR; U-PAR |
| 11     | ECM1        | Extracellular matrix protein 1 | URBWD |
| 12     | SERPINE1    | Serpin family E member 1 | PAI; PAI1; PAI-1; PLANH1 |
| 13     | MMP9        | Matrix metalloproteinase 9 | GELB; CLG4B; MMP-9; MANDP2 |
| 14     | TIMP1       | TIMP metalloproteinase inhibitor 1 | EPA; EPO; HCL; CLGI; TIMP; TIMP-1 |
| 15     | COL10A1     | Collagen type A alpha 1 chain | IRH; PBSF; SDF1; TLSF; TPAR1; SCYB12 |
| 16     | CXCL12      | C-X-C motif chemokine ligand 12 | DYT27; UCM1D1; BTHLM1 |
| 17     | COL6A3      | Collagen type VI alpha 3 chain | CLGI; CLGN |
| 18     | MMP1        | Matrix metalloproteinase 1 | COX2; COX-2; PHS-2; PGG-HS; PGHS-2; hCox-2; GRIPGHS |
| 19     | PTGS2       | Prostaglandin-endoperoxide synthase 2 | PI; A1A; AAT; PI1; A1AT; nNIF; PRO2275; alpha1AT |
| 20     | PLOD1       | Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 | LH; LHI; LLH; ED56; PLOD; ED5KCL1 |
| 21     | MMP13       | Matrix metalloproteinase 13 | CLG3; MSDT; MANDP1; MPP-13 |
| 22     | SPP1        | Secreted phosphoprotein 1pr | OPN; BNSP; BSPI; ETA-1 |
| 23     | SERPINA1    | Serpin family A member 1 | PI; A1A; AAT; PI1; A1AT; nNIF; PRO2275; alpha1AT |
| 24     | CFD         | Complement factor D | DF; ADN; PFD; ADPSIN |
| 25     | MMP3        | Matrix metalloproteinase 3 | SL-1; STR1; CHDS6; MPP-3; STMY1 |
| 26     | SPARC       | Secreted protein acidic and cysteine rich | OI17; BM-40 |
addition, mutations within tumor protein p53, notch receptor 1, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit α, X-ray repair cross complementing 1 and epidermal growth factor receptor have been reported to be involved in HNC (41-44). Patients with HNC that do not detect the cancer early have no effective treatments available except receiving palliative care, which leads to poor prognosis and quality of life, and a high rate of mortality (30). Therefore, there is an urgent need to identify potential target biomarkers that can be used to efficiently diagnose and treat HNC. Bioinformatics technology allows us to explore genetic differences between HNC and normal tissues, which can be used to identify potential biomarkers. Then, effective genes can be selected through screening and experimental validation for early diagnosis, clinical prognosis, and treatment of HNC.

In the present study, the dataset GSE58911 was analyzed to obtain differentially expressed genes between HNC and non-cancerous tissues. A total of 648 DEGs were identified.
GO and KEGG enrichment analyses were performed to explore interactions among these genes and they were mainly enriched in ‘extracellular matrix organization’, ‘actin binding’, ‘extracellular region’, ‘ECM-receptor interaction’, ‘drug metabolism-cytochrome P450’, and ‘chemical carcinogenesis’. Previous studies reported that ‘extracellular matrix organization’, ‘actin binding’, and ‘ECM-receptor interaction’ play important roles in the carcinogenesis, progression, and metastasis of tumors (45-48). In addition, previous data indicated that focal adhesion, drug metabolism-cytochrome P450, and chemical carcinogenesis are involved in radio- and chemotherapy (49‑52). Thus, the findings from the present study are consistent with results from previous studies. GO enrichment analysis indicated that changes in hub genes were mainly enriched in ‘extracellular matrix organization’, ‘collagen catabolic process’, ‘serine-type endopeptidase activity’, ‘extracellular matrix’, and ‘proteinaceous extracellular matrix’, while changes according to KEGG pathway analysis were mainly enriched in ‘complement and coagulation cascades’.

