New DNA Methylation Markers and Global DNA Hypomethylation Are Associated with Oral Cancer Development

Jean-Philippe Foy1,2,3, Curtis R. Pickering4, Vassiliki A. Papadimitrakopoulou5, Jaroslav Jelinek6, Steven H. Lin7, William N. William Jr8, Mitchell J. Frederick4, Jing Wang9, Wenhua Lang5, Lei Feng9, Li Zhang8, Edward S. Kim10, You H. Fan5, Waun K. Hong11, Adel K. El-Naggar12, J. Jack Lee9, Jeffrey N. Myers4, Jean-Pierre Issa6, Scott M. Lippman13, Li Mao14, and Pierre Saintigny1,2,15,16

Abstract

DNA promoter methylation of tumor suppressor genes and global DNA hypomethylation are common features of head and neck cancers. Our goal was to identify early DNA methylation changes in oral premalignant lesions (OPL) that may serve as predictive markers of developing oral squamous cell carcinoma (OSCC). Using high-throughput DNA methylation profiles of 24 OPLs, we found that the top 86 genes differentially methylated between patients who did or did not develop OSCC were simultaneously hypermethylated, suggesting that a CpG island methylation phenotype may occur early during OSCC development. The vast majority of the 86 genes were nonmethylated in normal tissues and hypermethylated in OSCC versus normal mucosa. We used pyrosequencing in a validation cohort of 44 patients to evaluate the degree of methylation of AGTR1, FOXI2, and PENK promoters CpG sites that were included in the top 86 genes and of LINE1 repetitive element methylation, a surrogate of global DNA methylation. A methylation index was developed by averaging the percent methylation of AGTR1, FOXI2, and PENK promoters; patients with a high methylation index had a worse oral cancer–free survival ($P = 0.0030$). On the other hand, patients with low levels of LINE1 methylation had a significantly worse oral cancer–free survival ($P = 0.0153$). In conclusion, AGTR1, FOXI2, and PENK promoter methylation and LINE1 hypomethylation may be associated with an increased risk of OSCC development in patients with OPLs. Cancer Prev Res; 8(11); 1027–35. ©2015 AACR.

Introduction

Oral squamous cell carcinoma (OSCC) remains frequent worldwide with an estimated 263,000 cases each year causing about 127,000 deaths (1). The treatment of OSCC with surgery and radiation often results in substantial morbidity. A comprehensive molecular characterization of oral premalignant lesions (OPL) may allow refining oral cancer risk assessment and identifying new potential targets for chemoprevention (2).

Besides the identification of individual biomarkers of risk to develop OSCC (3–6), we have previously performed whole transcriptome profiling of biopsies of OPLs that were prospectively collected in a chemoprevention trial (2). We found that gene expression profiles were strongly associated with the development of OSCC and multiple transcripts identified in our study tend to be differentially expressed between normal mucosa and head and neck squamous cell carcinoma (HNSCC). DNA methyltransferase 3B (DNMT3B) is a de novo DNA methyltransferase and its transcript was one of the most significant ones associated with OSCC risk and was included in a 29-transcript model predictive of oral cancer development (2). Furthermore, polymorphisms of the DNMT3B gene have been associated with risk of developing HNSCC (7), and DNMT3B expression has been associated with the CpG island methylator phenotype (CIMP; refs. 8, 9), which is defined by high levels of simultaneous gene promoter methylation. CIMP was originally described in colorectal cancer but subsequently reported in other solid tumors, including HNSCC (10, 11). Both of those observations lead us to suggest that a CpG...
island methylation phenotype may be an early event during OSCC development. Studies evaluating increased promoter methylation from normal mucosa to premalignant lesions and to HNSCC have included one or a handful of number of genes, including retinoic acid receptor (RAR) β, p16<sub>ink4a</sub>, p14<sub>arf</sub>, endothelin receptor type B (EDNRB). E-cadherin (CDH1), deleted in lung and esophageal cancer 1 (DLEC1), NDRG family member 2 (NDRG2), and O-6-methylguanine-DNA methyltransferase (MGMT; refs. 12, 13). More recently, genome-wide methylation studies have shown that overall patterns of epigenetic alteration can more reliably distinguish tumor from normal head and neck epithelial tissues than individual gene methylation events (14, 15). Global DNA hypomethylation, measured by the degree of methylation of repetitive sequences across the genome such as long-interspersed nuclear element-1 (LINE-1), is another important feature of malignant tumors and has been associated with genome instability (16).

