Transcriptomic dissection of tongue squamous cell carcinoma

Hui Ye, University of Illinois at Chicago
Tianwei Yu, Emory University
Stephane Temam, University of Texas

Journal Title: BMC Genomics
Volume: Volume 9, Number 69
Publisher: BioMed Central | 2008
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1186/1471-2164-9-69
Permanent URL: http://pid.emory.edu/ark:/25593/fkj5z

Final published version: http://www.biomedcentral.com/1471-2164/9/69

Copyright information:

© 2008 Ye et al; licensee BioMed Central Ltd.

This is an Open Access work distributed under the terms of the Creative Commons Attribution 2.0 Generic License (http://creativecommons.org/licenses/by/2.0/).

Accessed October 1, 2023 4:04 AM EDT
Abstract

Background: The head and neck/oral squamous cell carcinoma (HNOSCC) is a diverse group of cancers, which develop from many different anatomic sites and are associated with different risk factors and genetic characteristics. The oral tongue squamous cell carcinoma (OTSCC) is one of the most common types of HNOSCC. It is significantly more aggressive than other forms of HNOSCC, in terms of local invasion and spread. In this study, we aim to identify specific transcriptomic signatures that are associated with OTSCC.

Results: Genome-wide transcriptomic profiles were obtained for 53 primary OTSCCs and 22 matching normal tissues. Genes that exhibit statistically significant differences in expression between OTSCCs and normal were identified. These include up-regulated genes (MMP1, MMP10, MMP3, MMP12, PTHLH, INHBA, LAMC2, IL8, KRT17, COL1A2, IF6, ISG15, PLAU, GREM1, MMP9, IFI44, CXCL1), and down-regulated genes (KRT4, MAL, CRNN, SCEL, CRISP3, SPINK5, CLCA4, ADH1B, P11, TGM3, RHCG, PPP1R3C, CEACAM7, HPGD, CFD, ABCA8, CLU, CYP3A5). The expression difference of IL8 and MMP9 were further validated by real-time quantitative RT-PCR and immunohistochemistry. The Gene Ontology analysis suggested a number of altered biological processes in OTSCCs, including enhancements in phosphate transport, collagen catabolism, l-kappaB kinase/NF-κappaB signaling cascade, extracellular matrix organization and biogenesis, chemotaxis, as well as suppressions of superoxide release, hydrogen peroxide metabolism, cellular response to hydrogen peroxide, keratinization, and keratinocyte differentiation in OTSCCs.

Conclusion: In summary, our study provided a transcriptomic signature for OTSCC that may lead to a diagnosis or screening tool and provide the foundation for further functional validation of these specific candidate genes for OTSCC.
Background
Head and neck/oral squamous cell carcinoma (HNOSCC) is a complex disease arising in various organs, including oral cavity, tongue, pharynx, and larynx. Tumors from these different sites have distinct clinical presentations and clinical outcomes, and are associated with different risk factors [1] and genetic characteristics [2]. In this study, we focused on the oral tongue squamous cell carcinomas (OTSCC), one of the most common sites for HNOSCCs. The incidence of OTSCC is actually increasing in young and middle age groups [3-5]. OTSCC is significantly more aggressive than other forms of HNOSCCs, with a propensity for rapid local invasion and spread [6].

Cancer cells harbor genetic alterations which are translated into unique expression patterns. These patterns may segregate cancer cells from normal tissue of the same origin and serve as a molecular biomarker. Moreover, expression pattern changes may occur far earlier than clinical disease detection. The identification of such patterns has significant translational values for early detection and diagnosis, as well as for identifying novel therapeutic targets. While several recent studies have attempted to identify expression patterns for HNOSCCs [7-10], to our knowledge, no study has been devoted to identify the unique expression pattern for OTSCC. In this study, we aim to identify the specific transcriptomic/expression patterns that associated with OTSCC.

Results and discussion
Genome-wide gene expression profiles were obtained on 53 OTSCC samples and 22 normal matching samples. Principal Component Analysis (PCA) was performed based on all the probesets utilized in our microarray analysis. Apparent separation between OTSCC and normal groups was observed with a few outliers (Figure 1). Genes showing statistically significant differences in expression level were identified using RMA and a mixed-effects model as described in the Materials and Methods section. A signature gene set that consists of 35 genes was created using stringent statistical criteria (fold change > 4, and FDR values < 0.0001) (Table 1). Comprehensive lists of genes showing statistically significant upregulations (fold change > 2, and FDR values < 0.01) or downregulations in expression in OTSCC were presented in Supplement Table S1 [see additional file 1] and S2 [see additional file 2], respectively.