A total of 26 DEGs were selected as hub genes, among which survival rates and disease-free survival rates between patients with head and neck tumors and patients without tumors were significantly associated with the expression of IL8, IL1B, and SERPINA1. IL8 is a chemotactic factor that attracts neutrophils, basophils, and T-cells, but not monocytes (53) and can be released by several cell types in response to inflammatory stimuli (53). Higher IL8 expression was observed in HNSCC tissue (54,55). Furthermore, IL8 stimulated the proliferation of HNSC cells (55,56). In addition, a previous study showed that the tumor microenvironment plays a vital role in HNC initiation, progression, and metastasis (57). Tumor-associated macrophages can promote cancer initiation and progression by releasing cytokines and may facilitate papillary thyroid carcinoma (PTC) cell metastasis through IL8 and its paracrine interaction with C-X-C chemokine receptor CXCR1 and CXCR2 (58). Thus, IL8 may be a potential therapeutic target.

IL1B is a potent pro-inflammatory cytokine. Initially discovered as the major endogenous pyrogen, IL1B induces prostaglandin synthesis, neutrophil influx and activation, cytokine production, T cell and B cell activation, antibody production, collagen production, and fibroblast proliferation (59). A recent study of IL1B has shown that it plays a major role in tumor chemotherapy resistance. Anakinra can block the IL-1 pathway and overcome erlotinib resistance in HNSCC, which may represent a novel strategy to overcome EGFR inhibitor resistance, allowing for more effective treatment of patients with...
Furthermore, high expression of inflammatory cytokines (IL8, IL1B) and shorter progression-free survival are significantly associated. The expression level of inflammatory cytokines may help to identify which patients with recurrent and/or metastatic squamous cell carcinoma of the head and neck are likely to benefit from dacomitinib (61).

SERPINA1, an inhibitor of serine proteases, irreversibly inhibits trypsin, chymotrypsin, and plasminogen activator (62). Its primary target is elastase, but it also has a moderate affinity for plasmin and thrombin (62). A recent study showed a higher abundance of SERPINA1 candidate biomarkers in the saliva of patients with oral squamous cell carcinoma (OSCC), demonstrating that SERPINA1 is related to OSCC development (63). Moreover, SERPINA1 may be related to PTC by responding to steroid hormone stimuli and regulating the epithelial-to-mesenchymal transition (64). Based on these associations, SERPINA1 may be an effective mRNA marker of PTC (65). In addition to IL1B, IL8 and SERPINA1, which were associated with the survival rate of patients with head and neck cancer, other relevant hub genes that were identified in the present study are discussed.

Tenascin C (TNC), a gene associated with tumor metastatic potential, was upregulated in the OSCC cell line LNMTca8113 (66). In addition, vascular density and higher tumor stage were associated with differences in immuno-expression of stromal TNC, demonstrating its role in the tumorigenesis of juvenile nasopharyngeal angiofibroma (67).

A previous study showed that microRNA-29a/b could regulate the expression of collagen type III alpha 1 chain to enhance migration and invasion ability of nasopharyngeal carcinoma cells (68). The markers, the combination of collagen type V alpha 1 chain (COL5A1) and hemoglobin subunit beta and COL5A1 itself can better predict the treatment response in patients with oral tongue squamous cell carcinoma (69). Poor disease-free survival and increased progression or relapse risk were associated with high plasminogen activator, urokinase (PLAU) expression. Moreover, circulating PLAU