In the current study, we used high-throughput CpG island methylation profiles of OPLs for which clinical outcomes were available to determine the use of methylation profiles as predictive markers of OPLs progression and the potential importance of early DNA methylation changes to the development of OSCC. We found that (i) the top 86 genes differentially methylated between patients who did or did not develop OSCC were simultaneously hypermethylated in patients who develop OSCC, (ii) the vast majority of those 86 genes were nonmethylated in normal tissues and hypermethylated in HNSCC versus normal mucosa, (iii) among them, angiogenin B receptor type 1 (AGTR1), forkhead box I2 (FOXI2), and proenkephalin (PENK) promoters methylation was associated with OSCC development, and (iv) LINE-1 hypomethylation was also positively associated with OSCC development. We conclude that AGTR1, FOXI2, and PENK promoter methylation and global hypomethylation are promising markers of risk of OPLs progression.

**Materials and Methods**

**Patients and specimens**

From 1992 to 2001, 162 randomized and eligible patients were enrolled in a randomized OSCC chemoprevention trial at The University of Texas MD Anderson Cancer Center (MDACC; Houston, TX). The patients had been diagnosed with OPLs and randomized to intervention with 13-cis-retinoic acid (13-cisRA) versus retinyl palmitate (RP) with or without β-catenine. Detailed information has been previously described (17). A total of 153 frozen samples were available at baseline or 3 months after enrollment but before any event (defined as the diagnosis of OSCC). DNA was available for 148 cases, including 38 from patients who developed OSCC. All samples selected were collected at baseline. The study was approved by the institutional review board, and written informed consent was obtained from all patients. For the discovery part of the study, 24 samples were selected, including 12 samples from patients who did not develop OSCC (group A, median follow-up, 7.64 years) and 12 samples from patients who did develop OSCC (group B, median follow-up, 2.15 years).

**Identification of genes differentially methylated in OPLs from patients who developed OSCC**

DNA was extracted and purified from OCT-embedded tissue using the QIAamp DNA Mini Kit (Qiagen), including on-column RNase (Qiagen) digestion, as described by the manufacturer’s protocol. A hematoxylin and eosin (H&E)-stained section of all the samples was available to check for the presence of oral stratified epithelium. After DNA quantification using a ND-1000 spectrophotometer (Nanodrop Technologies), all DNAs were serially diluted in DNase-free water to obtain a 300 ng/μL stock solution. Aliquots were prepared and stored at −80°C. Each aliquot was used once.

High-throughput DNA methylation profiles were generated using a PCR-based method, methylated CpG island amplification (MCA) method described in detail elsewhere (18). All restriction enzymes were from New England Biolabs. In brief, 2 μg DNA was digested with Smal (methylation sensitive), followed by digestion with Xmal (methylation insensitive). The fragmented DNA was precipitated with ethanol and ligated with RXMA PCR adaptors RXMA24 (5’-AGCAGCTTCTACGCTGCCTACCGCA C-3’) and RXMA12 (5’-CCGGGCTCGGTCA-3’). The Xmal DNA fragments were ligated to RXMA adapters using T4 DNA ligase (New England Biolabs). PCR amplification was performed using a thermal cycler. Samples were subjected to 20 cycles of amplification. PCR products were then resolved on a 1.5% agarose gel and observed under ultraviolet light after ethidium bromide staining. Successfully methylated CpG island amplification reactions resulted in amplicons smear ranging from 300 bp to 3 kb, with most amplicons at 1 kb. PCR products were purified using the PCR purification kit from Qiagen.

Amplified CpG islands were then labeled with deoxyuridine and hybridized to comparative genomc hybridization (CGH) microarray slides. Incorporation of amino-allyl deoxyuridine triphosphate (dUTP; Sigma) into the DNA from potentially malignant lesions was conducted using the BioPrime Array CGH labeling system protocol (Life Technologies). Cy5 and Cy3 fluorescent dyes were coupled to amino-allyl dUTP-labeled oral leukoplakia (OL) from group A or group B, respectively. Each group A oral leukoplakia was randomly cohybridized with one group B oral leukoplakia to Agilent 4×44K custom CGH microarray slides, allowing the analysis and comparison of 12 pairs of oral leukoplakia. The array included 44,674 probes corresponding to 8,369 unique genes. The probes were selected to recognize Smal/Xmal restriction fragments from regions close to gene transcriptional start sites (19). Microarray protocols, including the hybridization and post-hybridization procedures, were performed as previously reported (20). Hybridized slides were scanned with the GenePix 4000A scanner (Axon Instruments), and the acquired images were analyzed with the software GenePix Pro 3.0. A two-step global lowess normalization was performed using the background-subtracted median intensity of each spot and the resultant Cy5/Cy3 log<sub>2</sub> (ratio) were obtained for the 12 pairs of OPLs. To identify genes differentially methylated in patients with OPLs who develop OSCC, we filtered out probes with no Smal fragment ID, on chromosomes X and Y, with no CpG island, and with an absolute minimal distance to transcription starting site > 1 kb. After filtering out those probes to identify genes differentially methylated in oral leukoplakia with malignant transformation, it remained 17,098 probes corresponding to 4,441 unique genes. We then selected probes with an average log<sub>2</sub> (ratio) ≥ 1 or ≤ −1 in at least 7 of the 12 pairs of oral leukoplakia studied to determine our candidate genes.

