

## Tumor and Salivary Matrix Metalloproteinase Levels Are Strong Diagnostic Markers of Oral Squamous Cell Carcinoma

Marni Stott-Miller<sup>1,2</sup>, John R. Houck<sup>1</sup>, Pawadee Lohavanichbutr<sup>1</sup>, Eduardo Méndez<sup>1,3,5</sup>, Melissa P. Upton<sup>4</sup>, Neal D. Futran<sup>3</sup>, Stephen M. Schwartz<sup>1,2</sup>, and Chu Chen<sup>1,2,3</sup>

### Abstract

**Background:** The matrix metalloproteinases (MMP) cause degradation of the extracellular matrix and basement membranes, and thus may play a key role in cancer development.

**Methods:** In our search for biomarkers for oral squamous cell carcinomas (OSCC), we compared primary OSCC, oral dysplasia and control subjects with respect to: (i) expression of *MMP1*, *MMP3*, *MMP10*, and *MMP12* in oral epithelial tissue using Affymetrix U133 2.0 Plus GeneChip arrays, followed by quantitative reverse transcription-PCR (qRT-PCR) for *MMP1*, and (ii) determination of MMP1 and MMP3 concentrations in saliva.

**Results:** *MMP1* expression in primary OSCC ( $n = 119$ ) was >200-fold higher ( $P = 7.16 \times 10^{-40}$ ) compared with expression levels in nonneoplastic oral epithelium from controls ( $n = 35$ ). qRT-PCR results on 30 cases and 22 controls confirmed this substantial differential expression. The exceptional discriminatory power to separate OSCC from controls was validated in two independent testing sets (AUC% = 100; 95% CI: 100–100 and AUC% = 98.4; 95% CI: 95.6–100). Salivary concentrations of MMP1 and MMP3 in OSCC patients (33 stage I/II, 26 stage III/IV) were 6.2 times (95% CI: 3.32–11.73) and 14.8 times (95% CI: 6.75–32.56) higher, respectively, than in controls, and displayed an increasing trend with higher stage disease.

**Conclusion:** Tumor and salivary MMPs are robust diagnostic biomarkers of OSCC.

**Impact:** The capacity of MMP gene expression to identify OSCC provides support for further investigation into MMPs as potential markers for OSCC development. Detection of MMP proteins in saliva in particular may provide a promising means to detect and monitor OSCC noninvasively. *Cancer Epidemiol Biomarkers Prev*; 20(12); 2628–36. ©2011 AACR.

### Introduction

Oral squamous cell carcinoma (OSCC) of the oral cavity and oropharynx is one of the most common cancers in the world, with an estimated 400,000 new cases and 200,000 deaths in 2008 worldwide (1). In 2010, an estimated 36,540 new cases and 7,880 deaths from OSCC occurred in the United States (2). Despite advances in chemotherapy, radiation treatments, screening tools such as the VELscope (LED Dental, Inc.; ref. 3), and improvements in surgical techniques, survival remains extremely poor. The

5-year survival is estimated at 63% for Whites and 43% for Blacks in the United States (2). Prognosis is heavily based on the AJCC staging system, which has limited utility for predicting survival because patients with tumors of the same AJCC stage often have heterogeneous responses to treatment (4).

To improve the mortality and morbidity burden associated with OSCC, there is an urgent need to identify sensitive and specific biomarkers for early detection and prognosis. For utmost clinical utility, these biomarkers should be measurable in specimens that are collected with minimal discomfort to the patient (5).

The matrix metalloproteinases (MMP) have frequently been studied as potential cancer biomarkers, including in OSCC, and have been associated with tumor invasion and metastases (6–9). Insertion/deletion (ins/del) polymorphisms in the promoter region of *MMP1* (–1,607, 1G/2G, rs1799750; ref. 10) and *MMP3* (–1,171, 5A/6A, rs3025039; ref. 11) influence the level of MMP expression. However, epidemiologic studies investigating these polymorphisms and oral cancer risk have not been entirely consistent (12–14). Overexpression of MMP transcripts in OSCC, have been observed repeatedly (7, 15–18). The protein concentrations of

**Authors' Affiliations:** <sup>1</sup>Program in Epidemiology, Fred Hutchinson Cancer Research Center; Departments of <sup>2</sup>Epidemiology, <sup>3</sup>Otolaryngology-Head and Neck Surgery, and <sup>4</sup>Pathology, University of Washington; and <sup>5</sup>Surgery and Perioperative Care Service, VA Puget Sound Health Care System, Seattle, Washington

**Note:** Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

**Corresponding Author:** Chu Chen, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, M5-C800 P.O. Box 19024, Seattle, WA 98109. Phone: 206-667-6644; Fax: 206-667-2537; E-mail: cchen@fhcrc.org

doi: 10.1158/1055-9965.EPI-11-0503

©2011 American Association for Cancer Research.

MMP1, MMP2, MMP3, and MMP9 have been observed to be higher in OSCC tumor tissue compared with control tissue (19). However, many of these studies are limited by having very small numbers of patients (15, 16) or by use of control tissue from OSCC patients instead of OSCC-free controls (7, 17–19). *MMP1* transcript levels in saliva have been shown to distinguish OSCC cases from controls (AUC = 98.9%; ref. 20). In addition, overexpression of *MMP1* and *MMP9* have been associated with progression of dysplasia to cancer (21).

This study sought to focus on the MMPs as potential diagnostic and prognostic biomarkers. Our primary goal was to determine whether salivary concentrations of the most highly differentially expressed MMPs might be useful as a diagnostic aid. Although others have reported elevated MMP transcript levels in saliva (20), protein biomarkers in saliva should provide more accurate representations of molecular function. There are numerous posttranscriptional processes, including transcript de/stabilization, translation, protein modification, and degradation. As such, mRNA levels may not correlate with protein activity (22, 23). In addition, saliva is an ideal diagnostic tool for biomarker assessment due to its accessibility and ease of collection (5, 24). An additional goal was to evaluate the use of the MMPs as a prognostic aid by investigating whether MMP expression was associated with an increased risk of death.

## Methods

### Study population

As described by Chen and colleagues (25), cases were previously untreated primary OSCC patients scheduled for biopsy or surgery at the University of Washington Medical Center, Harborview Medical Center, or Veterans Affairs (VA) Puget Sound Health Care System. Patients diagnosed with oral dysplastic lesions were also enrolled through these medical centers during the same time period. Controls were patients who received oral surgery, such as uvulopalatopharyngoplasty and tonsillectomy, for a nonmalignant or nonpre-malignant condition, at the same institutions and during the same time period. Subjects were enrolled between December 2003 and April 2007.