Figure 4. Oncomine analysis of cancer vs. normal tissue for IL8, IL1B and SERPINA1. Heat maps of IL8, IL1B, and SERPINA1 gene expression in clinical HNC samples vs. normal tissues. Red represents high expression and blue represents low expression and the color reflects the median rank, not the expression value. (A) 1. Head and neck squamous cell carcinoma vs. normal tissue (21). 2. Tongue squamous cell carcinoma vs. normal tissue (22). 3. Head and neck squamous cell carcinoma vs. normal tissue (23). 4. Thyroid gland papillary carcinoma vs. normal tissue (24). 5. Oral cavity squamous cell carcinoma vs. normal tissue (25). 6. Floor of the mouth carcinoma vs. normal tissue (26). 7. Oral cavity carcinoma vs. normal tissue (26). 8. Oropharyngeal carcinoma vs. normal tissue (26). 9. Tongue carcinoma vs. normal tissue (26). 10. Tonsillar carcinoma vs. normal tissue (26). 11. Tongue squamous cell carcinoma vs. normal tissue (27). (B) 1. Head and neck squamous cell carcinoma vs. normal tissue (21). 2. Tongue squamous cell carcinoma vs. normal tissue (22). 3. Head and neck squamous cell carcinoma vs. normal tissue (23). 4. Oral cavity squamous cell carcinoma vs. normal tissue (23). 5. Tongue squamous cell carcinoma vs. normal tissue (27). (C) 1. Head and neck squamous cell carcinoma vs. normal tissue (21). 2. Tongue squamous cell carcinoma vs. normal tissue (22). 3. Head and neck squamous cell carcinoma vs. normal tissue (23). 4. Oral cavity squamous cell carcinoma vs. normal tissue (23). 5. Tongue squamous cell carcinoma vs. normal tissue (27). 2. Tall cell variant thyroid gland papillary carcinoma vs. normal tissue (26). 3. Thyroid gland papillary carcinoma vs. normal tissue (28). 4. Thyroid gland papillary carcinoma vs. normal tissue (28). 5. Oral cavity squamous cell carcinoma vs. normal tissue (28). 6. Oropharyngeal carcinoma vs. normal tissue (28). 7. Thyroid gland papillary carcinoma vs. normal tissue (28).
levels were significantly higher in the plasma of patients with HNSCC compared with that in healthy individuals (70).

Extracellular matrix protein 1 (ECM1) levels gradually increased from benign laryngeal lesions to precancerous to malignant lesions, and ECM1 was expressed at lower levels in laryngeal carcinomas without metastasis (71,72). These results demonstrated that ECM1 facilitated development and metastasis of laryngeal carcinoma.

Overexpression of SERPINE1 promotes tumor migration and invasion and plays an important role in metastasis and poor prognosis of HNSCC (73). In addition, many researchers regard SERPINE1 as a prognostic marker based on its ability to stratify patients with HNSCC according to their recurrence risk (74).

Matrix metalloproteinases (MMPs) are a family of proteolytic enzymes that promote invasion and metastasis of various cancers due to their ability to degrade components of the extracellular matrix. MMP1, MMP3, MMP9, and MMP13 are predictors of poor clinical outcomes in patients with HNC (75-78). Furthermore, specific tissue inhibitors of matrix metalloproteinases (TIMPs) can regulate MMP activity. In addition to HNC, the majority of tumors are associated with alterations in MMPs and TIMPs. Imbalance between matrix metalloproteinases and their inhibitors contributes greatly to the progression and prognosis of HNC (76,79).

Compared with normal oral mucosa, secreted phosphoprotein 1 was expressed at significantly higher levels in OSCC (80). According to a previous study, secreted protein acidic and cysteine rich had significant prognostic value, especially in the stroma surrounding OSCC (81). A literature search revealed that the interaction between HNC and the hub genes procollagen-lysine 2-oxoglutarate 5-dioxygenase (PLOD)-2, collagen type XII alpha 1 chain, multimerin 1 (MMRN1), plasminogen activator urokinase receptor, collagen type X α 1 chain, collagen type VI α 3 chain, prostaglandin-endoperoxide synthase 2, PLOD1, and complement factor D (CFD) has not been widely reported.