**In silico validation of candidate genes**

For in silico validation of candidate genes, we used the following 3 assumptions: (i) genes associated with OSCC development should be unmethylated in normal tissues, including normal
head and neck mucosa, but methylated in HNSCC; (ii) promoter methylation of genes associated with OSCC development may be downregulated in HNSCC when compared with normal mucosa; and (iii) genes associated with oral cancer development may also be methylated in lung cancer when compared with normal lung as oral cavity has been reported as a surrogate tissue for lung smoking-induced molecular alterations, that is, field effect, in particular for promoter methylation markers (21).

We tested those assumptions using the following datasets: 1–10, oral tongue SCC, blood, and normal mucosa from the same patients with available genome-wide DNA methylation profiles by Digital Restriction Enzyme Analysis of Methylation (DREAM), a method based on next-generation sequencing analysis of methylation-specific signatures created by sequential digestion of gDNA with Smal and Xmal enzymes, that has the ability to absolutely quantitate methylation of 13,000 unique genes (30,000 CpG sites) at transcription start sites individual genes (22); 2, publicly available DNA methylation and gene expression profiles of the following studies were downloaded from Gene Expression Omnibus ( GEO): GSE19434 (23), GSE25083 (24), GSE13601 (25), GSE27902 (26), GSE9844 (27), and GSE46802 (28). A detailed description of the samples and platforms used in those studies is provided in Supplementary Table S1. DNA methylation studies used the Illumina Methylation Assay and BeadChip technology (GSE19434, GSE25083, and GSE46802). Normalized data, presented as β values, and representing the degree of methylation at each CpG site (with 0 being unmethylated and 1 being fully methylated), were used to evaluate the degree of methylation of our candidate genes overlapping with genes analyzed on the Illumina platform. All probes were preserved irrespective of the number of probes per gene symbol. For the gene expression study (GSE13601, GSE6791, and GSE9844), raw microarray data were processed using quantile normalization and robust multi-array average algorithm. Using an Affymetrix platform, a unique gene can be represented by more than one probesets that were all preserved.

Validation of 5 candidate genes and of global DNA methylation in 37 patients with OPLs, using pyrosequencing

The validation set consisted of 37 OPLs collected in patients from the previously mentioned chemoprevention trial, including 23 from patients who did not develop OSCC and 14 from patients who did develop OSCC. Comprehensive data, including demographics, smoking and alcohol history, histology, OPLs anatomical site, treatment arm, and LOH at various sites (d9s171, d3s1285, tp53, d951747, d1751716, and d85254), were available. The degree of promoter methylation of angiotensin II receptor, type 1 (AGTR1), FOX12, homeobox A9 (HOXA9), proenkephalin (PENK), and Zic family member 1 (ZIC1) and the degree of methylation of the repetitive sequence Long Interspersed Elements (LINE1) were evaluated using pyrosequencing-based methylation analysis in the validation set of OPLs. All primers for pyrosequencing were designed using the Pyrosequencing Assay Design 2.0 software (Qiagen/Biotage). To validate our pyrosequencing assays, the degree of AGTR1, FOX12, HOXA9, PENK, and ZIC1 promoter methylation obtained by pyrosequencing in 61 head and neck cell lines described in Supplementary Table S2 was correlated with the results obtained in the same set of cell lines and the appropriate controls [HCT116-DKO (colon cancer cell line double knockdown for DNMT1 and DNMT3B), C42 (primary skin fibroblast cell line), HBECE16 (normal bronchial epithelial cell line), and HMVEC (primary dermal lymphatic endothelial cell line)] using the Illumina HumanMethylation27 Beadchip (Illumina, Inc.). This platform analyzes 27,579 CpG sites around promoters of 14,475 consensus coding sequences.