Study participants were interviewed using a structured questionnaire eliciting demographic characteristics, medical and lifestyle history, including tobacco and alcohol use. Data on tumor stage and other characteristics were abstracted from medical records. Participants gave informed consent and study procedures were approved by the Institutional Review Boards of the Fred Hutchinson Cancer Research Center, University of Washington, and VA Puget Sound Health Care System.

### Tissue and saliva collection

Tumor tissue from cases was obtained at time of resection prior to other treatment. Tissue from dysplasia

patients and controls was obtained at time of biopsy or surgery. Oral epithelial tissue from controls was collected from the uvula or anterior tonsillar pillar, avoiding contamination with surrounding lymphoid tissue. The removed tissue was immersed in RNALater (Applied Biosystems, Inc.) for at least 12 hours at 4°C before storage at –80°C until use.

Saliva samples were collected preoperatively during the clinic visit. Patients were asked not to eat or drink for at least 1 hour before collection and to spit into a centrifuge tube. The saliva was stored at 4°C for up to 2 hours, then centrifuged for 10 minutes at 1,300 × g and aliquoted into cryovials for storage at –80°C.

### Patient follow-up

Patient follow-up was done from July, 2004 to January, 2011. Subjects were followed through periodic telephone contacts at as close to 6-month intervals as possible following surgery, review of medical records and linkage to data sources. Vital status was checked against the Social Security Death Index (SSDI) and Fred Hutchinson Cancer Research Center's Cancer Surveillance System (CSS), which is part of the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute, and is updated with the Washington State Death Certificate database and National Death Index. A death was classified as due to OSCC or not due to OSCC based on review of medical records and death certificates by head and neck surgeons involved in the study. If there was no indication of death, we censored that subject at the last known date of follow-up. Only 2 subjects were lost to follow-up.

### Laboratory studies

As described previously (25), total RNA was extracted from tumor and nonneoplastic oral epithelium using a TRIzol method (Invitrogen), purified with an RNeasy mini kit (Qiagen), processed using a GeneChip Expression 3'-Amplification Reagents Kit (Affymetrix), and examined with Affymetrix GeneChip Human Genome U133 Plus 2.0 Arrays.

Tumor and nonneoplastic oral epithelium samples were screened for the presence of HPV DNA using a nested PCR protocol (26). All samples that showed a positive result were tested for HPV types using the LINEAR ARRAY HPV Genotyping Test (Roche), under a research-use only agreement. "High-risk" HPV was defined by 13 types, including HPV-16 and HPV-18; "low-risk" HPV corresponded to 24 types. To verify HPV type calls, a subset of the samples were amplified and sequenced using HPV-16-specific primers, and compared against a known HPV-16 sequence (GenBank 333031).

We verified *MMP1* gene expression results by quantitative reverse transcription-PCR (qRT-PCR) on a subset of 30 OSCC cases and 22 controls using a QuantiTect SYBR Green RT-PCR kit (Qiagen) and bioinformatically validated QuantiTect primers (Qiagen) on an ABI 7900HT Sequence Detection System (Applied Biosystems, Inc.).

Standard curves were generated using Universal Human Reference RNA (Stratagene) for all genes, with the linear correlation coefficient ( $r^2$ )  $\geq 0.99$  for all runs. The mean threshold cycle ( $C_t$ ) values were calculated from triplicate  $C_t$  values. Samples that had  $C_t$  values with a SD  $> 0.3$  in the triplicate run were repeated. Mean  $C_t$  values were standardized to the mean  $C_t$  value of the reference gene  $\beta$ -actin. The cycling conditions consisted of 30 minutes at 50°C, 15 minutes at 95°C, and 40 cycles each of 15 seconds at 94°C, 30 seconds at 55°C, and 30 seconds at 72°C.

Saliva samples from 100 subjects (60 primary OSCC, 15 dysplasia, and 25 controls) were tested for MMP1 and MMP3 concentration. These MMPs were selected because they exhibited the largest fold-difference out of MMPs analyzed in gene expression analyses. Each sample in duplicate was analyzed blinded to patient status by Aushon Biosciences Inc. using a SearchLight multiplex sandwich-ELISA proteome array (Aushon BioSystems), with images analyzed using the SearchLight Array Analyst software. Protein concentrations from 2 subjects could not be measured because of excessive viscosity.

### MMP expression analyses for primary OSCC compared with control epithelium

Two rounds of quality control checks were done on the microarray gene expression data. First, recommendations by Affymetrix were followed to determine if any GeneChips needed to be excluded (27). Second, the "affyQCReport" and "affyPLM" software in the Bioconductor package within R statistical programming language were used to identify any poor-quality chips (28). In total, 167 cases and 45 controls passed both quality control procedures. The data were divided into a training set (119 cases and 35 controls) and a testing set (48 cases and 10 controls). We used the training set in analyses that identified the top differentially expressed genes, validated results in an independent testing set to evaluate discriminatory capacity of identified genes, and then merged the training and testing sets to provide greater numbers of subjects for secondary analyses on stage, survival, and stratified by site. We extracted gene expression values for ~54,000 probe sets in the training set from CEL files and normalized the data using the RMA algorithm in Partek Genomics Suite (29).

Although overexpression of the MMPs in OSCC had previously been reported by others (6–9, 15–19), our prior gene expression analyses did not identify the MMPs among the list of top differentially expressed genes because we had excluded probe sets based on magnitude (probe sets were excluded if the expression value for that probe set in any of the samples was  $< 3$  on  $\log_2$  scale; ref. 25). This study did not eliminate such probe sets. Using the training set, we eliminated probe sets if expression showed little variation (defined as interquartile range of expression levels less than 0.3 on the  $\log_2$  scale), resulting in ~35,000 probe sets for further analyses.

To compare expression levels of tumor tissue to tissue from control subjects in the training set, we conducted

ANCOVA (adjusting for age and sex) and a Student  $t$  test using Partek Genomics Suite. We set the false discovery rate at 1%. We then created a list for further analyses that contained only those probe sets with (i)  $t$  score greater than 6, and (ii) a 1.5-fold or greater difference in gene expression between cases and controls (on the  $\log_2$  scale). We identified the top MMPs by sorting this list according to case–control fold-difference.