Figure 5. Association between the expression of IL8, IL1B, and SERPINA1, tumor grade, and HPV infection status in the different datasets. The datasets were obtained from the Oncomine database. The box plot represents the maximum, the 75th percentile, the median, the 25th percentile and the minimum value of data. (A-C) IL8, IL1B and SERPINA1 mRNA expression of HNC samples in the Cromer Head and Neck dataset. There were four patients in group 0, representing with no grade; seven patients in group 1, representing Grade 1; nineteen patients in group 2, representing Grade 2; and eight patients in group 3, representing Grade 3. (D-F) IL8, IL1B, and SERPINA1 mRNA expression of HNC samples in the Slebos Head and Neck dataset. There were twenty-eight patients in group 1, representing HPV negative; and eight patients in group 2, representing HPV positive. IL, interleukin; SERPINA1, serpin family A member 1; HPV, human papillomavirus, HNC, head and neck cancer.
There were several limitations associated with the present study. First, only one series (GSE58911) downloaded and used from the GEO database. The number of tumor and normal samples in this series were both 15. This sample size was insufficient. Second, genes were analyzed that may be related to the carcinogenesis or progression of head and neck tumors from the results of the bioinformatics analyses. The functions of these genes have not been verified in vitro and in vivo. In the present study, the expression levels of IL8 (C-X-C motif chemokine ligand 8), IL1B and SERPINA1 in tumor tissues and normal tissues of patients with head and neck tumors were not verified further. In addition, phenotypic function was also not verified in head and neck tumor cell lines. In future studies this will be investigated. Third, the number of hub genes (modes with bold black circles) in Fig. 2A is 21, and there are 5 hub genes (MMPR1, EMT1, TIMP metallopeptidase inhibitor 1, SERPINA1 and CFD) that do not appear in the network map. In particular, SERPINA1 is among one of the identified three genes following further analysis of the data. This could be due to the following reasons: i) These 5 genes may not be closely related to other genes, and have other roles and mechanisms in the occurrence and development of tumors, so they were excluded from the network map; ii) the 26 hub genes were obtained by analyzing 648 DEGs using the Cytoscape plug-in, MCODE. cBioPortal for Cancer Genomics was subsequently used to analyze these 26 hub genes to obtain a network map of hub genes and their co-expressed genes. The computer algorithms that each database performed for analysis may differ, and may also cause differences; iii) in addition, 3 hub genes (IL8, IL1B and SERPINA1) were selected for more in-depth analysis as cBioPortal for Cancer Genomics was used to analyze overall survival and disease-free survival rate for the 26 hub genes. Changes in the expression of IL8, IL1B and SERPINA1 in patients with head and neck tumors were associated with overall survival and disease-free survival rate, and were statistically significant; iv), a holistic analysis of 26 hub genes from 628 DEGs was performed in an attempt to obtain an inductive result (Fig. 2). In addition, 2 detection methods (FISH and P16 tests) of HPV infection were used to obtain more accurate results for further analysis. However, the profiles of IL8, IL1B and SERPINA1 in Fig. 2C and D are not consistent. This may be due to different detection methods or detection of HPV subtypes. This may require a more accurate method to detect the infection status of HPV for a more accurate analysis; and v), there are differences in the monitoring of overall survival and disease-free survival of patients between the gene alterations of IL8, IL1B and SERPINA1 (Fig. 3). Monitoring was performed for over 180 months (over 15 years) when there were no alteration(s) in these 3 genes while alterations in these genes was monitored for only 20 months. This may be due to the following reasons: i) These samples were obtained at different time points over a 10-year period and as such the samples may have degraded; ii) some patients cannot be contacted during follow-up or have died due to illness. An open source, free database was used therefore information pertaining to when the samples were collected and whether it was in different decades. In this regard, the lack of data beyond 20 months may be considered a limitation associated with the present study. However, using this database for analysis is reliable and credible. This database is used in many articles on bioinformatics analysis (82-84). These limitations will be addressed in future studies. Despite these limitations, research in the present study is important as it elucidated molecular mechanisms underlying development of HNC, and also provides potential target genes for clinical diagnosis and targeted therapy. In addition, this provides direction for future studies of HNC.