Bisulfite treatment of OPLs gDNA from the validation set was performed using the EpiTect Bisulfite Kit (Qiagen) according to manufacturer’s instructions. Bisulfite-treated DNA was used for each PCR. After an initial hotstart at 95°C for 5 minutes, all PCR reactions were run at 95°C for 30 seconds, annealed at various temperatures for 30 seconds, and subjected to an extension step at 72°C for 30 seconds and 50 cycles. All reactions were carried out with a nested PCR step during which a biotinylated universal primer [5’-GGGACACCGCTGATCGTT TA-3’] was added. After PCR, the biotinylated strand was captured on streptavidin-coated beads (Amersham Bioscience) and incubated with sequencing primers. Pyrosequencing was performed with PSQ HS 96 Gold reagents on a PS QHS 96 pyrosequencer (Biotage) as published previously (29). Methylation quantification was performed using the provided Pyro Q-CpG software. The program calculates for each single CpG dinucleotide the ratio between its methylated and nonmethylated forms, resulting in percentage of methylation. The methylation degree of each CpG island area was then determined by calculating the average of the methylation differences of all CpG sites analyzed in each gene region.

Statistical analysis

Data analysis was performed using GraphPad Prism version 5.00. Summary statistics, including median and range values, were used to describe the distribution of candidate genes in different datasets. Cancer and normal samples were compared using a log, fold change and a paired or nonpaired two-sided Student t test. Correlations were computed using the Spearman rank correlation coefficient. The Pearson correlation coefficient and Ward linkage method were used for hierarchical clustering of samples. The Kaplan–Meier method was used to construct oral cancer–free survival curves, and the log-rank test was used to test the difference by covariate levels. The univariate Cox proportional hazards model was fitted for all available variables, and multivariable Cox model analysis, including smoking status and significant variables in univariate analysis, was performed. We applied Firth correction in the multivariate analysis to obtain a reasonable estimate of HR for variable methylation index without adding any dummy observations (30). All statistical tests were two-sided, and P ≤ 0.05 was considered to be statistically significant.

Results

Identification of 86 candidate genes simultaneously methylated in patients developing OSCC

In each of the 12 pairs, the vast majority of the genes differentially methylated [log2(ratio) ≥ +1 or ≤ −1] were found to be hypermethylated [log2(ratio) ≥ +1] in OPLs from patients who subsequently went on to develop OSCC (Fig. 1). To select our candidate genes, we selected probes with a log2(ratio) ≥ +1 or ≤ −1 in at least 7 of the 12 pairs of OPLs (Supplementary Fig. S1). We identified 146 probes corresponding to 86 unique genes that were all hypermethylated in OPLs subsequently transforming to OSCC (Table 1). A total of 41 of 86 (48%) candidate genes were represented by at least 2 probes. Simultaneous methylation of these 86 genes in OPLs that subsequently underwent malignant
transformation is consistent with CIMP being an early event in oral tumorigenesis (11).

In silico validation of candidate genes

Our first assumption is that candidate genes should be unmethylated in normal tissues and frequently methylated in OSCC. Table 1 and Fig. 2A show the percent methylation of 32 of 86 (37%) candidate genes overlapping with those analyzed with DREAM in 10 oral tongue SCC, blood, and normal oral mucosa. The 32 genes queried were found to have very low levels of methylation in the blood (median, 0.83%; range, 0%–9.9%) and normal mucosa (median, 1.72%; range, 0.10%–11.46%) as opposed to the degree of methylation found in oral tongue SCC (median, 25.22%; range, 4.18%–63.97%). For further validation, we analyzed methylation data from a set of 146 blood samples, 18 head and neck normal mucosa, and 91 HNSCC that had been evaluated with the Illumina HumanMethylation27 platform which interrogates the promoter regions of 14,000 annotated genes, including 62 of 86 (72%) candidate genes (Supplementary Table S1; ref. 24). With few exceptions, genes were found to have lower methylation levels in the blood (median, 9%; range, 2%–92%) and normal head and neck mucosa (median, 14%; range, 0%–92%).

