Epidermal Growth Factor Receptor Expression and Gene Copy Number in the Risk of Oral Cancer

Mohammed Taoudi Benchekroun¹,⁴, Pierre Saintigny¹,⁴, Sufi M. Thomas⁵,⁷, Adel K. El-Naggar², Vassiliki Papadimitrakopoulou¹, Hening Ren⁶, Wenhua Lang¹, You-Hong Fan¹, Jianhua Huang¹, Lei Feng³, J. Jack Lee³, Edward S. Kim¹, Waun Ki Hong¹,⁴, Faye M. Johnson¹, Jennifer R. Grandis⁵,⁶,⁷, and Li Mao¹,⁸

Abstract

Leukoplakia is the most common premalignant lesion of the oral cavity. Epidermal growth factor receptor (EGFR) abnormalities are associated with oral tumorigenesis and progression. We hypothesized that EGFR expression and gene copy number changes are predictors of the risk of an oral premalignant lesion (OPL) progressing to oral squamous cell carcinoma (OSCC). A formalin-fixed, paraffin-embedded OPL biopsy specimen was collected from each of 162 patients in a randomized controlled clinical trial. We assessed EGFR expression by immunohistochemistry with two methods: a semiquantitative analysis (145 evaluable specimens) and an automated quantitative analysis (127 evaluable specimens). EGFR gene copy number was assessed by fluorescence in situ hybridization (FISH) in a subset of 49 OPLs with high EGFR expression defined by the semiquantitative analysis. We analyzed EGFR abnormalities for associations with OSCC development. High EGFR expression occurred in 103 (71%) of the 145 OPLs and was associated with a nonsignificantly higher risk of OSCC (P = 0.10). Twenty (41%) of 49 OPLs assessed by FISH had an increased EGFR gene copy number (FISH-positive). Patients with FISH-positive lesions had a significantly higher incidence of OSCC than did patients with FISH-negative (a normal copy number) lesions (P = 0.0007). Of note, 10 of 11 OSCCs that developed at the site of the examined OPL were in the FISH-positive group, leaving only one FISH-negative OPL that did so (P < 0.0001). Our data indicate that an increased EGFR gene copy number is common in and associated with OSCC development in patients with OPLs expressing high EGFR, particularly OSCC developing at the site of a high-expression OPL; they also suggest that EGFR inhibitors may prevent oral cancer in patients with OPLs having an increased EGFR gene copy number.

Introduction

Head and neck squamous cell carcinoma (HNSCC) is second only to lung cancer as the most common smoking-related cancer worldwide. Oral squamous cell carcinoma (OSCC) is the most common anatomic site of HNSCC, accounting for ~50% of all HNSCC. Despite the tremendous effort to reduce tobacco use, HNSCC remains one of the leading causes of the ~443,000 deaths in the United States that were attributable to smoking in 2000 to 2004 (1). The only standard therapeutic option for earlier stage HNSCC is surgery, but it is a debilitating, substantially morbid procedure that severely impairs the quality of life for many patients. Despite recent progress in developing targeted therapies for patients with recurrent and metastatic disease, their prognosis remains poor (2). In light of its continuing burden and evasion of substantial control, HNSCC requires new approaches including prevention.

Leukoplakia and erythroplakia are the most commonly diagnosed oral premalignant lesions (OPL), with a 17% to 24% rate of malignant transformation over a period of up to 30 years (3–6). OPLs also are associated with hyperkeratosis, dysplasia, or in situ carcinoma. OPL histology has little value for marking the risk of oral SCC (OSCC). Although loss of heterozygosity (7), podoplanin (8), and p63 expression (9) are associated with an increased risk of OSCC, none of these biomarkers is targetable through currently available drugs.

Epidermal growth factor receptor (EGFR) is believed to play an important role in HNSCC development (10–14); EGFR expression and abnormal gene copy number are associated with a poor prognosis of HNSCC patients (15, 16);
and the anti-EGFR antibody cetuximab is approved for treating HNSCC (17, 18). An EGFR inhibitor for HNSCC prevention is being tested currently in a randomized, placebo-controlled trial (19). The role of EGFR as a marker of OSCC risk, however, has not been evaluated previously in a large series of OPL patients. Establishing EGFR as a reliable marker of OSCC risk would allow selecting a higher risk population with a potentially higher likelihood of benefiting from EGFR inhibitors.

We hypothesized that changes in EGFR protein expression and gene copy number marked the risk of developing OSCC and tested the hypothesis in a large population of OPL patients enrolled in a long-term prospective, randomized controlled trial, the first OPL trial that included long-term oral cancer incidence as a prespecified secondary end point (20). We show that EGFR protein expression and gene copy number may be effective markers of the risk of OPLs for progressing to OSCC.

Patients and Methods

Patients and specimens

All of the 162 randomized and eligible patients who were enrolled in a randomized chemoprevention trial at The University of Texas M. D. Anderson Cancer Center were eligible for this study. From 1992 to 2001, the patients had been diagnosed with OPL and randomly assigned to intervention with 13-cis-retinoic acid (13cRA) versus β-carotene (BC) + retinyl palmitate (RP) versus RP alone. Formalin-fixed, paraffin-embedded biopsy specimens were obtained at enrollment or after enrollment but before any event (defined as the diagnosis of OSCC). Clinicopathologic parameters were obtained from the clinical trial database. The follow-up data were obtained from a combination of chart review and a telephone interview. More detailed clinical information has been previously described in Papadimitrakopoulou et al. (20). The definition of oral cancer development at the same site as an OPL required the cancer and baseline OPL to be on the same side and in the same anatomic structure of the oral cavity. The study was approved by the institutional review board, and written informed consent was obtained from all patients.