In conclusion, the present study identified DEGs that may be involved in the carcinogenesis or progression of HNC. A total of 648 DEGs and 26 hub genes were identified and may have potential as target biomarkers for HNC. Further studies are needed to elucidate the biological functions of these genes in HNC.

Acknowledgements

Not applicable.

Funding

This study was supported by the National Natural Science Foundation of China (grant nos. 81372880 and 81670910) and the Guidance fund of the Renmin Hospital of Wuhan University (grant no. RMYD2018Z12).

Availability of data and materials

All data are fully available without restriction.

Authors’ contributions

FC and ZT conceived and designed the study. FC, AZ, FL, SW, SC and ZT processed the data. FC wrote the paper. FC, AZ, FL, SW, SC, and ZT reviewed and edited the manuscript. All authors read and approved the manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D: Global cancer statistics. CA Cancer J Clin 61: 69-90, 2011.
2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J and Jemal A: Global cancer statistics, 2012. CA Cancer J Clin 65: 87-108, 2015.
3. Bray F, Ferlay J, Laversanne M, Brewster DH, Gombe MC, Kohler B, Prieur M, Steliarova-Foucher E, Swaminathan R, Antoni S, et al: cancer incidence in five continents: Inclusion criteria, highlights from volume X and the global status of cancer registration. Int J Cancer 137: 2060-2071, 2015.
4. Thibodeau BJ, Geddes T, Fortier LE, Ahmed S, Pruett BL, Wobb J, Chen P, Wilson GD and Akervall JA: Gene expression characterization of HPV positive head and neck cancer to predict response to chemoradiation. Head Neck Pathol 9: 345-353, 2015.
5. Min SK, Lee SK, Park JS, Lee J, Paeng JY, Lee SI, Lee HJ, Kim Y, Paes HO, Lee SK and Kim EC: Endoplasmic reticulum stress is involved in hydrogen peroxide induced apoptosis in immortalized and malignant human oral keratinocytes. J Oral Pathol Med 37: 490-498, 2008.
6. Skvortsov S, Dudus J, Eicherger P, Witsch-Baumgartner M, Loeffler-Ragg J, Fritz C, Schartinger VH, Maier H, Hall J, Debbage P, et al: Rac1 as a potential therapeutic target for chemoradiotherapy-induced squamous cell carcinomas (HNSCC). Br J Cancer 110: 2677-2687, 2014.

7. Wang Y, Chen C, Wang X, Jin F, Liu Y, Liu H, Li T and Fu J: Lower DSCI expression is related to the poor differentiation and prognosis of head and neck squamous cell carcinoma (HNSCC). J Cancer Res Clin Oncol 142: 2461-2468, 2016.

8. Edgar R, Domrachev M and Lash AE: Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30: 207-210, 2002.

9. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashovsky M, Marshall KA, Philipphi KH, Sherman PM, Holmgren P, et al: DAVID at Genome Studio: a comprehensive functional annotation server. Nucleic Acids Res 38:D199-D210, 2010.

10. Eberhart CE, Roses DR, den Boon JA, Sengupta S, Marsit CJ, Woodworth CD, Connor JP, Haugen TH, Smith EM, et al: Fundamental differences in cell cycle deregulation in human papillomavirus-positive and human papillomavirus-negative head/neck and cervical cancers. Cancer Discov 2: 401-404, 2012.

11. Eichberger P, Witsch-Baumgartner M, Santini J, Jendrossek D, and Koganti T: Prognostic biomarkers for HNSCC using quantitative real-time PCR and microarray analysis: β-tubulin isoforms and the p53 interactome. Cytoskeleton (Hoboken) 71: 628-637, 2014.

12. Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Smith EM, et al: Fundamental differences in cell cycle deregulation in human papillomavirus-positive and human papillomavirus-negative head/neck and cervical cancers. Cancer Discov 2: 401-404, 2012.