Table 1. Quantitative promoter methylation of 32 of 86 candidate genes in 10 OSCC and paired blood and normal tongue by DREAM (22)

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>% Methylation in OSCC</th>
<th>Number &gt;15%</th>
<th>Number measured</th>
<th>% Methylation #² blood</th>
<th>% Methylation #³ tongue</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZIC1</td>
<td>63.97</td>
<td>10</td>
<td>10</td>
<td>5.07</td>
<td>11.46</td>
</tr>
<tr>
<td>HOXA9</td>
<td>48.99</td>
<td>10</td>
<td>10</td>
<td>1.12</td>
<td>7.50</td>
</tr>
<tr>
<td>FOXI2</td>
<td>46.85</td>
<td>10</td>
<td>10</td>
<td>1.39</td>
<td>3.09</td>
</tr>
<tr>
<td>CLDN10</td>
<td>46.72</td>
<td>10</td>
<td>10</td>
<td>0.30</td>
<td>5.95</td>
</tr>
<tr>
<td>PENK</td>
<td>40.82</td>
<td>10</td>
<td>10</td>
<td>9.09</td>
<td>7.41</td>
</tr>
<tr>
<td>GAD2</td>
<td>38.97</td>
<td>9</td>
<td>10</td>
<td>0.96</td>
<td>2.31</td>
</tr>
<tr>
<td>V5X2</td>
<td>34.52</td>
<td>10</td>
<td>10</td>
<td>7.63</td>
<td>5.39</td>
</tr>
<tr>
<td>KCCN2</td>
<td>33.53</td>
<td>8</td>
<td>10</td>
<td>1.19</td>
<td>1.79</td>
</tr>
<tr>
<td>TMEM152D</td>
<td>33.24</td>
<td>10</td>
<td>10</td>
<td>2.86</td>
<td>4.16</td>
</tr>
<tr>
<td>SLC35F1</td>
<td>32.78</td>
<td>8</td>
<td>10</td>
<td>0.77</td>
<td>0.62</td>
</tr>
<tr>
<td>STRGALNACS5</td>
<td>31.06</td>
<td>7</td>
<td>9</td>
<td>0.45</td>
<td>5.64</td>
</tr>
<tr>
<td>CNTN4</td>
<td>28.98</td>
<td>7</td>
<td>10</td>
<td>2.00</td>
<td>3.38</td>
</tr>
<tr>
<td>NXPH1</td>
<td>28.29</td>
<td>8</td>
<td>10</td>
<td>0.73</td>
<td>1.13</td>
</tr>
<tr>
<td>NETO1</td>
<td>27.70</td>
<td>9</td>
<td>10</td>
<td>1.58</td>
<td>4.13</td>
</tr>
<tr>
<td>STAC2</td>
<td>26.23</td>
<td>7</td>
<td>10</td>
<td>0.85</td>
<td>0.86</td>
</tr>
<tr>
<td>KCTD1</td>
<td>26.13</td>
<td>8</td>
<td>10</td>
<td>3.20</td>
<td>4.59</td>
</tr>
<tr>
<td>SAL1</td>
<td>24.31</td>
<td>5</td>
<td>10</td>
<td>0.82</td>
<td>1.97</td>
</tr>
<tr>
<td>IFRB8</td>
<td>24.28</td>
<td>7</td>
<td>9</td>
<td>0.72</td>
<td>0.80</td>
</tr>
<tr>
<td>SNAP91</td>
<td>24.04</td>
<td>4</td>
<td>10</td>
<td>0.63</td>
<td>3.57</td>
</tr>
<tr>
<td>RYR2</td>
<td>22.43</td>
<td>7</td>
<td>10</td>
<td>0.65</td>
<td>1.12</td>
</tr>
<tr>
<td>NPY2R</td>
<td>22.33</td>
<td>5</td>
<td>8</td>
<td>1.34</td>
<td>1.24</td>
</tr>
<tr>
<td>GRIK5</td>
<td>22.05</td>
<td>4</td>
<td>10</td>
<td>1.33</td>
<td>1.35</td>
</tr>
<tr>
<td>HSST2</td>
<td>21.50</td>
<td>5</td>
<td>10</td>
<td>0.59</td>
<td>2.53</td>
</tr>
<tr>
<td>ZNF529</td>
<td>20.90</td>
<td>5</td>
<td>10</td>
<td>0.37</td>
<td>0.65</td>
</tr>
<tr>
<td>RGSL7</td>
<td>17.45</td>
<td>4</td>
<td>10</td>
<td>0.76</td>
<td>1.65</td>
</tr>
<tr>
<td>SOX14</td>
<td>14.10</td>
<td>4</td>
<td>10</td>
<td>1.30</td>
<td>1.10</td>
</tr>
<tr>
<td>ALDH1A2</td>
<td>12.53</td>
<td>3</td>
<td>7</td>
<td>0.00</td>
<td>0.33</td>
</tr>
<tr>
<td>WBSCR17</td>
<td>12.25</td>
<td>3</td>
<td>10</td>
<td>1.99</td>
<td>1.45</td>
</tr>
<tr>
<td>TMEFF2</td>
<td>11.03</td>
<td>1</td>
<td>10</td>
<td>0.49</td>
<td>0.60</td>
</tr>
<tr>
<td>ITGA4</td>
<td>9.18</td>
<td>2</td>
<td>10</td>
<td>0.33</td>
<td>0.33</td>
</tr>
<tr>
<td>PDEY0A</td>
<td>8.40</td>
<td>3</td>
<td>10</td>
<td>0.61</td>
<td>0.31</td>
</tr>
<tr>
<td>RND4C8</td>
<td>4.18</td>
<td>1</td>
<td>8</td>
<td>0.63</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Note: Four of the 5 genes tested in the validation cohort were analyzed on DREAM and are in bold.