EGFR protein expression

Tissue sections (4 μm thick) from formalin-fixed, paraffin-embedded tissue blocks of OPL were mounted on positively charged glass slides. EGFR immunostaining was done using the avidin-biotin peroxidase complex technique, as previously described (8). Briefly, slides were deparaffinized and rehydrated. To retrieve antigenicity, the slides were steamed with 10 mmol/L citrate buffer (pH 6.0; DakoCytomation) for 30 minutes. The slides were then incubated in 10% fetal bovine serum for 30 minutes at room temperature, incubated with monoclonal antibody 31G7 (Zymed Laboratories, Inc.) diluted 1:100 for 90 minutes at room temperature, and subjected to signal development processes using the Vectastain Elite ABC kit according to the manufacturer’s protocol (Vector Labora-

tories). The slides were counterstained with Mayer’s hematoxylin (DakoCytomation). One lung SCC sample and one OPL sample known to express high levels of EGFR were used as positive (primary antibody added) and negative (no primary antibody) controls. No staining was observed in negative controls.

For the semiquantitative evaluation of EGFR expression, each slide was scored for membranous expression as follows: 0, no membrane staining; 1, weak membrane staining in >10% of epithelial cells; 2, intermediate membrane staining in >10% of epithelial cells; and 3, intense membrane staining in >10% of epithelial cells. Because no or weak EGFR expression has been described in normal oral mucosa (21, 22), we classified scores 0 to 1 as low EGFR expression; we classified scores 2 to 3 as high EGFR expression. All scores were based on examining the whole section in each biopsy under a multihheaded microscope by three observers (M.T.B., P.S., and L.M.), who were blinded to clinical patient information.

A second evaluation through an automated analysis system provided an independent, blinded quantitative assessment of EGFR expression. Entire sections were scanned with an Olympus BX61 microscope. Images were acquired and analyzed in the Ariol SL-50 image analyses software. Cytoplasmic and membrane staining were assessed. Composite scores of membrane staining and cytoplasmic staining were obtained by multiplying the number of positive cells by the staining intensity. The sum of the membrane and cytoplasmic scores was used to generate a total composite score. Distribution plots showed that transforming the composite score to its square root divided by 10,000 stabilized variance and brought the data closer to the Gaussian distribution. Therefore, the transformed composite score was used for the final analysis.

EGFR gene copy number

We evaluated EGFR copy number through fluorescence in situ hybridization (FISH), as previously described in ref. (23) and in brief here. Tissue sections (4 μm thick) from formalin-fixed, paraffin-embedded tissue blocks were deparaffinized in citrazol washes, digested with proteinase K, and incubated in denature solution. Sections were then hybridized using the dual-target, dual-color LSI EGFR SpectrumOrange/CEP 7 SpectrumGreen probe (Vysis, Inc.); CEP 7 SpectrumGreen targets the chromosome 7 centromere and serves as a control for copy number normalization. The analysis was done on a BX61 brightfield and epifluorescence microscope (Olympus Bx61, Olympus America) equipped with the Quips XL genetic workstation (Applied Imaging). The EGFR sequence was visualized with a Texas red filter; the chromosome 7 centromere sequence was visualized with a FITC filter; and the nuclei were identified with a 4′,6-diamidino-2-phenylindole filter. Double (FITC and Texas red) and triple band pass filters (4′,6-diamidino-2-phenylindole, FITC, and Texas red; Chroma Technology) were also used. Representative images of each specimen were acquired with a SenSys cooled
CCD camera (Photometrics) in monochromatic layers that were subsequently merged by the SmartCapture software (Vysis). At least 100 nonoverlapping interphase nuclei from whole samples were scored by two independent observers (M.T.B. and P.S.) blinded to clinical information. The number of copies of $EGFR$ probes was assessed independently from that of chromosome 7 probes. Copies of probes for $EGFR$ are usually equal to and balanced in number with copies of probes for chromosome 7, except in the case of $EGFR$ amplification, defined by clustered unbalanced gains of $EGFR$, or a high ratio of $EGFR$ copy number to chromosome 7 copy number. Patients were classified according to the six following and previously described FISH patterns of balanced or unbalanced $EGFR$ and chromosome 7 copy number (24): (a) balanced disomy (in >90% of cells); (b) low, balanced trisomy (10-40% of cells with three copies and <10% of cells with four copies); c) high, balanced trisomy (≥40% of cells with 3 copies and <10% of cells with four copies); (d) low, balanced polysomy (10-40% of cells with four copies); (e) high, balanced polysomy (≥40% of cells with four copies); and (f) $EGFR$ amplification (clustered unbalanced gain of $EGFR$). Because of the preinvasive nature of our samples, we expected an extremely low frequency of polysomy (pattern e) and gene amplification (pattern f) compared with what has been reported for HNSCC. Therefore, our definition of FISH positivity (an increased $EGFR$ gene copy number) for OPLs was expanded (beyond HNSCC definitions) to include any one of patterns b through e (any increase in $EGFR$ gene and chromosome 7 copy number); pattern a was considered to reflect a normal $EGFR$ gene copy number.