In this study, we identified and validated several interesting potential biomarkers for OTSCC diagnosis. One interesting observation is that 5 members of the Matrix Metalloproteinase (MMP) family (MMP1, MMP3, MMP9, MMP10, and MMP12) are among the genes that most significantly upregulated, which may contribute to the aggressive nature of the OTSCC. MMPs are a large family of proteinases which remodel extracellular matrix (ECM) components and play a significant role in tumor development, survival, invasion and metastasis [11-13]. Several members of the MMP family have been considered to be important biomarkers for diagnosis and prognosis as well as potential therapeutic targets for many types of cancers, including HNOSCC [14]. Our recent study suggested that up regulation of the MMP9 gene is associated with advanced OTSCC and has predictive value for the identification of lymph node metastasis [15]. Here, our data further suggested that MMP9 is one of the biomarkers for the detection of OTSCC. We also observed 2 chemokines (IL8 and CXCL1) to be among the most significantly upregulated genes in OTSCC. The increase in the protein and mRNA of IL8 gene has been suggested as a biomarker for the early detection of oral cancer [16,17]. Our data provided independent validation for this biomarker at the disease tissue level, and suggested that the increase of IL8 molecules (mRNA and protein) is due at least in part to the increased expression of gene in the disease tissues. CXCL1, also known as growth-regulated oncogene 1 (Gro-1), is vital for the survival, progression and invasion of several cancer types [18,19], including oral cancer [20]. Our results here further confirmed the importance of CXCL1 in the tumorgenesis of tongue cancer. Other inter-
Testing observations include the up-regulation of KRT17 (associated with invasion and proliferation) and down-regulation of KRT4 (which is associated with squamous cell differentiation), suggesting the potentially distinct roles of KRT genes in tongue SCC development and progression. In addition to those identified biomarkers, our results will also serve as a valuable reference data set for future development and validation of biomarkers for detection, diagnosis and prognosis of tongue cancer.

To test the utility of this 35-gene signature gene set for classifying OTSCC and normal groups, average linkage hierarchical clustering analysis was performed. As illustrated in Figure 2, our results demonstrated that this 35-gene set provides classification power for OTSCC based on gene expression analyses, which misclassified two cases for each of two groups (a 95% overall accuracy rate).

Our analysis demonstrated that OTSCC can be identified based on the gene expression signature. This finding should provide a foundation for the creation of a specific screen tool for OTSCC. One of the major factors accounting for the poor outcome of OTSCC patients is that a great proportion of oral cancers are diagnosed at advanced stages. Patients diagnosed at an early stage of the disease typically have a better chance for cure and functional outcome. Early detection of tongue cancer lesion will greatly improve patient survival and the quality of life. Current clinical diagnosis and histopathologic examinations are usually based on biopsied material, which requires invasive procedures. However, the use of gene expression signatures for early diagnosis can provide a non-invasive and less invasive alternative.

| Symbol | Probe ID | Gene Name | Location |
|--------|----------|-----------|----------|
| MMP1   | 204475_at| matrix metalloproteinase 1 | 11q22.3 |
| MMP10  | 205680_at| matrix metalloproteinase 10 | 11q22.3 |
| MMP3   | 205828_at| matrix metalloproteinase 3 | 11q22.3 |
| MMP12  | 204580_at| matrix metalloproteinase 12 | 11q22.3 |
| PTHLH  | 211756_at| parathyroid hormone-like hormone | 12p12.1-p11.2 |
| INHBA  | 210511_s_at| inhibin, beta A | 7p15-p13 |
| LAMC2  | 202267_at| laminin, gamma 2 | 1q25-q31 |
| IL8    | 202859_x_at| interleukin 8 | 4q13-q21 |
| KRT17  | 205157_s_at| keratin 17 | 17q12-q21 |
| COL1A2 | 202404_s_at| collagen, type I, alpha 2 | 7q22.1 |
| IF6    | 204415_at| interferon, alpha-inducible protein 6 | 1p35 |
| ISG15  | 205483_s_at| ISG15 ubiquitin-like modifier | 1p36.33 |
| PLA2G7 | 205479_s_at| plasminogen activator, urokinase | 10q24 |
| GREM1  | 218468_s_at| gremlin 1 | 15q13-q15 |
| MMP9   | 203936_s_at| matrix metalloproteinase 9 | 20q1.2-q13.1 |
| IFI44  | 214453_s_at| interferon-induced protein 44 | 1p31.1 |
| CXCL1  | 204470_at| chemokine (C-X-C motif) ligand 1 | 4q21 |
| MMP1   | 204475_at| matrix metalloproteinase 1 | 11q22.3 |
| MMP10  | 205680_at| matrix metalloproteinase 10 | 11q22.3 |
| MMP3   | 205828_at| matrix metalloproteinase 3 | 11q22.3 |
| MMP12  | 204580_at| matrix metalloproteinase 12 | 11q22.3 |
| PTHLH  | 211756_at| parathyroid hormone-like hormone | 12p12.1-p11.2 |
| INHBA  | 210511_s_at| inhibin, beta A | 7p15-p13 |
| LAMC2  | 202267_at| laminin, gamma 2 | 1q25-q31 |
| IL8    | 202859_x_at| interleukin 8 | 4q13-q21 |
| KRT17  | 205157_s_at| keratin 17 | 17q12-q21 |
| COL1A2 | 202404_s_at| collagen, type I, alpha 2 | 7q22.1 |
| IF6    | 204415_at| interferon, alpha-inducible protein 6 | 1p35 |
| ISG15  | 205483_s_at| ISG15 ubiquitin-like modifier | 1p36.33 |
| PLA2G7 | 205479_s_at| plasminogen activator, urokinase | 10q24 |
| GREM1  | 218468_s_at| gremlin 1 | 15q13-q15 |
| MMP9   | 203936_s_at| matrix metalloproteinase 9 | 20q1.2-q13.1 |
| IFI44  | 214453_s_at| interferon-induced protein 44 | 1p31.1 |
| CXCL1  | 204470_at| chemokine (C-X-C motif) ligand 1 | 4q21 |