Figure 1. Genes differentially methylated in OPL undergoing malignant transformation are simultaneously methylated. Distribution of the log$_2$ ratio of 17,098 probes (4,441 unique genes) across 12 pairs of OPLs. In each pair, one OPL from a patient who developed OSCC was labeled with Cy5 and cohybridized with one OPL from a patient who did not develop OSCC labeled with Cy3. Probes on chromosomes X and Y, with no SmaI fragment ID, with no CpG island, and with an absolute minimal distance to transcription starting site > 1 kb were removed.
1%–95%), compared with HNSCC (median, 31%; range, 4%–88%; Fig. 2B). Details for each gene are provided in Supplementary Fig. S2. Consistent results were obtained using another set of 217 normal tissues (including among others blood and head and neck mucosa) analyzed with an earlier platform that interrogates 1,505 CpG sites from the promoter region of 807 genes, including 13 of 86 (14%) candidate genes (ref. 23; Supplementary Table S1 and Supplementary Fig. S3). Finally, in a cohort of 10 patients with paired normal, dysplastic lesions and carcinomas (in situ or invasive) profiled with the Illumina HumanMethylation27 platform (28), the methylation status of 62 of 86 candidate genes allowed to identify 2 clusters of samples: one included 7 of 10 (70%) carcinomas and 3 of 10 (30%) dysplastic samples, whereas the other one included 10 of 10 normal mucosa samples, 7 of 10 (70%) dysplastic mucosa and 3 of 10 (30%) carcinomas (Supplementary Fig. S4).

Our second assumption was that the expression of our candidate genes would be downregulated in OSCC compared with normal oral mucosa through promoter hypermethylation. We used gene expression profiles of 32 OSCC and 26 normal oral mucosa, including 20 matched samples (Supplementary Table S1; ref. 25). A total of 62 probesets corresponding to 42 of 86 (49%) candidate genes were analyzed. An expression index was computed in 20 matched samples by the average expression levels of the 62 probesets. The expression index of our candidate genes was significantly lower in OSCC than in normal oral mucosa (Fig. 3A). Hierarchical clustering of the 58 samples using this geneset identified 2 clusters of samples, the first one included 22 of 26 normal oral mucosa and 7 of 32 oral tongue SCC and the second one including 4 of 26 normal oral mucosa and 25 of 32 oral tongue SCC (Fisher exact test; P < 0.0001; Supplementary Fig. S5A).

Our third assumption was based on the fact that oral cavity has been reported as a surrogate tissue for lung smoking–induced molecular alterations. Genes methylated early during oral tumorigenesis may therefore be methylated well in lung cancer when compared with normal lung. To test this, we computed a methylation index in 47 matched non–small cell lung cancer (NSCLC) and normal lung using 29 probes corresponding to 12 of 86 (14%) candidate genes overlapping with the platform used (GSE27902). Fig. 3B shows that the methylation index was significantly higher in NSCLC than in normal lung. Hierarchical clustering of the 94 samples showed 2 clusters of samples, the first one including 38 of 47 normal lung and 1 of 47 NSCLC and the second one including 9 of 47 normal lung and 46 of 47 NSCLC (Fisher exact test, P < 0.0001; Supplementary Fig. S5B).