**Statistical methods**

The associations between the biomarker expressions, protocol response, and other patient prognostic factors were tested using the $\chi^2$ test or Fisher's exact test for categorical variables and the Wilcoxon rank-sum test for continuous variables. Time to event, such as time to death or time to oral cancer development, was calculated from the treatment randomization date to the event date or last follow-up date if no event had been recorded. The Kaplan-Meier method was used to estimate the event-free rate. The median time to event with 95% confidence intervals (CI) and the event-free rates at years 5 and 10 with 95% CIs by prognostic factors were provided. The log-rank test was used to compare the difference in survival between the prognostic factor groups. Cox proportional hazard models were used for multivariable analysis. Hazard ratios (HR) with 95% CIs and $P$ values were reported. Martingale residual plots were used to visually examine the nature of the relationship between the residuals from a null Cox proportional hazard model (without covariates) and the transformed composite $EGFR$ scores. All tests are two sided, and $P$ values of <0.05 are considered statistically significant.

**Results**

**Patient characteristics**

Biopsies from 17 (11%) of the 162 patients enrolled in the chemoprevention trial were excluded from the analyses because of a lack of sufficient tissue in blocks ($n = 12$) or the absence of evaluable epithelial cells in the H&E section ($n = 5$). In 18 (12%) of the 145 remaining patients, biopsy specimens were obtained after enrollment because the baseline-biopsy paraffin blocks were unavailable. Median follow-up was 7.5 years, with 35 (24%) of the 145 patients developed oral cancer. Seventeen oral cancers developed at the site of a baseline OPL, and 18 oral cancers developed at a site that was contralateral to and/or different from the site of any baseline OPL. The study population was balanced for gender; the majority of the patients were white, with current or former smoking and alcohol history. Half of the patients received 13cRA and half received RP or RP plus BC. Two thirds of the OPLs were classified as hyperplasia, and one-third was classified as dysplasia of various degrees.

**Semiquantitative EGFR protein expression in OPLs**

Analyzed in 145 specimens, EGFR expression was mostly membranous, predominant in the basal layers of the epithelium and was observed in the vast majority of samples. A total of 42 (29%) OPLs were scored 0 or 1—17 (12%) 0s and 25 (17%) 1s—and thus were considered to have a low EGFR expression. A total of 103 OPLs (71%) were scored 2 or 3—51 (35%) 2s and 52 (36%) 3s—and thus were considered to have a high EGFR expression. In most cases, EGFR expression was homogeneous in the epithelium. Examples of EGFR expression are shown in Fig. 1A to C.

**Semiquantitative EGFR expression and clinicopathologic parameters**

High EGFR expression was more frequent in OPLs from females (79.4%) versus males (63.6%, $P = 0.03$), whites (73.6%) versus nonwhites (50%, $P = 0.04$), and older versus younger patients (high expression in a median age of 59 y versus low expression in a median age of 49 y; $P = 0.003$). There was no association between EGFR expression and histologic status, smoking history, history of alcoholic consumption, or treatment arm.

**EGFR protein expression and oral cancer risk**

A trend between OPLs with high EGFR expression (scored 2 or 3) and a higher risk of oral cancer development was observed in the univariate analysis (Table 1; Fig. 2A), which also showed that oral cancer development had no significant association with sex, race, age, smoking, or alcohol history and had a borderline association with OPL histologic status ($P = 0.06$). In a multivariable analysis, neither OPL histology at baseline nor EGFR expression was significantly associated with time to oral cancer (data not shown).
To further study the association between EGFR expression and oral cancer risk, we quantitatively evaluated EGFR expression using an automated analysis system in 127 samples. Increased total transformed composite EGFR score was significantly associated with oral cancer development, with a HR of 1.187 ($P = 0.012; 95\%\ CI, 1.039-1.356$). Increased total transformed composite EGFR score also was significantly associated with oral cancer development in a multicovariate analysis including age, histology at baseline, and treatment arm (HR = 1.147; $P = 0.036; 95\%\ CI, 1.01-1.30$; Table 2). Time to oral cancer was not statistically different between patients with high-EGFR-expression OPLs (defined by median total transformed score) and patients with low-EGFR-expression OPLs (Fig. 3A). A Martingale residual analysis revealed a linear trend of increasing oral cancer risk beginning with a total transformed composite EGFR score of 7. With a cut-off point at the score 7, time to oral cancer was significantly worse in patients with high EGFR expression (Fig. 3B).

**EGFR gene copy number**

The borderline association between oral cancer development and high-EGFR-expression OPLs (Figs. 2A and 3) led us to hypothesize that this trend involved the subset of high-expression patients who also had an increased EGFR gene copy number. Because of a scarcity of tissue, we could only evaluate 60 of the 103 high-EGFR-expression OPL patients (semiquantitative evaluation), including 29 who developed oral cancer and 31 who did not (Table 3). Among these 60 patients, 49 exhibited at least 100 non-overlapping interphase nuclei from the whole sample and so were included in subsequent analyses. Of these 49 patients, 20 OPLs (41\%) were FISH positive, or had a high number of $EGFR$ gene and chromosome 7 copies, distributed as follows: 14 with low trisomy, 1 with high trisomy, 4 with low polysomy, and 1 with gene amplification. The remaining 29 OPLs (59\%) were FISH negative, or had a low copy number (disomy; Fig. 1D-F). We did not find any association between FISH positivity and the degree of dysplasia (Supplementary Table S4). This comparison, however, is limited by its small sample size ($n = 49$).