* The complete lists of upregulated and downregulated transcripts in OTSCCs were presented in Supplement Table S1 and S2, respectively.
The emerging technology of saliva-based diagnosis may provide an alternative strategy for early diagnosis and screening of the subjects at risk [21]. The markers identified here may be suitable for the saliva-based early diagnosis and screening strategy [17,21]. Additional validation studies will be needed to fully explore this possibility.

To gain a better understanding of the underlying molecular biological processes that dictate the observed expression pattern alterations in OTSCC development, the Gene Ontology analysis was performed. It was carried out using the GOstats package in the Bioconductor [22] and Gene Ontology Consortium database [23], based on the complete list of 365 differentially expressed transcripts (Supplement Table S1 and S2). Among those 365 genes, 306 were mapped to ENTREZ genes. The gene universe in the analysis consists of 13125 transcripts that were mapped to ENTREZ genes. The Gene Ontology analysis suggested a number of altered biological processes in OTSCC. These include enhancement in phosphate transport, collagen catabolism, I-kappaB kinase/NF-kappaB signaling cascade, extracellular matrix organization and biogenesis, chemotaxis, as well as suppression of superoxide release, hydrogen peroxide metabolism, cellular response to hydrogen peroxide, keratinization, keratinocyte differentiation in OTSCCs (Table 2). The complete lists of enhanced and suppressed biological processes, molecular functions and cellular components in OTSCCs were presented in Supplement Table S3 [see additional file 3] and S4 [see additional file 4], respectively.

Among the identified alteration in biological activities in OTSCC, the most significantly enhanced are related to the
extracellular matrix remodeling (GO:0030574, GO:0030198), I-kappaB kinase/NF-kappaB cascade (GO:0043123) and chemotaxis (GO:0050921), which are known to be related to tumorgenesis and progression of the cancer. One interesting observation is the enhancement in phosphate transport (GO:0006817) in OTSCC. This may be related directly to the enhanced metabolic activity and energy consumption rate in OTSCCs. It has also been suggested that phosphate can act as a signaling molecule on the extracellular signal-regulated kinase (ERK1/2) [24] and adenylate cyclase/cAMP signaling pathways [25], and ultimately affect cell growth. However, the precise role of enhanced phosphate transport in tumorgenesis is largely unclear. The significantly suppressed biological activities, such as superoxide release (GO:0042554), hydrogen peroxide metabolism (GO:0042743), and response to hydrogen peroxide (GO:0042542) are all appeared to be related to the cellular redox state. The effects of redox state in malignancies are somewhat contradictory. In theory, reducing the oxidative stress may prevent DNA degeneration and therefore prevent the development of cancer. However, doing so may also offer increased growth potential to tumor cells and protect them from excess of reactive oxygen species (ROS), which would otherwise lead to apoptosis or necrosis. At the center of this apparent controversy is superoxide dismutase 2 (SOD2), which has been considered as one of the most important antioxidant enzymes. The role of SODs in carcinogenesis has been widely studied but is still rather ambiguous. While the majority of in vitro studies have reported a protective role of SOD2 against tumor progression in cancer cell lines [26-30], including oral cancer cell lines [31], the in vivo studies indicate more complicated roles. Increased SOD2 levels have been observed from esophageal, gastric, brain astrocytic and colorectal carcinomas, and often associated with metastasis and poor prognosis [32-40]. The status of SOD2 in breast cancer is not clear, with some studies showing an increase [41], while others showing a decrease in SOD2 level [42]. Reduction in SOD2 level has been observed in prostatic carcinomas [43,44]. Our microarray results indicated a significant increase in expression of SOD2 gene (probeset: 215223_s_at; fold change = 2.37; p value = 0.0014; and probeset: 216841_s_at; fold change = 2.24; p value = 0.000197) in OTSCC. These findings are in agreement with the recent observation in oral cancer [45]. Additional studies will be needed to fully understand the role(s) of redox state and SOD2 in OTSCC.