We compared expression levels of these top differentially expressed MMPs in primary OSCC ( $n = 119$ ) with expression levels in oral epithelium from control subjects ( $n = 35$ ) by fitting linear regression models, adjusted for age and sex, and estimating 95% CI for all association estimates. We repeated these analyses with additional adjustment for pack years of smoking (continuous) and alcoholic drinks per day in the year prior to the date of diagnosis for OSCC cases or recruitment for controls (categorical). We calculated Pearson correlation coefficients for the top 4 differentially expressed MMPs.

In secondary analyses using data from both the training and testing sets, we conducted linear regression analyses of MMP1 case–control differences, adjusting for age and sex, after separating by (i) site: oral cavity (115 cases, 45 controls) versus oropharynx (52 cases, 45 controls), and (ii) HPV status: high-risk HPV positive (56 cases, 4 controls) versus HPV negative or low-risk HPV (111 cases, 40 controls).

### Validation of MMP1 results using independent testing sets

We extracted gene expression values from files of our independent testing set (48 OSCC cases and 10 controls) and a publicly available data set from Gene Expression Omnibus (GEO; GSE6791; 42 OSCC cases and 14 controls; ref. 30), and normalized the 2 data sets using the RMA algorithm. We then used the receiver operator characteristic (ROC) curve to calculate, for each data set, the area under the ROC curve (AUC) to evaluate the discriminatory performance of the statistical prediction models.

### Statistical Analyses

#### Salivary MMP1 and MMP3 concentration

We log-transformed the MMP1 and MMP3 concentrations to correct for skewed data. We calculated the ratio of geometric means of their concentrations in cases ( $n = 59$ ) and control subjects ( $n = 25$ ) by fitting adjusted regression models, adjusted for age and sex. We then repeated these analyses for MMP1 after stratifying by site: oral cavity (45 cases, 25 controls) versus oropharynx (15 cases, 25 controls), and HPV status: high-risk HPV positive (15 cases, 1 control) versus high-risk HPV negative (43 cases, 24 controls).

In secondary analyses, we calculated the geometric mean for control, dysplasia, stage I/II OSCC, and stage III/IV OSCC. To evaluate sensitivity and specificity, we conducted ROC analyses and calculated AUC for salivary MMP1 and MMP3. To determine whether tissue mRNA

levels of *MMP1* and *MMP3* correlated with corresponding salivary concentrations, we calculated the Pearson correlation coefficient.

In all analyses for which software is not specified, we used STATA statistical software (version 10.0, Stata Corp.)

### Associations of *MMP1* and *MMP3* gene expression with stage and survival

We merged the training and testing datasets to compare expression levels of *MMP1* and *MMP3* (the top 2 differentially expressed genes based on the initial microarray analyses) across different disease categories. We were unable to categorize 2 subjects because of missing data. After categorization, we calculated mean  $\log_2$  *MMP1* and *MMP3* expression levels for 45 controls, 17 dysplasia, 54 stage I/II primary OSCC, and 111 stage III/IV primary OSCC.

We conducted survival analyses for primary OSCC cases using data from our training and testing sets. The primary outcome variable was survival time, measured as days after diagnosis until a death occurred, or until subjects were lost to follow-up, or January 25, 2011. We started the accumulation of follow-up time and deaths at 4 months after diagnosis date to exclude any deaths due to treatment-related complications. There were 159 subjects with at least 4 months of follow-up time. To analyze risk of death based on gene expression values of *MMP1* and *MMP3*, we calculated OSCC-specific and all cause HR by Cox proportional hazards regression, adjusting for age, sex, and stage and HPV status. We then conducted mutual adjustment for *MMP1* and *MMP3*.

## Results

Cases tended to be older than control patients, and were more likely to be male, White and current smokers (Table 1). In the training set, 74% of cases had oral cavity tumors and 26% had oropharyngeal tumors. In our internal independent testing set, oral cavity tumors accounted for 60% and oropharyngeal tumors for 40% of the cases.

### Gene expression levels of MMPs

Results of ANCOVA analyses on the training set resulted in 16,228 significant probe sets out of a total of 33,057. After preprocessing and filtering, we identified 173 probe sets that were substantially differentially expressed between cases and controls (Supplementary Table S1). After sorting by fold-difference (cases vs. controls), the top overexpressed genes in this list were *MMP1*, *MMP3*, *MMP10*, and *MMP12*. *MMP1* had an exceptionally high fold-difference (>200-fold; 3.16-fold on  $\log_2$  scale) and very low adjusted ANCOVA *P* value ( $P = 7.16 \times 10^{-40}$ ). We observed significant differential expression for multiple additional MMPs, including *MMP2* (3.5-fold-difference;  $P = 2.01 \times 10^{-20}$ ) and *MMP9* (8.0-fold-difference;  $P = 2.29 \times 10^{-13}$ ), but elected to conduct further analyses only on the 4 MMPs that were at the top of our list.

The difference in mean  $\log_2$  expression levels between primary OSCC and epithelium from cancer-free controls in the training set, adjusting for age and sex, was 7.74 (95% CI: 7.05–8.44; Table 2) for *MMP1*, 5.30 (95% CI: 4.57–6.03) for *MMP3*, 5.19 (95% CI: 4.52–5.84) for *MMP10*, and 5.26 (95% CI: 4.47–6.06) for *MMP12*. Results were similar for analyses that additionally adjusted for pack years of smoking and alcoholic drinks per day (results not shown). The Pearson correlation coefficients between the 4 MMPs ranged from 0.72 to 0.92 ( $P < 0.001$ , Table 3).

Using the combined training and testing sets, after stratifying by site and adjusting for age and sex, the difference in means for *MMP1* (compared with control oral epithelium) was somewhat higher for oral cavity cancer (difference in means, 8.32; 95% CI: 7.75–8.90), than for oropharyngeal cancer (6.04; 95% CI: 5.09–6.99). Similarly, the difference in means was higher for OSCC among high-risk HPV-negative subjects (difference in means, 8.30; 95% CI: 7.70–8.90), than among high-risk HPV-positive subjects (6.00; 95% CI: 3.99–8.02). For oropharyngeal cancers only, the difference in means was 6.38 (95% CI: 4.54–8.22) among high-risk HPV-negative subjects and 6.10 (95% CI: 3.90–8.31) among high-risk HPV positive subjects. Similar patterns were observed for the other MMPs.