**Impact of FISH-positivity on oral cancer–free rates**

Patients with a FISH-positive OPL had a significantly higher incidence of OSCC than did those with a FISH-negative OPL (log-rank test, $P = 0.0007$; Fig. 2B). The difference was even more striking when considering OSCC that developed at the site of a baseline OPL (log-rank test, $P < 0.0001$; Fig. 2C). There was no difference between FISH-positive and FISH-negative OPLs with regard to the incidence of OSCC not at a baseline OPL site (data not shown). The oral cancer–free rate (OCF) was only 40\% for patients with FISH-positive OPLs (95\% CI, 0.23-0.68) versus 79\% for patients with FISH-negative OPLs (95\% CI, 0.65-0.95; $P = 0.0007$) at 5 years after biopsy. The difference in OCF rate was more pronounced at
10 years after biopsy—only 16% (95% CI, 0.05-0.53) in the FISH-positive group versus 67% (95% CI, 0.5-0.9) in the FISH-negative group ($P = 0.0007$). In a multivariate analysis, the only covariate significantly associated with OCF was FISH positivity, with a HR of 3.620 (95% CI, 1.439-9.104; Table 4).

**Discussion**

In the present study, patients with high-EGFR-expression OPLs had a statistically significantly decreased OCF if their OPL also carried increased chromosome 7 and $EGFR$ gene copy numbers (16%; 95% CI, 0.05-0.53) versus carrying a normal $EGFR$ gene copy number (67%; 95% CI, 0.5-0.9) at 10 years ($P = 0.0007$). This finding clearly shows that an increased chromosome 7 and $EGFR$ gene copy number is an early event in oral tumorigenesis that is strongly associated with oral cancer risk.

To our knowledge, our study is the first to report $EGFR$ expression and gene copy number in OPLs in a series of longitudinal and prospectively collected samples, which came from the largest, longest term randomized controlled trial ever conducted in OPL patients (20).

Assessing cancer risk has the potential to help lower cancer incidence and mortality by providing the most appropriate populations for clinical prevention research. Although podoplanin, loss of heterozygosity (7), and p63 (9) have been shown to associate with an OPL's increased risk of OSCC (8), there are no investigational or clinically approved agents for targeting these abnormalities. EGFR, on the other hand, is a validated cancer treatment target, and an anti-EGFR antibody, cetuximab, has been approved for treating patients with HNSCC and several other types of cancer (17, 18).

A seminal study in non–small cell lung carcinoma (25) defined the FISH patterns of chromosome 7 and $EGFR$ gene copy number that we used, but with substantially different definitions of FISH-positive and FISH-negative tumors. The definition for FISH positivity, or a high $EGFR$ copy number, only included high polysomy ($\geq 4$ gene copies in $\geq 40\%$ cells) or $EGFR$ amplification (unbalanced gene-to-chromosome copy number ratio of $>2e$ or $\geq 15$ gene copies in $\geq 10\%$ of cells); FISH negativity was defined as low polysomy, high or low trisomy, or disomy. These patterns and definitions of FISH status were applied in previous studies of HNSCC (15, 16). Chung et al. (15) found $EGFR$ gene amplification in 31%, high polysomy in 27%, low polysomy in 17%, trisomy in 17%, and balanced disomy in 8% of 81 cases of HNSCC, resulting in 58% FISH-positive cases of HNSCC. We did not expect

### Table 1. The median time (y) to oral cancer development and OCF rates at years 5 and 10 for all patients ($n = 145$)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level</th>
<th>N</th>
<th>Event</th>
<th>Median (95% CI)</th>
<th>OCF rate at 5 y (95% CI)</th>
<th>OCF rate at 10 y (95% CI)</th>
<th>$P^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients</td>
<td>145</td>
<td>35</td>
<td>NA</td>
<td>0.8 (0.74-0.87)</td>
<td>0.72 (0.65-0.81)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>Female</td>
<td>68</td>
<td>16</td>
<td>NA</td>
<td>0.78 (0.69-0.89)</td>
<td>0.73 (0.62-0.86)</td>
<td>0.843</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>77</td>
<td>19</td>
<td>NA</td>
<td>0.82 (0.74-0.92)</td>
<td>0.72 (0.62-0.84)</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>Other</td>
<td>16</td>
<td>3</td>
<td>NA</td>
<td>0.85 (0.67-1)</td>
<td>0.75 (0.54-1)</td>
<td>0.719</td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>129</td>
<td>32</td>
<td>NA</td>
<td>0.8 (0.73-0.87)</td>
<td>0.72 (0.64-0.81)</td>
<td></td>
</tr>
<tr>
<td>Histologic status</td>
<td>Hyperplasia</td>
<td>97</td>
<td>19</td>
<td>NA</td>
<td>0.85 (0.78-0.92)</td>
<td>0.78 (0.70-0.88)</td>
<td>0.064</td>
</tr>
<tr>
<td></td>
<td>Mild dysplasia</td>
<td>37</td>
<td>11 (6.65-NA)</td>
<td>0.74 (0.61-0.9)</td>
<td>0.65 (0.50-0.85)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate/severe dysplasia</td>
<td>11</td>
<td>5</td>
<td>6.67 (1.85-NA)</td>
<td>0.61 (0.38-1)</td>
<td>0.46 (0.22-0.97)</td>
<td></td>
</tr>
<tr>
<td>Treatment arm</td>
<td>13cRA</td>
<td>73</td>
<td>18</td>
<td>NA</td>
<td>0.78 (0.69-0.89)</td>
<td>0.72 (0.61-0.84)</td>
<td>0.565</td>
</tr>
<tr>
<td></td>
<td>BC + RP</td>
<td>40</td>
<td>8</td>
<td>NA</td>
<td>0.84 (0.73-0.97)</td>
<td>0.8 (0.68-0.95)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RP</td>
<td>32</td>
<td>9</td>
<td>(6.65-NA)</td>
<td>0.8 (0.67-0.96)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking history</td>
<td>Current</td>
<td>51</td>
<td>7</td>
<td>NA</td>
<td>0.87 (0.78-0.97)</td>
<td>0.83 (0.72-0.96)</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>Former</td>
<td>58</td>
<td>20</td>
<td>(6.65-NA)</td>
<td>0.75 (0.64-0.87)</td>
<td>0.62 (0.50-0.78)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Never</td>
<td>36</td>
<td>8</td>
<td>NA</td>
<td>0.8 (0.68-0.94)</td>
<td>0.76 (0.63-0.92)</td>
<td></td>
</tr>
<tr>
<td>Alcohol history</td>
<td>Current</td>
<td>80</td>
<td>18</td>
<td>NA</td>
<td>0.84 (0.76-0.93)</td>
<td>0.74 (0.64-0.86)</td>
<td>0.783</td>
</tr>
<tr>
<td></td>
<td>Former</td>
<td>19</td>
<td>5</td>
<td>(6.21-NA)</td>
<td>0.78 (0.61-1)</td>
<td>0.68 (0.48-0.98)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Never</td>
<td>46</td>
<td>12</td>
<td>NA</td>
<td>0.75 (0.63-0.89)</td>
<td>0.72 (0.6-0.87)</td>
<td></td>
</tr>
<tr>
<td>EGFR protein expression</td>
<td>Low</td>
<td>42</td>
<td>6</td>
<td>NA</td>
<td>0.89 (0.8-1)</td>
<td>0.82 (0.69-0.96)</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>103</td>
<td>29</td>
<td>(10.7-NA)</td>
<td>0.77 (0.69-0.86)</td>
<td>0.69 (0.6-0.8)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: NA, not applicable.