To visualize the changes in gene expression patterns in relationship with the alteration of biological processes and cellular functions in OTSCC, gene expression heat maps for each identified GO entities were generated based on the microarray results of 53 OTSCC and 22 normal samples (Figure 3). Apparent differences in expression patterns can be observed between OTSCC and normal groups for all the altered GO entities. The summarized statistical values on differential expression for each individual gene of GO:0030574 (collagen catabolism) and GO:0050921 (positive regulation of chemotaxis) are presented in Table 3. The complete expressional analysis for all the altered biological processes identified in Table 2 is presented in Supplement Table S5 [see additional file 5].

Among these differentially expressed genes, some have potential value as diagnosis and prognosis markers, and may be indicative of their respective biological pathways. For example, IL8 is a prototypical chemokine (chemotactic cytokine) and is known for its involvements in the positive regulation of chemotaxis (GO:0050921), which is
Enhanced in tongue cancers as indicated by our Gene Ontology analysis. IL8 has also been suggested to be a potential biomarker for the early detection of oral cancer [16,17]. The over-expression of the MMP9 gene has been shown to be associated with progression of oral dysplasia to cancer [14]. Our recent study suggested that over-expression of the MMP9 gene is associated with advanced OTSCC and has predictive value for OTSCC lymph node metastasis [15]. The upregulation of MMP9 in OTSCC is also involved in the enhanced collagen catabolism (GO:0030574), as indicated by our Gene Ontology analysis. The qRT-PCR analyze were performed to validate the expressional differences of the IL8 and MMP9 genes between tongue SCCs and normal matching tissues. As shown in Figure 4A and 4B, the differences of both IL8 and MMP9 mRNA levels are statistically significant (p < .05) between OTSCC and normal tissues. The expression of MMP9 and IL8 was further confirmed by immunohis-
tochemistry tests performed using monoclonal antibody to IL8 and MMP9 on 10 OTSCC cases (Figure 4C and 4D). Strong positive stained SCC cells were observed in 4 cases for IL8 and 8 cases for MMP9. The observation of positive staining of IL8 in OTSCC cells confirmed that SCC cells are one of the major sources of IL8 production at the site of oral cancer lesion.

**Conclusion**

The HNOSCCs are a diverse group of cancers, that develop from many different anatomic sites and are associated with different risk factors [1] and genetic characteristics [2]. This is the first high-resolution genomic profiling study to our knowledge that has focused on identifying unique expression patterns for tongue cancer (OTSCC). OTSCC is one of the most common types of HNOSCC, and is significantly more aggressive than other forms of HNOSCCs, with a propensity for rapid local invasion and spread [6]. Recent epidemiological studies suggested that the incidence of OTSCC is actually increasing in young and middle age groups [3-5]. In this study, we utilized a relatively large sample size (53 OTSCCs and 22 normal matching samples), which enabled us to capture a precise picture of the genome-wide expression pattern for this disease. It is possible that the genomic portrait of HNOSCC originating from different anatomic sites may be different. More studies will be needed to address this important question.

In summary, we identified the unique expression pattern for OTSCC. Several interesting candidate genes associated with OTSCC were identified. The Gene Ontology analysis indicated that several biological processes and cellular functions are consistently altered in OTSCC. Our results demonstrate the feasibility of utilizing biomarkers discovered by global expression profiling analyses for the detection and diagnosis of OTSCC. In addition, we also provided a valuable reference dataset for future identification and validation of biomarkers for detection, diagnosis and prognosis of OTSCC.

**Methods**

**Patients**

Microarray data were generated from 26 microdissected OTSCC tissues and 12 matching normal tissue samples. Three additional microarray datasets from 27 OTSCC cases and 10 matching normal control tissues that published previously were either downloaded from GEO database (GSE2280, [46] and GSE3524, [47]) or requested from the authors [7]. Clinical characterizations of these patients are outlined in Table 4. The tumor stages were determined according to the American Joint Committee on Cancer (AJCC) designated classification. This study is approved by Institutional Review Boards at University of California at Los Angeles and at University of Illinois at Chicago.

**Tumor procurement, RNA extraction and microarray hybridization**

The OTSCC tissues and their matching normal samples were obtained for this study. These tissues were snap frozen. Cancer tissues containing more than 80% tumor cells based on haematoxylin and eosin (H&E) staining and pathological examination were identified and selectively microdissected by a trained pathologist. The total RNA was isolated using RNeasy Mini kit (Qiagen), and quantified by the RiboGreen RNA Quantitation Reagent (Molecular Probes). A total of 150–200 ng of purified total RNA was amplified by a modified T7 RNA amplification protocol as described previously [15,17]. The total RNA was isolated using RNeasy Mini kit (Qiagen), and quantified by the RiboGreen RNA Quantitation Reagent (Molecular Probes). A total of 150–200 ng of purified total RNA was amplified by a modified T7 RNA amplification protocol as described previously [15,17]. The Enzo BioArray High Yield RNA Transcript Labeling System (Enzo) was used for labeling the sample prior to hybridization. The biotinylated cRNA (IVT product) was purified using the RNeasy kit (Qiagen). The quantity and purity of the biotinylated cRNA was determined by spectrophotometry and an aliquot of the sample was checked by gel electrophoresis. The samples were hybridized to the Affymetrix Human Genome U133 Plus 2.0 GeneChip arrays accord-