Results of ROC analyses on our testing set (48 cases, 10 controls) indicated that *MMP1* had exceptional discriminatory power to separate OSCC cases from controls (AUC = 100%; 95% CI: 100–100). The predictive accuracy of *MMP1* was confirmed in the GEO dataset, GSE6791 (AUC = 98.4%; 95% CI: 95.6–100).

In the combined Affymetrix array data from the training set and our testing set, the mean of the  $\log_2$  *MMP1* expression levels for 45 control oral epithelium was 3.49 (95% CI: 3.21–3.78), 17 dysplasia samples was 6.69 (95% CI: 5.21–8.17), 54 stage I/II primary OSCC was 11.43 (95% CI: 10.94–11.93) and 111 stage III/IV OSCC was 10.92 (95% CI: 10.46–11.38, Fig. 1). The pattern for *MMP3* expression was similar (Fig. 1).

Our qRT-PCR results confirmed the substantial differential expression of *MMP1*, with a difference in mean  $C_t$  value for controls versus cases of 10.43 (95% CI: 9.31–11.55,  $P < 0.001$ ). The Pearson correlation coefficient for the microarray gene expression values and the qRT-PCR results was  $-0.99$  ( $P < 0.001$ ).

### Salivary *MMP1* and *MMP3* concentration

After log-transformation, salivary concentrations of *MMP1* for primary OSCC patients ( $n = 59$ ) were 6.2 times higher than in control patients ( $n = 25$ ), adjusting for age and sex (95% CI: 3.32–11.73;  $P < 0.001$ ). For *MMP3*, adjusted concentrations were 14.8 times higher in cases than in control patients (95% CI: 6.75–32.56;  $P < 0.001$ ). Gene expression values in tissue correlated well with salivary concentrations for *MMP1* (Pearson correlation coefficient, 0.63;  $P < 0.001$ ) and *MMP3* (0.67;  $P < 0.001$ ).

In stratified analyses of *MMP1* salivary concentrations, results were stronger for patients with oral cavity



**Table 1.** Selected characteristics of OSCC patients and controls, University of Washington Affiliated Hospitals, 2003 to 2007

|  | Training set                  |                            | Testing set                  |                            |
|--|-------------------------------|----------------------------|------------------------------|----------------------------|
|  | OSCC case<br>(n = 119)<br>n % | Control<br>(n = 35)<br>n % | OSCC case<br>(n = 48)<br>n % | Control<br>(n = 10)<br>n % |
| Age (y)  |                               |                            |                              |                            |
| 19–39  | 5 (4.2)                       | 14 (40.0)                  | 2 (4.2)                      | 3 (30.0)                   |
| 40–49  | 18 (15.1)                     | 8 (22.9)                   | 8 (16.7)                     | 6 (60.0)                   |
| 50–59  | 40 (33.6)                     | 4 (11.4)                   | 17 (35.4)                    | 1 (10.0)                   |
| 60–88  | 56 (47.1)                     | 9 (25.7)                   | 21 (43.8)                    | 0 (0.0)                    |
| Sex  |                               |                            |                              |                            |
| Male   | 84 (70.6)                     | 25 (71.4)                  | 36 (75.0)                    | 7 (70.0)                   |
| Female   | 35 (29.4)                     | 10 (28.6)                  | 12 (25.0)                    | 3 (30.0)                   |
| Race   |                               |                            |                              |                            |
| White  | 106 (93.0)                    | 24 (68.6)                  | 40 (85.1)                    | 6 (66.7)                   |
| Non-White  | 8 (7.0)                       | 11 (31.4)                  | 7 (14.9)                     | 3 (33.3)                   |
| Unknown  | 5 (0)                         | 1 (1.0)                    | 0 (0)                        | 0 (0)                      |
| Cigarette smoking status <sup>a</sup>            |                               |                            |                              |                            |
| Never/former                                     | 59 (49.6)                     | 25 (71.4)                  | 27 (56.3)                    | 8 (80.0)                   |
| Current  | 60 (50.4)                     | 10 (28.6)                  | 21 (43.7)                    | 2 (20.0)                   |
| Alcoholic drinking status <sup>a</sup>           |                               |                            |                              |                            |
| Never/former                                     | 36 (30.8)                     | 9 (25.7)                   | 19 (40.4)                    | 3 (30.0)                   |
| Current  | 81 (69.2)                     | 26 (74.3)                  | 28 (59.6)                    | 7 (70.0)                   |
| Unknown  | 2 (0)                         | 1 (0)                      | 0 (0)                        | 0 (0)                      |
| AJCC stage                                       |                               |                            |                              |                            |
| I/II   | 42 (35.3)                     |                            | 13 (27.1)                    |                            |
| III/IV   | 77 (64.7)                     |                            | 35 (72.9)                    |                            |
| Tumor site                                       |                               |                            |                              |                            |
| Oral cavity                                      | 87 (73.1)                     |                            | 28 (58.3)                    |                            |
| Oropharynx                                       | 32 (26.9)                     |                            | 20 (41.7)                    |                            |
| High risk HPV status, by tumor site <sup>b</sup> |                               |                            |                              |                            |
| Oral cavity                                      |                               |                            |                              |                            |
| HPV+   | 14 (16.1)                     |                            | 5 (17.9)                     |                            |
| HPV–   | 73 (83.9)                     |                            | 23 (82.1)                    |                            |
| Oropharynx                                       |                               |                            |                              |                            |
| HPV+   | 25 (78.1)                     |                            | 12 (60.0)                    |                            |
| HPV–   | 7 (21.9)                      |                            | 8 (40.0)                     |                            |

<sup>a</sup>As of the date of diagnosis (OSCC cases) or recruitment (controls).

<sup>b</sup>High risk samples contained HPV-16, HPV-35, or HPV-45. Percentages estimated within site.

cancers compared with controls (ratio of geometric means, 8.86; 95% CI: 1.54–16.82) versus patients with oropharyngeal cancers compared with controls (2.51; 95% CI: 0.90–6.97). Similarly, results were stronger among high-risk HPV-negative subjects (ratio of geometric means, 6.79; 95% CI: 3.40–13.56) versus high-risk HPV-positive subjects (5.86; 95% CI: 2.07–16.55). The pattern for MMP3 was similar.