$P^*$ values are from the log-rank test (univariate analysis).

$^1$Semiquantitative evaluation; EGFR low: score 0-1 at immunohistochemistry; EGFR high: score 2-3 at immunohistochemistry.
to see this high HNSCC frequency of high polysomy or gene amplification in preinvasive lesions. Therefore, we could not use the same definition of FISH positivity in our study, in which the majority of OPLs exhibited either a disomy (59%) or trisomy (31%). Therefore, our definition for FISH positivity, or an increased \textit{EGFR} gene copy number, included trisomy along with polysomy or gene amplification.

Using an independent prospective population, a dual-target and dual-color FISH assay, and recently described FISH patterns, our results are consistent with our earlier findings (26, 27) showing that any increase in copy number of chromosome 7 (called “polysomy” in these early articles) is a major risk factor for oral cancer in OPL patients. These earlier studies built on the then-established association between chromosome polysomy and an increased risk of oral cancer (28, 29), evaluating chromosomes 7 and 17 centromeres through chromogenic \textit{in situ} hybridization. The frequency of chromosome polysomy in the tumor field was shown to increase as the tissue progressed from normal morphology (33% frequency) to hyperplasia (67%) to dysplasia (95%) to SCC (96%). Subsequently, we analyzed OPL biopsies collected in a randomized chemoprevention trial with a median follow-up of 7 years (6). Patients with more than three chromosome 7 copy numbers in at least 3% of epithelial cells were at an increased risk of oral cancer versus patients with lesser copy numbers (HR = 1.85; 95% CI, 1.05-3.25; \( P = 0.03 \)).

Our findings also suggest that an increased \textit{EGFR} gene copy number in OPLs is a precursor to \textit{EGFR} gene amplification in HNSCC (as is chromosome 7 increased copy number) and an important oncogenesis-driving effector in oral oncogenesis. Sheu et al. (12) recently conducted functional genomic analyses showing that 7p11.2 was the most frequently amplified region in OSCC, and mapping this region showed a unique amplicon containing \textit{SEC61G} and \textit{EGFR} genes. The expression level of \textit{EGFR} but not of \textit{SEC61G} was upregulated and tightly correlated with DNA copy number. Furthermore, \textit{EGFR} downstream effectors such as K-ras, mitogen-activated protein kinase 1, and cyclin D1 also were amplified or mutated, resulting in the activation of \textit{EGFR} signaling in 55% of OSCC patients. Another study validated these findings through array-comparative gene hybridization in which amplification of 7p12 including \textit{EGFR} was frequent in HNSCC (30).

Taken together, all these findings support our current data indicating that an increased \textit{EGFR} gene copy number is an early event in oral oncogenesis, consistent with its effect on the oral cancer development of OPL patients. Our and these other data also strongly suggest that \textit{EGFR} is a major independent driver of oral oncogenesis as it progresses continuously from chromosomal instability to \textit{EGFR} trisomy to \textit{EGFR} polysomy and ultimately to \textit{EGFR} amplification, all resulting in an unbalanced chromosome 7 polysomy.

Previous reports have shown that \textit{EGFR} expression increases dramatically with progression from dysplastic lesions to HNSCC (22, 31), although \textit{EGFR} expression as a prognostic factor in HNSCC is controversial (32); it also increases in normal epithelium adjacent to HNSCC compared with normal epithelium of healthy controls, as is consistent with “field cancerization” (31, 33). Analyzing only OPLs in the present study, we found \textit{EGFR} expression in 88% and high \textit{EGFR} expression in 71%. Consistent with previous reports (22, 31), \textit{EGFR} expression did not differ between hyperplasia and dysplasia in the present study. To the best of our knowledge, the effect of \textit{EGFR} expression on oral cancer development has never been reported in a prospective series of patients (from a randomized controlled trial, in this case). The trend that our semiquantitative evaluation of \textit{EGFR} expression showed between high \textit{EGFR} protein expression and oral cancer development remained in our automated analysis of \textit{EGFR}.