Validation of 5 specific candidate genes and global DNA methylation changes in oral tumorigenesis

Among the 86 candidate genes methylated in OPLs that subsequently underwent malignant transformation, we selected AGTR1, FOXI2, HOXA9, PENK, and ZIC1 for further validation using pyrosequencing. Their average \( \beta \) values in 146 blood samples, 18 head and neck normal mucosa, and 91 HNSCC are...
shown in Fig. 4. To validate our pyrosequencing assays, we tested 61 head and neck cell lines and found a significant correlation between the degree of methylation at the promoter of AGTR1, FOXI2, HOXA9, PENK, and ZIC1 measured by pyrosequencing and their corresponding \( \beta \) values extracted from methylation profiles using the Illumina HumanMethylation27 Beadchip, thus validating our pyrosequencing assays (Supplementary Fig. S6).

The percent methylation of AGTR1, FOXI2, HOXA9, PENK, and ZIC1 was then evaluated in a validation set of 37 OPLs. Figure 5A–C shows that the percent methylation of AGTR1, FOXI2, and PENK in patients who developed OSCC (group B) was significantly higher than in patients who did not develop oral cancer (group A). The difference between group A and B was not as pronounced for ZIC1 and was not significant for HOXA9 (data not shown). We then computed a methylation index by averaging the percent methylation of AGTR1, FOXI2, and PENK. The group with high methylation index as defined by the median had a significantly worse oral cancer–free survival than the group with low methylation index (log-rank, \( P = 0.003 \); Fig. 5E).

We then tested whether the degree of methylation of LINE1, a group of genetic elements that composes 17% of the human genome and is being used as a surrogate of global DNA hypomethylation, is associated with OSCC development. Percentage methylation of AGTR1 (A), FOXI2 (B), PENK (C), and LINE1 (D) were obtained using pyrosequencing. ANOVA was used to compare both groups. Figure 4. Promoter methylation of AGTR1, FOXI2, HOXA9, PENK, and ZIC1 in HNSCC, normal mucosa, and blood. Promoter methylation was given by the average \( \beta \) values of AGTR1, FOXI2, HOXA9, PENK, and ZIC1 in a set of 146 blood samples, 18 head and neck normal mucosa, and 91 HNSCC (GSE25083; ref. 24). ANOVA was used to compare both groups.

Figure 5. Promoter methylation of AGTR1, FOXI2, and PENK and global DNA hypomethylation are associated with OSCC development. Percentage methylation of AGTR1 (A), FOXI2 (B), and PENK (C) were obtained using pyrosequencing from 37 OPLs, including 23 who did not undergo malignant transformation (group A) and 14 who did develop OSCC (group B). D, global DNA methylation was measured by the percentage methylation of LINE1 by pyrosequencing and compared in patients from groups A and B. E, Kaplan–Meier curve in patients with high versus low methylation index (MI) defined by the median of the average percentage methylation of AGTR1, FOXI2, and PENK. F, Kaplan–Meier curve in patients with high versus low LINE1 percentage methylation by pyrosequencing. Statistical significance was determined by unpaired two-sided Student t test (A–D) and log-rank test (E and F).
methylation, was associated with the development of OSCC. Using pyrosequencing, we measured the degree of methylation of LINE1 by pyrosequencing in 37 OPLs included in the validation set. LINE1 methylation was significantly lower in patients who developed OSCC (group B) than in patients who did not develop OSCC (group A; Fig. 5D). Using the median to define patients with high versus low LINE1 methylation, this translated to a worse oral cancer–free survival in patients with low LINE1 methylation (log-rank, \( P = 0.0153 \); Fig. 5F).

Histology (dysplasia vs. hyperplasia) was the only other variable reaching statistical significance in a univariate Cox proportional hazards model (Supplementary Table S3). Because of the small sample size, multivariate Cox models, including histology, smoking status, and methylation index or histology and LINE1 methylation status, were marginally significant and only LINE1 methylation was significant at the 5% level with \( P = 0.0368 \) (Supplementary Table S4).

**Discussion**

Promoter hypermethylation of specific genes have been already studied in oral premalignancy (31, 32), but description of early DNA methylation changes during oral carcinogenesis using genome-wide profiles is scarce. Towle and colleagues have reported early DNA promoter methylation alterations associated with histologic changes during oral tumorigenesis (28). They confirmed that those alterations are early events during oral tumorigenesis and offer an opportunity for biomarker development (31). In our study, we used an alternative approach and identified candidate promoter methylation markers based on patients’ outcome rather than histologic changes, because dysplastic lesions may progress to cancer, remain stable for years and/or may regress spontaneously or after tobacco/alcohol cessation, and OPLs without dysplasia may still progress to cancer (34). We also report the value of global DNA methylation to identify patients at high risk to develop OSCC.