In analyses across disease categories (controls, dysplasia, stage I/II OSCC, and stage III/IV OSCC), the geometric means of salivary MMP1 and MMP3 concentrations displayed an increasing trend with higher stage of disease

(Fig. 2). However, differences between categories did not reach statistical significance.

In ROC analyses, the AUC was 84.5% (95% CI: 76.00–92.95%) for MMP1 and 87.66% (95% CI: 80.16–95.17%) for MMP3. A model that included both MMP1 and MMP3 did not improve upon the AUC obtained for either MMP1 or MMP3 alone.

#### Associations of MMP1 and MMP3 transcript levels with survival

The median follow-up time for primary OSCC cases (n = 159) was 51 months. For each unit increase in log<sub>2</sub>

**Table 2.** Difference in mean MMP expression levels between OSCC cases and controls, University of Washington Affiliated Hospitals, 2003 to 2007

| Gene/protein       | mRNA <sup>b</sup> (119 cases, 35 controls)        |           | Protein (pg/mL; <sup>c</sup> 59 cases, 25 controls) |            |
|--------------------|---|-----------|---|------------|
|                    | Difference in log <sub>2</sub> means <sup>a</sup> | 95% CI    | Ratio of geometric means <sup>a</sup>               | 95% CI     |
| MMP1               | 7.74  | 7.05–8.44 | 6.24  | 3.32–11.73 |
| MMP3               | 5.30  | 4.57–6.03 | 14.83   | 6.75–32.56 |
| MMP10 <sup>d</sup> | 5.19  | 4.52–5.84 | — <sup>d</sup>                                      | —          |
| MMP12 <sup>d</sup> | 5.26  | 4.47–6.06 | — <sup>d</sup>                                      | —          |

<sup>a</sup>Adjusted for age and sex.<sup>b</sup>Measured in OSCC tissue of cases and oral epithelium of controls.<sup>c</sup>Measured in saliva of OSCC cases and controls.<sup>d</sup>Protein concentrations not measured in saliva.

expression value, the HR for *MMP1*, adjusted for age, sex, stage, and HPV status was 1.24 (95% CI: 1.02–1.52) for OSCC-specific death and 1.20 (95% CI: 1.03–1.40) for death from all causes. The adjusted HR was 1.09 (0.86–1.38) for death from known non-OSCC causes. The adjusted HR for *MMP3* expression for each unit increase in log<sub>2</sub> value was 1.26 (95% CI: 1.09–1.46) for OSCC-specific death and 1.25 (95% CI: 1.10–1.41) for death from all causes. We obtained similar results in analyses with adjustment for age and sex only.

HRs for OSCC-specific death were lowered for *MMP1* expression levels that were additionally adjusted for *MMP3* expression (HR, 0.96; 95% CI: 0.62–1.48), and were increased for *MMP3* expression values that were additionally adjusted for *MMP1* expression (HR, 1.30; 95% CI: 0.89–1.92). These patterns were similar for all-cause mortality.

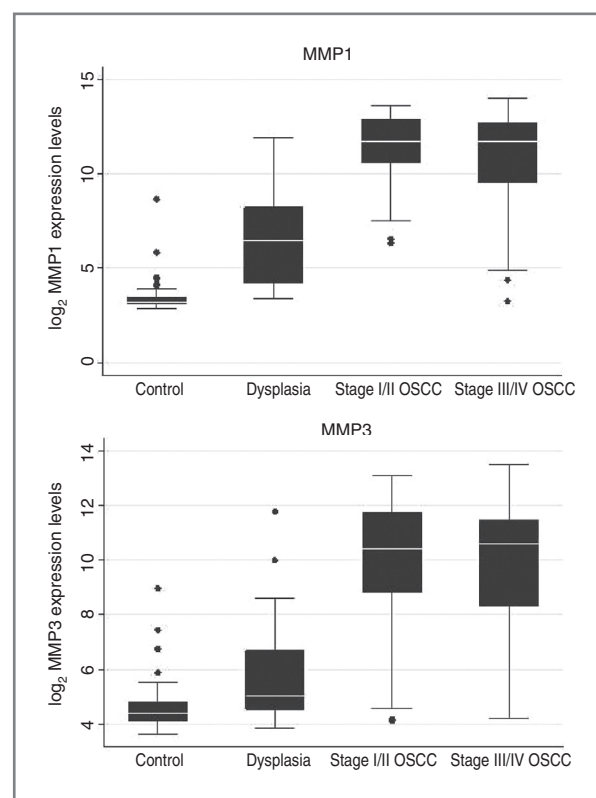
## Discussion

The MMPs are believed to play a key role in tumor metastasis, cell migration, cancer cell growth, and angiogenesis (31–33). They promote cell invasion by cleaving components of the extracellular matrix, with MMP collagenases having the unique ability to cleave native inter-

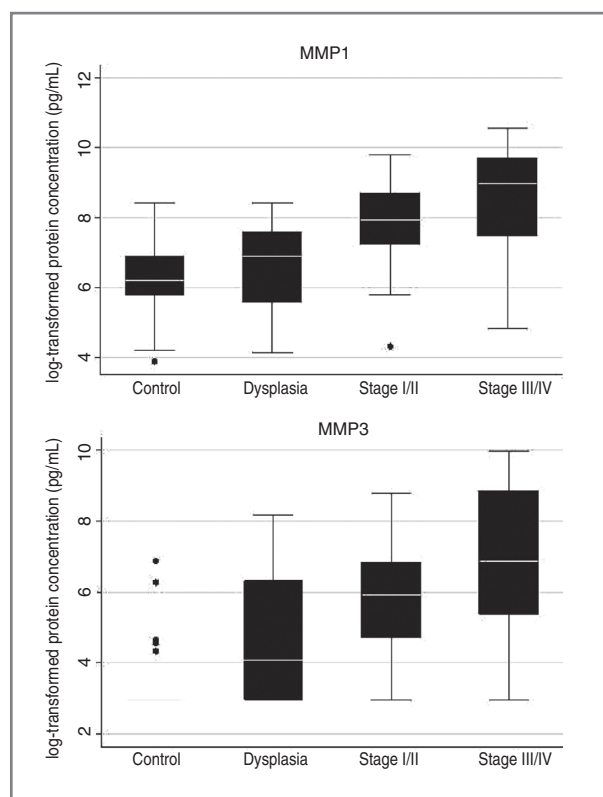
stitial collagen, such as collagen types I, II, III, and IV (34, 35). Among the MMPs, *MMP1* is the most ubiquitously expressed interstitial collagenase and plays a key role in initial cleavage of the extracellular matrix.