---

**Table 3: Expression values of genes that constitute the collagen catabolism (GO:0030574) and positive regulation of chemotaxis (GO:0050921) processes in OTSCC.**

| GO term        | entrezID | Gene Symbol | Fold change | FDR level |
|----------------|----------|-------------|-------------|-----------|
| GO:0030574     |          |             |             |           |
| 4312 MMP1      | 57.6     | 0           |             |           |
| 4313 MMP2      | 1.44     | 0.109       |             |           |
| 4314 MMP3      | 8.43     | 1.01E-08    |             |           |
| 4316 MMP7      | 2.83     | 0.0049      |             |           |
| 4317 MMP8      | 1.01     | 0.7963      |             |           |
| 4318 MMP9      | 4.08     | 5.85E-05    |             |           |
| 4319 MMP10     | 8.45     | 9.95E-06    |             |           |
| 4320 MMP11     | 2.02     | 0.00085     |             |           |
| 4322 MMP13     | 3.79     | 0.0007      |             |           |
| 4325 MMP16     | 0.93     | 0.1117      |             |           |
| 4327 MMP19     | 1.08     | 0.4404      |             |           |
| 5184 PEPD      | 0.92     | 0.4602      |             |           |
| 5645 PRSS2     | 0.76     | 0.0187      |             |           |
| 5653 KLK6      | 0.73     | 0.4218      |             |           |
| 5657 PRTN3     | 0.96     | 0.2656      |             |           |
| 9508 ADAMTS3   | 1.05     | 0.5858      |             |           |
| 9509 ADAMTS2   | 1.19     | 0.2451      |             |           |
| 56547 MMP26    | 0.94     | 0.2394      |             |           |
| 9353 SLIT2     | 0.88     | 0.1700      |             |           |
| 566 AZU1       | 0.95     | 0.1643      |             |           |
| 3576 IL8       | 5.87     | 1.54E-06    |             |           |
| 6696 SPP1      | 3.23     | 0.0027      |             |           |
| 7422 VEGF      | 1.20     | 0.4614      |             |           |
| 7857 SCG2      | 1.00     | 0.9901      |             |           |
ing to the Affymetrix protocols. The arrays were scanned with a GeneChip Scanner 3000. The scanned array images were processed with GeneChip Operating software (GCOS), and the CEL files were extracted for further analysis.

Array data analysis and gene ontology analysis

The CEL files from all datasets (newly generated array data from 26 OTSCCs and 12 matching normals, and additional 27 OTSCCs and 10 normals from published studies [7,46,47]) were imported into the statistical software R 2.4.1 [48] using Bioconductor [49]. The meta-analysis was performed as described [50]. In brief, the Robust Multi-Array Average (RMA) expression measures [51] were computed after background correction and quantile normalization for each microarray dataset. Then, expression values of the overlapping probesets between U133A and U133 Plus 2.0 arrays were extracted. Probeset-level quantile normalization was performed across all samples to make the effect sizes similar among the four datasets. To visualize the overall expression patterns, we performed Principal Component Analysis (PCA) after removing the normal group mean vector separately from each of the four datasets. Finally, for every probeset, a mixed effects model was applied to identify differential expression. For gene $i$ in sample $k$ of experiment $j$, 

Figure 4
Elevated expressions of IL8 and MMP9 genes in OTSCC. The qRT-PCR tests were performed with primer sets specific for IL8 gene (A) and MMP9 gene (B) on 25 OTSCC cases and 12 normal matching samples. The $p$ values (Wilcoxon test) were presented. Immunohistochemistry tests were performed using monoclonal antibody to IL8 and MMP9 and detected using peroxidase-antiperoxidase and diaminobenzidine (DAB) on 10 OTSCC cases. Positive stained SCC cells were observed in 4 cases for IL8 and 8 cases for MMP9. Representative images for IL8 (C) and MMP9 (D) were presented.

Table 4: Clinical Characterization of the OTSCC Patients*

|            | OTSCC (n = 53) | Normal (n = 22) |
|------------|---------------|-----------------|
| **Age**    |               |                 |
| Average    | 57            | 56              |
| (Range)    | 32–82         | 37–78           |
| **Gender** |               |                 |
| Male (%)   | 73.6          | 59.1            |
| Female (%) | 26.4          | 40.9            |
| **Anatomic Site** |     |                 |
| Tongue (%) | 100           | 100             |
| **Pathological T Stage**  |     |                 |
| Stage 4 (%) | 56.6          |                 |
| Stage 3 (%) | 7.5           |                 |
| Stage 2 (%) | 22.6          |                 |
| Stage 1 (%) | 13.2          |                 |
| **Pathological N Stage** |   |                 |
| Stage 2 (%) | 43.4          |                 |
| Stage 1 (%) | 7.5           |                 |
| Stage 0 (%) | 49.1          |                 |