Fig. 2. Time to oral cancer in oral leukoplakias (OPL) by \textit{EGFR} protein expression scored 0 to 1+ versus 2 to 3+ (A); time to oral cancer (B) and time to same-site oral cancer (C) in OPL by \textit{EGFR} gene copy number. E/N, number of events and number of patients.
expression, which allowed a quantitative, less-subjective evaluation. Our study possibly was underpowered to detect a statistically significant difference in time to oral cancer based on this factor alone. EGFR expression is associated with smoking history and is significantly higher in lung SCC (a histologic subtype strongly associated with smoking habits) than in lung adenocarcinoma (34). In our study, we did not observe any association between EGFR expression and smoking history. Tobacco smoke exposure induced an EGFR-centered subnetwork, in which EGFR and its ligands were all significantly induced, in a cellular model of oral leukoplakia (35). It is possible that EGFR activation by tobacco smoking preferentially stimulates the expression of EGFR ligands such as amphiregulin rather than EGFR itself (36). Fundamental differences have been reported between human papillomavirus (HPV)-positive and HPV-negative murine models of HNSCC (37) and human oropharyngeal or oral HNSCC (37, 38). In North America, the overall HPV prevalence is 16% in OSCC and 47% in oropharyngeal cancer (39), and HPV-positive oropharyngeal cancer has been associated with a low EGFR expression (40). Therefore, some of the OSCC cases with low EGFR expression in our study possibly were HPV related, which may have decreased our study power for identifying EGFR-driven OSCC.

It also should be noted that the effect of chemopreventive agents in the clinical trial of the present study may be a confounding factor that influenced our results. Although there was no significant difference in oral cancer development among the treatment groups (20), it remains to be determined whether any treatment or treatments in this trial interacted in any way with the status of EGFR expression and copy number in OPLs.

Anti-EGFR antibodies and EGFR tyrosine kinase inhibitors (TKIs) are the most widely used strategies for inhibiting EGFR. For chemoprevention, however, EGFR TKIs, which are oral and convenient, are preferred over the antibodies, which require the inconvenience of i.v. administration. Increased EGFR protein expression and gene copy number, and TK domain–activating mutations are the most studied mechanisms associated with response to EGFR inhibitors (41, 42). EGFR gene copy number abnormalities are well known in HNSCC (15, 16). Because past reports have not shown differences of EGFR gene copy number in laryngeal, pharyngeal, and oral carcinogenesis (15, 16), we speculate

Table 2. Multicovariate analysis of time to oral cancer on quantitative EGFR expression (E/N = 29/127)

<table>
<thead>
<tr>
<th>Variable</th>
<th>HR</th>
<th>95% CI for HR</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.05</td>
<td>1.01-1.08</td>
<td>0.005</td>
</tr>
<tr>
<td>13cRA vs BC-RP/RP only</td>
<td>0.83</td>
<td>0.39-1.76</td>
<td>0.626</td>
</tr>
<tr>
<td>Histologic status at baseline: dysplasia vs hyperplasia</td>
<td>2.99</td>
<td>1.42-6.29</td>
<td>0.004</td>
</tr>
<tr>
<td>EGFR transformed total composite score</td>
<td>1.15</td>
<td>1.01-1.30</td>
<td>0.036</td>
</tr>
</tbody>
</table>

NOTE: For this analysis, the Cox proportional models were fitted with other important factors in the models. Composite scores of membrane staining and cytoplasmic staining were obtained by multiplying the number of positive cells by the intensity. The sum of the membrane and cytoplasmic scores was used to generate a total composite EGFR score, which then was transformed by dividing its square root by 10,000. This analysis produced the maximum likelihood estimates.

Abbreviations: E, number of events; N, number of patients.

Fig. 3. Time to oral cancer in oral leukoplakias (OPL) by EGFR expression evaluated quantitatively by automated analysis. A, the median of the EGFR total transformed composite (TTC) score was used to dichotomize low versus high EGFR expression. B, an EGFR total transformed composite score of 7 was used as a cutoff, based on Martingale analysis.
that our findings may be relevant to more head and neck sites than just the oral cavity. On the other hand, EGFR mutations have rarely been described in HNSCC tumors or HNSCC cell lines and therefore probably are irrelevant in OPLs (16).

Early reports of a correlation between EGFR expression and response to EGFR TKIs (17, 43, 44) led to controversial subsequent results. Preclinical and clinical HNSCC studies have found an association between EGFR expression or gene copy number and response to EGFR TKIs. Sheu et al. (12) found that only cell lines with EGFR gene amplification or EGFR overexpression were sensitive to an EGFR TKI. In 18 HNSCC cell lines, EGFR overexpression correlated with sensitivity to the EGFR TKI gefitinib and EGFR gene amplification occurred in the most sensitive cell lines (13). Other studies found similar results (45). EGFR overexpression and increased gene copy number were associated with a trend toward higher objective clinical response rates in a phase I/II trial of the EGFR TKI erlotinib combined with cisplatin in 37 patients with recurrent or metastatic HNSCC (46). These patient response rates were 36% with strong, 12% with medium, and 0% with

### Table 3. The median time (y) to oral cancer development and OCF rates at years 5 and 10 in the subgroup of 49 patients with high EGFR expression and evaluation of EGFR gene copy number