13. Mekhitarian K, Zhai Y, Zhou Q, Xu X, Jiang H, Ma Y, et al: Identification of a gene expression signature associated with head and neck squamous cell carcinoma. BMC Genomics 9: 69, 2008.

14. Giordano TJ, Assay AY, Kuick R, Thomas DG, Rhodes DR, Wilhelm KJ Jr, Vinco M, Misek DE, Sanders D, Zhu Z, et al: Delineation, functional validation, and bioinformatic evaluation of gene expression in thyroid follicular carcinomas with the PAX8-PPARG translocation. Clin Cancer Res 12: 1983-1993, 2006.

15. Roayaei J, Stephens R, Baseler MW, Lane HC and Lempicki RA: The cBioPortal for comparative cancer genomics and clinical profiles using the cBio Cancer Genomics Portal. Sci Signal 6: pl1, 2013.

16. Sun Y, Jacobsen A, Sinha R, Larsson E, Bandettini WP, Kellman P, Mancini C, Booker OJ, Vasu S, Kleinman N, et al: The UCSC Genome Browser database: 2015 update. Nucleic Acids Res 43(D1): D670-D681, 2015.

17. Wang Y, Chen C, Wang X, Jin F, Liu Y, Liu H, Li T and Fu J: Lower DSCI expression is related to the poor differentiation and prognosis of head and neck squamous cell carcinoma (HNSCC). J Cancer Res Clin Oncol 142: 2461-2468, 2016.

18. Garnier J, Crucès G, Escobas M, Giraud D, Laviolle B, Mangeat P, Dumas B, et al: STRING v10: Protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43: D670-D681, 2015.

19. Eichberger P, Witsch-Baumgartner M, Santini J, Jendrossek D, and Koganti T: Prognostic biomarkers for HNSCC using quantitative real-time PCR and microarray analysis: β-tubulin isoforms and the p53 interactome. Cytoskeleton (Hoboken) 71: 628-637, 2014.

20. Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Smith EM, et al: Fundamental differences in cell cycle deregulation in human papillomavirus-positive and human papillomavirus-negative head/neck and cervical cancers. Cancer Discov 2: 401-404, 2012.

21. Mekhitarian K, Zhai Y, Zhou Q, Xu X, Jiang H, Ma Y, et al: Identification of a gene expression signature associated with head and neck squamous cell carcinoma. BMC Genomics 9: 69, 2008.

22. Giordano TJ, Assay AY, Kuick R, Thomas DG, Rhodes DR, Wilhelm KJ Jr, Vinco M, Misek DE, Sanders D, Zhu Z, et al: Delineation, functional validation, and bioinformatic evaluation of gene expression in thyroid follicular carcinomas with the PAX8-PPARG translocation. Clin Cancer Res 12: 1983-1993, 2006.

23. Roayaei J, Stephens R, Baseler MW, Lane HC and Lempicki RA: The cBioPortal for comparative cancer genomics and clinical profiles using the cBio Cancer Genomics Portal. Sci Signal 6: pl1, 2013.

24. Sun Y, Jacobsen A, Sinha R, Larsson E, Bandettini WP, Kellman P, Mancini C, Booker OJ, Vasu S, Kleinman N, et al: The UCSC Genome Browser database: 2015 update. Nucleic Acids Res 43(D1): D670-D681, 2015.

25. Wang Y, Chen C, Wang X, Jin F, Liu Y, Liu H, Li T and Fu J: Lower DSCI expression is related to the poor differentiation and prognosis of head and neck squamous cell carcinoma (HNSCC). J Cancer Res Clin Oncol 142: 2461-2468, 2016.

26. Sun Y, Jacobsen A, Sinha R, Larsson E, Bandettini WP, Kellman P, Mancini C, Booker OJ, Vasu S, Kleinman N, et al: The UCSC Genome Browser database: 2015 update. Nucleic Acids Res 43(D1): D670-D681, 2015.