* The M stage data is not available.
\[ y_{ijk} = m_i + a_j + \beta_i x_{ijk} + e_{ijk}, \quad a_j \sim N(0, \sigma_a), \quad e_{ijk} \sim N(0, \sigma_e) \]

In the model, the random effect \( a_j \) is the laboratory effect, and \( \beta_i \) is the first-order cancer effect, which is our major focus in the identification of cancer-associated genes. After obtaining the estimates and the p-values of the \( \beta_i \)'s of each probe set, we corrected the p-values for false discovery rate (FDR) [52]. We selected genes at the FDR level of 0.01, and with cancer effect size > 1 (> 2 fold change between cancer and normal samples). Functional analysis of the differentially expressed genes was carried out using the GOstats package in Bioconductor [22] based on the Gene Ontology Consortium database [23].

Quantitative RT-PCR (qRT-PCR)
The mRNA levels of interleukin-8 (IL8) and matrix metalloproteinases 9 (MMP9) in OTSCCs and normal tissues were further validated using qRT-PCR as previously described [15,17]. The RNA was converted to first strand cDNA using MuLV reverse transcriptase (Applied Biosystems) and the quantitative PCR was performed using iQ SYBR Green Supermix (Bio-Rad) in a BIO-RAD iCycler iQ real-time PCR detection system. The primer sets specific for IL8 (Forward: 5'-GAGGGTTGTGAGAAGTTTTTG-3', Reverse: 5'-CTGGCATCTTCACTGATTCTTG-3') and for MMP9 (Forward: 5'-GCACCACGTCTTCCAGTACC-3', Reverse: 5'-TCAACTCAGCCTCAGGAACTC-3') were used. All reactions were performed in triplicate. The melting curve analyses were performed to ensure the specificity of the qRT-PCR reactions. The data analysis was performed using the \( 2^{-\Delta \Delta Ct} \) method described previously [53], where beta-actin was used as reference gene. The qRT-PCR based gene expression values between two groups were compared by the nonparametric Wilcoxon test.

Immunohistochemistry
The expression of IL8 and MMP9 in OTSCCs were further examined using immunohistochemistry tests as previously described [54]. In brief, the OTSCC tissues were processed, embedded, and sectioned at 5 \( \mu \)m. Tissue sections were stained using monoclonal antibody to IL8 (MAB208) (R & D Systems) and MMP9 (ab51203) (Abcam, Inc) and detected using peroxidase-antiperoxidase and diaminobenzadine (DAB) with a Discovery XT automated instrument (Ventana Medical Systems, Inc).

Authors' contributions
BZ, LM, DW, and XZ conceived the idea for the project and drafted the manuscript. HY, TY, ST, JS, JW and XZ performed the laboratory analyses and conducted statistical analyses. JS and JW provided discussions on clinical relevance. BZ, JS, JW, LM, DW and XZ revised the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1  
Supplement Table S1: Up-regulated transcripts in OTSCC. The table showing the complete list of the up-regulated transcripts in OTSCC (p value < 0.01; fold increase > 2.0). 
Click here for file  
[http://www.biomedcentral.com/content-supplementary/1471-2164-9-69-S1.doc]

Additional file 2  
Supplement Table S2: Down-regulated transcripts in OTSCC. The table showing the complete list of the down-regulated transcripts in OTSCC (p value < 0.01; fold increase < 0.5). 
Click here for file  
[http://www.biomedcentral.com/content-supplementary/1471-2164-9-69-S2.doc]

Additional file 3  
Supplement Table S3: Enhanced Biological Processes (BP), Molecular Functions (MF) and Cellular Components (CC) in OTSCC. The table showing the complete list of the enhanced biological processes (BP), molecular functions (MF) and cellular components (CC) in OTSCC (p value < 0.01). 
Click here for file  
[http://www.biomedcentral.com/content-supplementary/1471-2164-9-69-S3.doc]

Additional file 4  
Supplement Table S4: Suppressed Biological Processes (BP), Molecular Functions (MF) and Cellular Components (CC) in OTSCC. The table showing the complete list of the suppressed biological processes (BP), molecular functions (MF) and cellular components (CC) in OTSCC (p value < 0.01). 
Click here for file  
[http://www.biomedcentral.com/content-supplementary/1471-2164-9-69-S4.doc]