**Table 3.** Pearson correlation coefficients between *MMP1*, *MMP3*, *MMP10*, and *MMP12*<sup>a</sup>

|       | MMP1 | MMP3 | MMP10 | MMP12 |
|-------|------|------|-------|-------|
| MMP1  | 1.00 |      |       |       |
| MMP3  | 0.92 | 1.00 |       |       |
| MMP10 | 0.89 | 0.85 | 1.00  |       |
| MMP12 | 0.84 | 0.76 | 0.72  | 1.00  |

<sup>a</sup>All correlations  $P < 0.001$ .

**Figure 1.** log<sub>2</sub> *MMP1* and *MMP3* expression levels across disease categories, University of Washington Affiliated Hospitals, 2003 to 2007. *MMP1* and *MMP3* mRNA expression levels were measured from Affymetrix microarray data, and are expressed as expression levels in the log<sub>2</sub> scale. Plots are shown for expression levels in oral epithelium for: 45 controls, 17 dysplasia patients, 54 stage I/II invasive OSCC patients, and 111 stage III/IV invasive OSCC patients.



**Figure 2.** MMP1 and MMP3 protein concentration levels in saliva across disease categories, University of Washington Affiliated Hospitals, 2003 to 2007. MMP1 and MMP3 protein concentrations were measured in saliva, and are expressed as log-transformed protein concentrations in pg/mL. Plots are shown for protein concentrations in saliva for: 25 controls, 14 dysplasia patients, 33 stage I/II invasive OSCC patients, and 26 stage III/IV invasive OSCC patients.

Results of this study highlight the importance of the MMPs in the development of OSCC. The prominence of these genes is concordant with similar microarray experiments; for example, Ziober and colleagues (7) reported MMPs on their list of 25 genes overexpressed in OSCC. In particular, *MMP1* emerged as an exceptionally strong marker in our study. This gene was expressed at a level more than 200-fold greater in primary OSCC compared with oral epithelium from controls. Importantly, *MMP1* perfectly distinguished OSCC from control tissue in ROC analyses in our independent testing set (AUC = 100%), and almost perfectly discriminated OSCC from control tissue in an external independent testing set (AUC = 98.4%). The substantial differential expression of *MMP1* was confirmed by qRT-PCR. In addition, we analyzed expression of *MMP1* and *MMP3* across different stages in the natural history of OSCC and observed increasing expression corresponding to progression from normal tissue to dysplasia to OSCC (although not from stage I/II to stage III/IV OSCC). The MMPs may thus be ideal molecular markers for monitoring progression from dysplasia to invasive OSCC.

The gene expression levels of the MMPs were highly correlated. In particular, we observed a correlation of 0.92 ( $P < 0.001$ ) for *MMP1* and *MMP3*. Therefore, although each of these genes is a strong biomarker for OSCC, the combination of the 2 does not add to their predictive ability. *MMP1* and *MMP3* are both on chromosome 11 in the region 11q22.3 (36). They are often observed to be coordinately expressed (37, 38), and are activated by similar factors, such as interleukin-1 (39).

Protein concentrations of both MMP1 and MMP3 were observed to be highly elevated in the saliva of OSCC patients compared with saliva from cancer-free controls. Although results did not reach statistical significance, we observed a trend toward higher concentrations for both MMP1 and MMP3 with increasing disease severity. Several studies have evaluated salivary proteins as potential diagnostic tools in OSCC (24, 40–43). Hu and colleagues identified a panel of protein biomarkers in the human salivary proteome and observed that a combination of 5 proteins yielded an AUC of 93% for OSCC detection (40). Shpitzer and colleagues observed increased salivary protein concentrations of 75% for MMP2 and 35% for MMP9 in cases versus controls (42, 43). The MMPs investigated in our study were not among the proteins reported in these previous studies.

Oral cavity and oropharyngeal squamous cell carcinoma are particularly suitable for salivary biomarker measurements due to the presence of potentially abnormal cells/markers sloughed off directly into saliva. Identification of a sensitive and specific marker of OSCC by noninvasive means such as salivary analysis has great potential to assist diagnosis and prognosis (5). People at high risk of developing OSCC, such as heavy smokers, could potentially be monitored for the first indications of OSCC development. Salivary marker analysis could be done in between biopsies to assist in the monitoring of disease status of dysplasia patients (24). Salivary monitoring has the advantage of being less invasive, and provides a mechanism to possibly detect lesions in locations that are difficult to be visualized by general examination. In addition, the best strategy for management of patients presenting with mild to moderate dysplasia or with other atypical lesions is currently unclear. Markers associated with invasive disease (in saliva or the biopsy itself) could assist clinicians in determining which patients should undergo further screening or more invasive treatment.

We obtained different results after stratification by site, with greater case–control differences for MMP gene expression levels and salivary protein concentrations in oral cavity cancer versus oropharyngeal cancer. Similarly, case–control differences for MMP gene expression levels and salivary concentrations were greater among high-risk HPV-negative subjects versus high-risk HPV positive subjects, although the very small number of high-risk HPV-positive controls limit meaningful interpretation of these results. This pattern is congruent with results obtained after stratification by site, as cancers occurring

in the oropharynx, as opposed to the oral cavity, are most strongly associated with HPV infection (44, 45). Higher MMP expression in HPV-negative tumors may be reflective of biological processes that differ from HPV-positive tumors. HPV-negative OSCC is more frequently associated with smoking, and tobacco smoke has been shown to induce *MMP1* mRNA expression in fibroblasts (46, 47) and skin (48). However, we did not observe increased MMP expression associated with smoking for the cell types in our data (results not shown), and the reasons for higher MMP expression in HPV-negative tumors remain unknown. Despite these results, the difference in observed means for gene expression and salivary protein concentrations compared with controls was still substantial within both site and HPV status categories, supporting the use of these MMPs as biomarkers, regardless of site or HPV infection status.