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level</th>
<th>N</th>
<th>Event</th>
<th>Median (95% CI)</th>
<th>OCF rate at 5 y (95% CI)</th>
<th>OCF rate at 10 y (95% CI)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients</td>
<td></td>
<td>49</td>
<td>24</td>
<td>6.65 (3.52-NA)</td>
<td>0.62 (0.5-0.78)</td>
<td>0.46 (0.32-0.65)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>Female</td>
<td>23</td>
<td>11</td>
<td>6.67 (2.24-NA)</td>
<td>0.54 (0.37-0.8)</td>
<td>0.47 (0.28-0.76)</td>
<td>0.737</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>26</td>
<td>13</td>
<td>6.65 (5.22-NA)</td>
<td>0.69 (0.53-0.89)</td>
<td>0.46 (0.28-0.74)</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>Other</td>
<td>1</td>
<td>0</td>
<td>NA</td>
<td></td>
<td></td>
<td>0.489</td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>48</td>
<td>24</td>
<td>6.65 (3.52-NA)</td>
<td>0.61 (0.49-0.77)</td>
<td>0.45 (0.32-0.64)</td>
<td></td>
</tr>
<tr>
<td>Histologic status</td>
<td>Hyperplasia</td>
<td>33</td>
<td>14</td>
<td>10.7 (5.22-NA)</td>
<td>0.68 (0.54-0.87)</td>
<td>0.53 (0.36-0.77)</td>
<td>0.454</td>
</tr>
<tr>
<td></td>
<td>Mild dysplasia</td>
<td>10</td>
<td>6</td>
<td>5.09 (2.05-NA)</td>
<td>0.5 (0.27-0.93)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moderate/severe dysplasia</td>
<td>6</td>
<td>4</td>
<td>4.45 (1.85-NA)</td>
<td>0.5 (0.22-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment arm</td>
<td>13cRA</td>
<td>24</td>
<td>13</td>
<td>6.21 (2.88-NA)</td>
<td>0.53 (0.36-0.78)</td>
<td>0.41 (0.24-0.69)</td>
<td>0.452</td>
</tr>
<tr>
<td></td>
<td>BC + RP</td>
<td>18</td>
<td>7</td>
<td>10.7 (5.3-NA)</td>
<td>0.72 (0.54-0.96)</td>
<td>0.62 (0.4-0.94)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RP</td>
<td>7</td>
<td>4</td>
<td>5.22 (2.05-NA)</td>
<td>0.71 (0.45-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking history</td>
<td>Current</td>
<td>16</td>
<td>5</td>
<td>NA (6.67-NA)</td>
<td>0.75 (0.57-1)</td>
<td></td>
<td>0.471</td>
</tr>
<tr>
<td></td>
<td>Former</td>
<td>25</td>
<td>15</td>
<td>6.21 (2.88-NA)</td>
<td>0.56 (0.4-0.79)</td>
<td>0.39 (0.23-0.67)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Never</td>
<td>8</td>
<td>4</td>
<td>5.3 (3.36-NA)</td>
<td>0.6 (0.33-1)</td>
<td>0.45 (0.2-1)</td>
<td></td>
</tr>
<tr>
<td>Alcohol history</td>
<td>Current</td>
<td>29</td>
<td>12</td>
<td>10.7 (6.65-NA)</td>
<td>0.72 (0.58-0.91)</td>
<td>0.56 (0.38-0.81)</td>
<td>0.272</td>
</tr>
<tr>
<td></td>
<td>Former</td>
<td>8</td>
<td>5</td>
<td>3.32 (2.27-NA)</td>
<td>0.47 (0.21-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Never</td>
<td>12</td>
<td>7</td>
<td>3.36 (1.85-NA)</td>
<td>0.44 (0.21-0.92)</td>
<td>0.29 (0.1-0.87)</td>
<td></td>
</tr>
<tr>
<td>EGFR gene copy number§</td>
<td>FISH−</td>
<td>29</td>
<td>8</td>
<td>NA</td>
<td>0.79 (0.65-0.95)</td>
<td>0.67 (0.5-0.9)</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td>FISH+</td>
<td>20</td>
<td>16</td>
<td>2.88 (2.05-NA)</td>
<td>0.4 (0.23-0.68)</td>
<td>0.16 (0.05-0.53)</td>
<td></td>
</tr>
</tbody>
</table>

*P values are from the log-rank test (univariate analysis).
§FISH−: disomy; FISH+: low trisomy, high trisomy, low polysomy, or gene amplification.

that our findings may be relevant to more head and neck sites than just the oral cavity. On the other hand, EGFR mutations have rarely been described in HNSCC tumors or HNSCC cell lines and therefore probably are irrelevant in OPLs (16).

Early reports of a correlation between EGFR expression and response to EGFR TKIs (17, 43, 44) led to controversial subsequent results. Preclinical and clinical HNSCC studies have found an association between EGFR expression or gene copy number and response to EGFR TKIs. Sheu et al. (12) found that only cell lines with EGFR gene amplification or EGFR overexpression were sensitive to an EGFR TKI. In 18 HNSCC cell lines, EGFR overexpression correlated with sensitivity to the EGFR TKI gefitinib and EGFR gene amplification occurred in the most sensitive cell lines (13). Other studies found similar results (45). EGFR overexpression and increased gene copy number were associated with a trend toward higher objective clinical response rates in a phase I/II trial of the EGFR TKI erlotinib combined with cisplatin in 37 patients with recurrent or metastatic HNSCC (46). These patient response rates were 36% with strong, 12% with medium, and 0% with

### Table 4. Multicovariate analysis for time to oral cancer on FISH status for EGFR gene and chromosome 7 copies (E/N = 24/49)