Acknowledgements
This work was supported in part by NIH PHS grants K22 DE014847, RO3 DE016569, and RO3 CA114688 (to X. Zhou), R01 DE05970 (to D. Wong), P01 CA106451 Project 3 (to L. Mao), and a Foundation de France grant (to S. Temam). We thank Drs. O’Donnell and Muschel at Children’s Hospital of Philadelphia and Dr Dermody at University of Medicine and Dentistry of New Jersey for sharing their microarray datasets with us. We thank Mr. Hui Zhou and Ms. Nisa Pungpravat for their excellent technical assistance, and Ms. Katherine Long for editorial assistance.
References
1. Dobrossy L: Epidemiology of head and neck cancer: magnitude of the problem. Cancer Metastasis Rev 2005, 24(1):9-17.
2. Timar J, Csoka O, Remenar E, Repassy G, Kasler M: Progression of head and neck squamous cell cancer. Cancer Metastasis Rev 2005, 24(1):107-127.
3. Mackenzie J, Ah-See K, Thakker N, Sloan P, Maran AG, Birch J, Macfarlane GJ: Increasing incidence of oral cancer amongst young persons: what is the aetiology? Oral Oncol 2000, 36(4):387-389.
4. Amettz A, Anderson H, Bjorklund N, Holler T, Kantola S, Mark J, Olsen JH, Wennerberg J: Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. Int J Cancer 2002, 101(1):95-99.
5. Schantz SP, Yu GP: Head and neck cancer incidence trends in young persons: what is the aetiology? Oral Oncol 2002, 38(1):123-129.

4. Timear J, Csuka O, Remenar E, Repassy G, Kasler M: Progression of head and neck squamous cell cancer. Cancer Metastasis Rev 2005, 24(1):107-127.
5. Mackenzie J, Ah-See K, Thakker N, Sloan P, Maran AG, Birch J, Macfarlane GJ: Increasing incidence of oral cancer amongst young persons: what is the aetiology? Oral Oncol 2000, 36(4):387-389.
6. Amettz A, Anderson H, Bjorklund N, Holler T, Kantola S, Mark J, Olsen JH, Wennerberg J: Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. Int J Cancer 2002, 101(1):95-99.
7. Schantz SP, Yu GP: Head and neck cancer incidence trends in young persons: what is the aetiology? Oral Oncol 2002, 38(1):123-129.

4. Timear J, Csuka O, Remenar E, Repassy G, Kasler M: Progression of head and neck squamous cell cancer. Cancer Metastasis Rev 2005, 24(1):107-127.
5. Mackenzie J, Ah-See K, Thakker N, Sloan P, Maran AG, Birch J, Macfarlane GJ: Increasing incidence of oral cancer amongst young persons: what is the aetiology? Oral Oncol 2000, 36(4):387-389.
6. Amettz A, Anderson H, Bjorklund N, Holler T, Kantola S, Mark J, Olsen JH, Wennerberg J: Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. Int J Cancer 2002, 101(1):95-99.
7. Schantz SP, Yu GP: Head and neck cancer incidence trends in young persons: what is the aetiology? Oral Oncol 2002, 38(1):123-129.

4. Timear J, Csuka O, Remenar E, Repassy G, Kasler M: Progression of head and neck squamous cell cancer. Cancer Metastasis Rev 2005, 24(1):107-127.
5. Mackenzie J, Ah-See K, Thakker N, Sloan P, Maran AG, Birch J, Macfarlane GJ: Increasing incidence of oral cancer amongst young persons: what is the aetiology? Oral Oncol 2000, 36(4):387-389.
6. Amettz A, Anderson H, Bjorklund N, Holler T, Kantola S, Mark J, Olsen JH, Wennerberg J: Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. Int J Cancer 2002, 101(1):95-99.
7. Schantz SP, Yu GP: Head and neck cancer incidence trends in young persons: what is the aetiology? Oral Oncol 2002, 38(1):123-129.

4. Timear J, Csuka O, Remenar E, Repassy G, Kasler M: Progression of head and neck squamous cell cancer. Cancer Metastasis Rev 2005, 24(1):107-127.
5. Mackenzie J, Ah-See K, Thakker N, Sloan P, Maran AG, Birch J, Macfarlane GJ: Increasing incidence of oral cancer amongst young persons: what is the aetiology? Oral Oncol 2000, 36(4):387-389.
6. Amettz A, Anderson H, Bjorklund N, Holler T, Kantola S, Mark J, Olsen JH, Wennerberg J: Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. Int J Cancer 2002, 101(1):95-99.
7. Schantz SP, Yu GP: Head and neck cancer incidence trends in young persons: what is the aetiology? Oral Oncol 2002, 38(1):123-129.

4. Timear J, Csuka O, Remenar E, Repassy G, Kasler M: Progression of head and neck squamous cell cancer. Cancer Metastasis Rev 2005, 24(1):107-127.
5. Mackenzie J, Ah-See K, Thakker N, Sloan P, Maran AG, Birch J, Macfarlane GJ: Increasing incidence of oral cancer amongst young persons: what is the aetiology? Oral Oncol 2000, 36(4):387-389.
6. Amettz A, Anderson H, Bjorklund N, Holler T, Kantola S, Mark J, Olsen JH, Wennerberg J: Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. Int J Cancer 2002, 101(1):95-99.
7. Schantz SP, Yu GP: Head and neck cancer incidence trends in young persons: what is the aetiology? Oral Oncol 2002, 38(1):123-129.