We previously observed *LAMC2* expression level to be highly predictive of OSCC-specific survival (AUC = 80%, CI: 69–91 for *LAMC2* combined with stage; ref. 49); elevated expression of *MMP1* and *MMP3* observed in this study was not as strongly associated with poor survival as *LAMC2*. However, we did observe a moderate elevation in risk of all-cause and OSCC-specific death associated with expression levels of *MMP1* and *MMP3*. However, when *MMP1* and *MMP3* were mutually adjusted, there was not as much evidence of their association with prognosis.

Using more stringent filtering criteria and a different normalization algorithm, we previously identified 131 probe sets (111 unique genes) with mRNA expression that were substantially different between OSCC cases and controls (25). The top predictive model comprised of the genes *LAMC2* and *COL4A1* was able to distinguish cases from controls almost perfectly in our testing set (AUC = 99.8%) and an independent data set (AUC = 97.6%). Taken together, these results suggest that markers identified using different statistical analyses may have the same or similar ability to predict phenotype and represent

the underlying network of biological processes that are involved. Unlike the situation for MMPs, a commercially available immunoassay against well-characterized epitope(s) of *LAMC2* that can be used to test the presence of *LAMC2* in saliva is not yet available.

## Conclusion

Results from our investigations into MMPs lend support to the importance of this family of genes in the pathogenesis and progression of OSCC. The very strong association of *MMP1* with OSCC in particular provides a robust candidate marker for further investigations into its use as a potential marker for OSCC development. The strong associations of MMP salivary protein concentrations with OSCC warrant further investigations into their use as salivary biomarkers.

## Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

## Acknowledgments

We thank the study participants and their families and acknowledge the resources from and use of facilities at the VA Puget Sound Health Care System, University of Washington Medical Center and Harborview Medical Center, Seattle, WA.

## Grant Support

U.S. NIH (R01CA095419 from the National Cancer Institute), NIH 5T32DE007132-28 (Comprehensive training in inter-disciplinary oral health research), and institutional funds from the Fred Hutchinson Cancer Research Center.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 27, 2011; revised September 8, 2011; accepted September 23, 2011; published OnlineFirst September 28, 2011.

## References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;127:2893–917.
2. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA Cancer J Clin* 2010;60:277–300.
3. Poh CF, Zhang L, Anderson DW, Durham JS, Williams PM, Priddy RW, et al. Fluorescence visualization detection of field alterations in tumor margins of oral cancer patients. *Clin Cancer Res* 2006;12:6716–22.
4. Groome PA, Schulze K, Boysen M, Hall SF, Mackillop WJ. A comparison of published head and neck stage groupings in carcinomas of the oral cavity. *Head Neck* 2001;23:613–24.
5. Lee YH, Wong DT. Saliva: an emerging biofluid for early detection of diseases. *Am J Dent* 2009;22:241–8.
6. Charous SJ, Stricklin GP, Nannery LB, Netteville JL, Burkey BB. Expression of matrix metalloproteinases and tissue inhibitor of metalloproteinases in head and neck squamous cell carcinoma. *Ann Otol Rhinol Laryngol* 1997;106:271–8.
7. Ziober AF, Patel KR, Alawi F, Gimotty P, Weber RS, Feldman MM, et al. Identification of a gene signature for rapid screening of oral squamous cell carcinoma. *Clin Cancer Res* 2006;12:5960–71.
8. Nagata M, Fujita H, Ida H, Hoshina H, Inoue T, Seki Y, et al. Identification of potential biomarkers of lymph node metastasis in oral squamous cell carcinoma by cDNA microarray analysis. *Int J Cancer* 2003;106:683–9.
9. Kurahara S, Shinohara M, Ikebe T, Nakamura S, Beppu M, Hiraki A, et al. Expression of MMPs, MT-MMP, and TIMPs in squamous cell carcinoma of the oral cavity: correlations with tumor invasion and metastasis. *Head Neck* 1999;21:627–38.
10. Rutter JL, Mitchell TI, Buttice G, Meyers J, Gusella JF, Ozelius LJ, et al. A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter creates an Ets binding site and augments transcription. *Cancer Res* 1998;58:5321–5.
11. Ye S, Eriksson P, Hamsten A, Kurkinen M, Humphries SE, Henney AM. Progression of coronary atherosclerosis is associated with a common genetic variant of the human stromelysin-1 promoter which results in reduced gene expression. *J Biol Chem* 1996;271:13055–60.
12. Zinzindohoué F, Blons H, Hans S, Lorient MA, Houllier AM, Brasnu D, et al. Single nucleotide polymorphisms in *MMP1* and *MMP3* gene promoters as risk factor in head and neck squamous cell carcinoma. *Anticancer Res* 2004;24:2021–6.