<table>
<thead>
<tr>
<th>Variable</th>
<th>HR</th>
<th>95% CI for HR</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.01</td>
<td>0.98-1.04</td>
<td>0.577</td>
</tr>
<tr>
<td>BC+RP/RP-only vs 13cRA</td>
<td>1.19</td>
<td>0.52-2.76</td>
<td>0.682</td>
</tr>
<tr>
<td>Histologic status at baseline: dysplasia vs hyperplasia</td>
<td>1.02</td>
<td>0.42-2.46</td>
<td>0.965</td>
</tr>
<tr>
<td>FISH: positive vs negative</td>
<td>3.62</td>
<td>1.44-9.10</td>
<td>0.006</td>
</tr>
</tbody>
</table>

NOTE: This analysis produced the maximum likelihood estimates.
Abbreviation: E/N, number of events and number of patients.
weak or absent EGFR expression and 50% with, versus 15% without, high EGFR polysomy or gene amplification.

Tang et al. (47) reported that EGFR levels were elevated in OPLs versus in control samples, suggesting that increased EGFR protein expression marks carcinogenesis, in a murine model of oral carcinogenesis induced by the tobacco surrogate 4-nitroquinoline-1-oxide; the level of DNA damage in this model is associated with the development of OPLs and SCC (21). Sheu et al. (12) showed that the EGFR TKI AG1478 dramatically reduced the incidence of OSCC and high-grade dysplasia in mice that developed oral leukoplakia of the tongue within a model of OSCC induced by 4 weeks of combined arecoline and 4-nitroquinoline-1-oxide. EGFR gene copy number, however, has not been assessed in 4-nitroquinoline-1-oxide–based oral mouse models. Last, EGFR inhibition has been shown to downregulate signaling molecules such as cyclin D1 (which is implicated in genetic instability) downstream of the EGFR-signal transducer and activator of transcription 3 pathway (48, 49).

Taken together, these results from in vitro and in vivo models allow us to hypothesize that EGFR gene copy number may be valuable in predicting response to EGFR TKIs in the chemoprevention setting. It is also possible, however, that a change in EGFR copy number may reflect mainly chromosome 7 polysomy or aneuploidy and not a change directly linked to EGFR. If this is the case, increased EGFR gene copy number may be a marker only of oral cancer risk and not of drug sensitivity. Correlative studies in the ongoing phase III Erlotinib Prevention of Oral Cancer trial in OPL patients with a high OSCC risk marked by a prospective longitudinal clinical trial that an increased EGFR gene copy number marks the risk of OSCC development in the substantial subgroup of OPL patients who have EGFR overexpression (71% of our total study population). Follow-on study in larger cohorts will be necessary to validate these findings. Assessments of EGFR protein expression and gene copy number could lead to selecting patients most in need of and most likely to respond to EGFR inhibitors in future oral cancer chemoprevention trials.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

National Cancer Institute grants P01 CA106451, P50 CA97007, and P30 CA16672 and a grant from the Kadence Charitable Foundation.

Received 08/06/2009; revised 12/08/2009; accepted 12/22/2009; published OnlineFirst 06/22/2010.

References

7. Mao L, Lee JS, Fan YH, et al. Frequent microsatellite alterations at both EGFR-dependent and EGFR-independent mechanisms (50, 51). Therefore, it is possible that the effectiveness of EGFR inhibition may be increased by combining it with Akt-mamalian target of rapamycin inhibition. Another approach in development is EGFR antisense DNA therapy, which has been tested in a phase I (dose escalation) clinical trial involving 17 assessable patients with HNSCC that was injected with EGFR antisense DNA therapy weekly for 4 weeks (14). No grades 3 to 4 or dose-limiting toxicities were reported, and a maximum-tolerated dose was not reached. Five of the 17 patients had an objective response, including two complete responses. Disease control (objective response plus stable disease) was associated with baseline EGFR expression. This alternative approach of EGFR targeting may lend itself to oral leukoplakia, which is easily accessible and frequently involves only one or a few lesions.

In conclusion, the present study provides the first demonstration in a prospective longitudinal clinical trial that an increased EGFR gene copy number marks the risk of OSCC development in the substantial subgroup of OPL patients who have EGFR overexpression (71% of our total study population). Follow-on study in larger cohorts will be necessary to validate these findings. Assessments of EGFR protein expression and gene copy number could lead to selecting patients most in need of and most likely to respond to EGFR inhibitors in future oral cancer chemoprevention trials.

Disclosed Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

National Cancer Institute grants P01 CA106451, P50 CA97007, and P30 CA16672 and a grant from the Kadence Charitable Foundation.

Received 08/06/2009; revised 12/08/2009; accepted 12/22/2009; published OnlineFirst 06/22/2010.
EGFR Expression and Gene Copy Number in Oral Preneoplasia


Epidermal Growth Factor Receptor Expression and Gene Copy Number in the Risk of Oral Cancer


Updated version
Access the most recent version of this article at:
doi:10.1158/1940-6207.CAPR-09-0163

Supplementary Material
Access the most recent supplemental material at:
http://cancerpreventionresearch.aacrjournals.org/content/suppl/2010/06/21/1940-6207.CAPR-09-0163.DC1
http://cancerpreventionresearch.aacrjournals.org/content/suppl/2010/06/29/1940-6207.CAPR-09-0163.DC2

Cited articles
This article cites 51 articles, 30 of which you can access for free at:
http://cancerpreventionresearch.aacrjournals.org/content/3/7/800.full.html#ref-list-1

Citing articles
This article has been cited by 16 HighWire-hosted articles. Access the articles at:
/content/3/7/800.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.