4. Timear J, Csuka O, Remenar E, Repassy G, Kasler M: Progression of head and neck squamous cell cancer. Cancer Metastasis Rev 2005, 24(1):107-127.
5. Mackenzie J, Ah-See K, Thakker N, Sloan P, Maran AG, Birch J, Macfarlane GJ: Increasing incidence of oral cancer amongst young persons: what is the aetiology? Oral Oncol 2000, 36(4):387-389.
6. Amettz A, Anderson H, Bjorklund N, Holler T, Kantola S, Mark J, Olsen JH, Wennerberg J: Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. Int J Cancer 2002, 101(1):95-99.
7. Schantz SP, Yu GP: Head and neck cancer incidence trends in young persons: what is the aetiology? Oral Oncol 2002, 38(1):123-129.

4. Timear J, Csuka O, Remenar E, Repassy G, Kasler M: Progression of head and neck squamous cell cancer. Cancer Metastasis Rev 2005, 24(1):107-127.
5. Mackenzie J, Ah-See K, Thakker N, Sloan P, Maran AG, Birch J, Macfarlane GJ: Increasing incidence of oral cancer amongst young persons: what is the aetiology? Oral Oncol 2000, 36(4):387-389.
6. Amettz A, Anderson H, Bjorklund N, Holler T, Kantola S, Mark J, Olsen JH, Wennerberg J: Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. Int J Cancer 2002, 101(1):95-99.
7. Schantz SP, Yu GP: Head and neck cancer incidence trends in young persons: what is the aetiology? Oral Oncol 2002, 38(1):123-129.

4. Timear J, Csuka O, Remenar E, Repassy G, Kasler M: Progression of head and neck squamous cell cancer. Cancer Metastasis Rev 2005, 24(1):107-127.
5. Mackenzie J, Ah-See K, Thakker N, Sloan P, Maran AG, Birch J, Macfarlane GJ: Increasing incidence of oral cancer amongst young persons: what is the aetiology? Oral Oncol 2000, 36(4):387-389.
6. Amettz A, Anderson H, Bjorklund N, Holler T, Kantola S, Mark J, Olsen JH, Wennerberg J: Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. Int J Cancer 2002, 101(1):95-99.
7. Schantz SP, Yu GP: Head and neck cancer incidence trends in young persons: what is the aetiology? Oral Oncol 2002, 38(1):123-129.
42. Soini Y, Vakkala M, Kahlos K, Paakkko P, Kinnula V. MnSOD expression is less frequent in tumour cells of invasive breast carcinomas than in in situ carcinomas or non-neoplastic breast epithelial cells. J Pathol 2001, 195(2):156-162.

43. Baker AM, Oberley LW, Cohen MB. Expression of antioxidant enzymes in human prostatic adenocarcinoma. Prostate 1997, 32(4):229-233.

44. Bostwick DG, Alexander EE, Singh R, Shan A, Qian J, Santella RM, Oberley LW, Yan T, Zhong W, Jiang X, Oberley TD. Antioxidant enzyme expression and reactive oxygen species damage in prostatic intraepithelial neoplasia and cancer. Cancer 2000, 89(1):123-134.

45. Lo WY, Tsai MH, Tsai Y, Hua CH, Tsai FJ, Huang SY, Tsai CH, Lai CC. Identification of over-expressed proteins in oral squamous cell carcinoma (OSCC) patients by clinical proteomic analysis. Clin Chem Acta 2007, 376(1-2):101-107.

46. O'Donnell RK, Kupferman M, Wei SJ, Singhal S, Weber R, O'Malley B, Cheng Y, Putt M, Feldman M, Ziober B, Muschel RJ. Gene expression signature predicts lymphatic metastasis in squamous cell carcinoma of the oral cavity. Oncogene 2005, 24(7):1244-1251.

47. Toruner GA, Ulger C, Alkan M, Galante AT, Rinaggio J, Wilk R, Tian B, Sotopoulos P, Hameed MR, Schwab MN, Dernody JJ. Association between gene expression profile and tumor invasion in oral squamous cell carcinoma. Cancer Genet Cytogenet 2004, 154(1):27-35.

48. R_Development_Core_Team. R: A language and environment for statistical computing. In R Foundation for Statistical Computing Vienna, Austria : 2006.

49. Gentleman RC, Carey VJ, Bates DM, Bolstad B,Destlring M, Dubois S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang Y, Zhang J. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004, 5(10):R80.

50. Yu T, Ye H, Chen Z, Ziober BL, Zhou X. Dimension reduction and mixed-effects model for microarray meta-analysis of cancer. Front Biosci 2008, 13:2714-2720.

51. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003, 31(4):e15.

52. Benjamin Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 1995, 57:289-300.

53. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta CT(T)) Method. Methods 2001, 25(4):402-408.

54. Peng H, Shintani S, Kim Y, Wong DT. Loss of p12CDK2-AP1 expression in human oral squamous cell carcinoma with disrupted transforming growth factor-beta-Smad signaling pathway. Neoplasia 2006, 8(12):1028-1036.