27. Wang Y, Chen C, Wang X, Jin F, Liu Y, Liu H, Li T and Fu J: Lower DSCI expression is related to the poor differentiation and prognosis of head and neck squamous cell carcinoma (HNSCC). J Cancer Res Clin Oncol 142: 2461-2468, 2016.

28. Wang Y, Chen C, Wang X, Jin F, Liu Y, Liu H, Li T and Fu J: Lower DSCI expression is related to the poor differentiation and prognosis of head and neck squamous cell carcinoma (HNSCC). J Cancer Res Clin Oncol 142: 2461-2468, 2016.
40. Boeckx C, Weyn C, Vanden Bempt I, Deschoolmeester V, Wouters A, Specenier P, Van Laer C, Van den Weyngaert D, Kocks M, Vermorken JB, et al: Mutation analysis of genes in the EGFR pathway in head and neck cancer patients. Implications for anti-EGFR treatment response. BMC Res Notes 7: 337. 2014.

41. Gilkes DM, Semenza GL and Wirtz D: Focal adhesion signaling and therapy resistance in nasopharyngeal carcinoma progression by regulating SPARC and COL3A1 gene expression. PLoS One 10: e0120969, 2015.

42. Suresh A, Vannan M, Kumaran D, Gümüs ZH, Sivadas P, Muruganuk E, Keetakpure V, Iyer S, Thangaraj K and Kurtakose MA: Resistant cells over-express more for oral tongue squamous cell carcinoma. Dis Markers 32(5): 51-64, 2012.

43. Peşişhavili L, Hui A, Ignatchenko V, Shi W, Su S, Xu W, Huang SH, O'Sullivan B, Waldron J, Irish JC, et al: Potentially novel candidate biomarkers for head and neck squamous cell carcinoma identified using an integrated cell line-based discovery strategy. Mol Cell Proteomics 11: 1404-1415, 2012.

44. Gu M, Guan J, Zhao L, Ni K, Li X and Han Z: Correlation of ECM1 expression level with the pathogenesis and metastasis of laryngeal carcinoma. Int J Clin Exp Pathol 6: 1132-1137, 2013.

45. Meng XY, Liu J, Lv F, Liu MQ and Wan JN: Study on the correlation between extracellular matrix proteins interleukin 8 and beta-thromboglobulin: In vitro and in vivo comparison of NH2-terminally processed forms. Eur J Immunol 20: 2113-2118, 1990.

46. Johnson AL, Edson KZ, Totah RA and Rettie AE: Cytochrome P450 o-hydroxylases in inflammation and cancer. Adv Cancer Res 73: 273-262, 2005. 2012.

47. Trulsson M, Hu H, Gisselsson L, Chao Y, Aittis S, Mossberg AK and Svanborg C: HAMLET binding to α-helix bundles triggers apoptosis. PLoS One 6: e17179, 2011.

48. Zhang HJ, Tao J, Sheng L, Hu X, Dong RM, Xu M and Zhu Y: Twist2 promotes kidney cancer cell proliferation and invasion by regulating ITGA6 and CD44 expression in the ECM-receptor interaction pathway. Onco Targets Ther 9: 1801-1812, 2016.

49. Eke I and Cortes N: Focal adhesion signaling and therapy resistance in cancer. Semin Cancer Biol 31: 65‑75, 2015.

50. Blackstone BN, Li R, Ackerman WT, Ghadiiali SN, Powell HM and Knis DA: Myoferlin depletion elevates focal adhesion kinase and paxillin phosphorylation and enhances cell-matrix adhesion in breast cancer cells. Am J Physiol Cell Physiol 308: C642-C649, 2015.

51. Johnson AL, Edson KZ, Totah RA and Rettie AE: Cytochrome P450 o-hydroxylases in inflammation and cancer. Adv Cancer Res 73: 273-262, 2005. 2012.

52._rgb_img