13. Vairaktaris E, Yapijakis C, Derka S, Serefoglou Z, Vassiliou S, Nkenke E, et al. Association of matrix metalloproteinase-1 (-1607 1G/2G) polymorphism with increased risk for oral squamous cell carcinoma. *Anticancer Res* 2007;27:459–64.
14. A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter enhances oral squamous cell carcinoma susceptibility in a Chinese population. *Oral Oncol* 2006;42:32–8.
15. Kainuma K, Katsuno S, Hashimoto S, Oguchi T, Suzuki N, Asamura K, et al. Differences in the expression of genes between normal tissue and squamous cell carcinomas of head and neck using cancer-related gene cDNA microarray. *Acta Otolaryngol* 2006;126:967–74.
16. Suhr ML, Dysvik B, Bruland O, Warnakulasuriya S, Amaratunga AN, Jonassen I, et al. Gene expression profile of oral squamous cell carcinomas from Sri Lankan betel quid users. *Oncol Rep* 2007;18:1061–75.
17. Yen CY, Chen CH, Chang CH, Tseng HF, Liu SY, Chuang LY, et al. Matrix metalloproteinases (MMP) 1 and MMP10 but not MMP12 are potential oral cancer markers. *Biomarkers* 2009;14:244–9.
18. Ye H, Yu T, Temam S, Ziober BL, Wang J, Schwartz JL, et al. Transcriptomic dissection of tongue squamous cell carcinoma. *BMC Genomics* 2008;9:69.
19. Baker EA, Leaper DJ, Hayter JP, Dickenson AJ. The matrix metalloproteinase system in oral squamous cell carcinoma. *Br J Oral Maxillofac Surg* 2006;44:482–6.
20. Lallemand B, Evrard A, Combescure C, Chapuis H, Chambon G, Raynal C, et al. Clinical relevance of nine transcriptional molecular markers for the diagnosis of head and neck squamous cell carcinoma in tissue and saliva rinse. *BMC Cancer* 2009;9:370.
21. Jordan RC, Macabeo-Ong M, Shiboski CH, Dekker N, Ginzinger DG, Wong DT, et al. Overexpression of matrix metalloproteinase-1 and -9 mRNA is associated with progression of oral dysplasia to cancer. *Clin Cancer Res* 2004;10:6460–5.
22. Hack CJ. Integrated transcriptome and proteome data: the challenges ahead. *Brief Funct Genom Proteom* 2004;3:212–9.
23. Brockmann R, Beyer S, Heinisch JJ, Wilhelm T. Posttranscriptional expression regulation: what determines translation rates? *PLoS Comput Biol* 2007;3:e57.
24. Nagler R, Bahar G, Shpitzer T, Feinmesser R. Concomitant analysis of salivary tumor markers—a new diagnostic tool for oral cancer. *Clin Cancer Res* 2006;12:3979–84.
25. Chen C, Méndez E, Houck J, Fan W, Lohavanichbutr P, Doody D, et al. Gene expression profiling identifies genes predictive of oral squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev* 2008;17:2152–62.
26. Lohavanichbutr P, Houck J, Fan W, Yueh B, Mendez E, Futran N, et al. Genomewide gene expression profiles of HPV-positive and HPV-negative oropharyngeal cancer: potential implications for treatment choices. *Arch Otolaryngol Head Neck Surg* 2009;135:180–8.
27. Pyeon D, Newton MA, Lambert PF, den Boon JA, Sengupta S, Marsit CJ, et al. Fundamental differences in cell cycle deregulation in human papillomavirus-positive and human papillomavirus-negative head/neck and cervical cancers. *Cancer Res* 2007;67:4605–19.
28. GeneChip Expression Analysis; Data Analysis Fundamentals. Affymetrix. [http://www.affymetrix.com/support/downloads/manuals/data\\_analysis\\_fundamentals\\_manual.pdf](http://www.affymetrix.com/support/downloads/manuals/data_analysis_fundamentals_manual.pdf). Accessed October 11, 2011.
29. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;5:R80.
30. Partek Genomics Suite. Partek Inc. <http://www.partek.com/partekgs>. Accessed October 11, 2011.
31. Visse R, Nagase H. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res* 2003;92:827–39.
32. Kleiner DE, Stetler-Stevenson WG. Matrix metalloproteinases and metastasis. *Cancer Chemother Pharmacol* 1999;43:S42–51.
33. Rosenthal EL, Matrisian LM. Matrix metalloproteinases in head and neck cancer. *Head Neck* 2006;28:639–48.
34. Egeblad M, Werb Z. New functions for the matrix metalloproteinases in cancer progression. *Nat Rev Cancer* 2002;2:161–74.
35. Kerkelä E, Saarialho-Kere U. Matrix metalloproteinases in tumor progression: focus on basal and squamous cell skin cancer. *Exp Dermatol* 2003;12:109–25.
36. Spurr NK, Gough AC, Gosden J, Rout D, Porteous DJ, van Heyningen V, et al. Restriction fragment length polymorphism analysis and assignment of the metalloproteinases stromelysin and collagenase to the long arm of chromosome 11. *Genomics* 1988;2:119–27.
37. White LA, Maute C, Brinckerhoff CE. ETS sites in the promoters of the matrix metalloproteinases collagenase (MMP-1) and stromelysin (MMP-3) are auxiliary elements that regulate basal and phorbol-induced transcription. *Connect Tissue Res* 1997;36:321–35.
38. Dörr S, Lechtenböhrer N, Rau R, Herborn G, Wagner U, Müller-Myhök B, et al. Association of a specific haplotype across the genes MMP1 and MMP3 with radiographic joint destruction in rheumatoid arthritis. *Arthritis Res Ther* 2004;6:R199–207.
39. Hui A, Min WX, Tang J, Cruz TF. Inhibition of activator protein 1 activity by paclitaxel suppresses interleukin-1-induced collagenase and stromelysin expression by bovine chondrocytes. *Arthritis Rheum* 1998;41:869–76.
40. Hu S, Arellano M, Boonthung P, Wang J, Zhou H, Jiang J, et al. Salivary proteomics for oral cancer biomarker discovery. *Clin Cancer Res* 2008;14:6246–52.
41. St John MA, Li Y, Zhou X, Denny P, Ho CM, Montemagno C, et al. Interleukin 6 and interleukin 8 as potential biomarkers for oral cavity and oropharyngeal squamous cell carcinoma. *Arch Otolaryngol Head Neck Surg* 2004;130:929–35.
42. Shpitzer T, Bahar G, Feinmesser R, Nagler RM. A comprehensive salivary analysis for oral cancer diagnosis. *J Cancer Res Clin Oncol* 2007;133:613–7.
43. Shpitzer T, Hamzany Y, Bahar G, Feinmesser R, Savulescu D, Borovoi I, et al. Salivary analysis of oral cancer biomarkers. *Br J Cancer* 2009;101:1194–8.
44. Mork J, Lie AK, Glatre E, Hallmans G, Jellum E, Koskela P, et al. Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck. *N Engl J Med* 2001;344:1125–9.
45. D'Souza G, Kreimer AR, Viscidi R, Pawlita M, Fakhry C, Koch WM, et al. Case-control study of human papillomavirus and oropharyngeal cancer. *N Engl J Med* 2007;356:1944–56.
46. Kim H, Liu X, Kohyama T, Kobayashi T, Conner H, Abe S, et al. Cigarette smoke stimulates MMP-1 production by human lung fibroblasts through the ERK1/2 pathway. *COPD* 2004;1:13–23.
47. Yin L, Morita A, Tsuji T. Alterations of extracellular matrix induced by tobacco smoke extract. *Arch Dermatol Res* 2000;292:188–94.
48. Lahmann C, Bergemann J, Harrison G, Young AR. Matrix metalloproteinase-1 and skin ageing in smokers. *Lancet* 2001;357:935–6.
49. Méndez E, Houck JR, Doody DR, Fan W, Lohavanichbutr P, Rue TC, et al. A genetic expression profile associated with oral cancer identifies a group of patients at high risk of poor survival. *Clin Cancer Res* 2009;15:1353–61.