We identified a set of 86 candidate genes with high levels of simultaneous gene promoter methylation associated with the development of OSCC in patients with OPLs. A majority were consistently found to be methylated in OSCC but not in normal tissues. Although our approach may have been biased by the use of a restriction enzyme inherent to the MCA method and limited to 8,369 unique genes analyzed on the Agilent 4×44k custom CGH microarray, we validated our findings in multiple independent datasets either found in the public domain and using a different array-based technology (GEO), as well as in samples profiled by DREAM, a method based on next-generation sequencing. We believe that consistent observations from independent studies using various platforms strengthen our conclusion.

Our observation of simultaneous gene promoter methylation in a large number of genes may be consistent with the emergence of a CIMP in the early steps of oral tumorigenesis. CIMP has been initially reported in colorectal cancer (11). Recent reports based on whole-genome methylation analysis support the concept of some tumors having a high propensity for CpG island DNA methylation (41). However, the absence of objective definition of CIMP may explain the challenge to determine its real clinical impact. In patients with breast cancer or glioblastoma, CIMP-positive tumors have been associated with an improved outcome (42), although in lung adenocarcinoma and colorectal cancer CIMP-positive tumors have been associated with worse outcome (43). The mechanism(s) associated with the onset of CIMP tumors remain(s) to be discovered.

In order to confirm the relevance of some of our 86 candidate genes, we focused on 5 specific ones: FOXI2, ZIC1, and HOXA9 because they were on the top of the list, PENK because its aberrant promoter methylation has already been reported in various cancers and it has recently been reported as a potential blood-based diagnostic markers in colorectal cancer (35). Finally, AGTR1 was prioritized over CLDN10 due to intriguing reports that are discussed below, linking the use of angiotensin receptor blockers (ARB) for the treatment of blood hypertension and the risk of cancer. We found that promoter methylation of FOXI2, PENK, and AGTR1 were associated with the development of OSCC. FOXI2 plays a role during development, in particular during early craniofacial development (36). PENK plays a role in cell death and survival and is downregulated by the proto-oncogenes FOS and JUN in the central nervous system. However, their role in cancer is not well understood (37).

Our finding of AGTR1 promoter methylation as a marker associated with OSCC development is consistent with a recent study of new DNA methylation markers for early diagnosis of colorectal cancer (35). This may also be consistent with provocative results of a recent meta-analysis of randomized controlled trials evaluating ARBs for the treatment of various medical conditions, suggesting that ARBs may be associated with a modestly increased risk of new cancer diagnosis. Among specific solid organ cancers examined, only new lung cancer occurrence was significantly higher in patients randomly assigned to receive ARBs than in those assigned to receive control (0.9% vs. 0.7%; ref. 38).

Together with our observation that AGTR1 promoter methylation is associated with OSCC development and is methylated as well in NSCLC, it is tempting to hypothesize that a loss of function of AGTR1 through pharmacologic intervention or DNA promoter methylation plays a role in lung and head and neck tumorigenesis. A complete understanding of the interaction between the two types of angiotensin receptors, ATR1 and ATR2, and functional studies in head and neck and lung preclinical models are warranted as those data contrast with other reports suggesting that AGTR1 may be a target for the treatment of gynecological malignancies (39).

Another aspect of DNA methylation aberrations in cancer is global DNA hypomethylation. In this report, we show that global DNA hypomethylation is an early event associated with the risk of developing OSCC in patients with OPLs. Global DNA hypomethylation has been associated with genome instability (16). Conflicting results have been reported between global DNA hypomethylation in peripheral blood and risk of cancer (40). The study of global DNA hypomethylation in the target tissue may be more relevant for cancer risk assessment.

The most robust and validated biomarker for risk prediction of OSCC development in patients with OPLs remains the LOH first reported by Mao and colleagues (5). The value of LOH profiles to predict OCSS development was validated in multiple retrospective studies and more recently in a prospective Canadian cohort (32). The value of AGTR1, FOXI2, and PENK promoter methylation and LINE1 hypomethylation to predict OSCC development and whether it can improve the risk assessment obtained by the study of LOH needs to be validated in larger cohorts of patients. One of the advantages of methylation over LOH markers is that analyzing constitutional DNA may not be necessary. On the other hand, LOH status can be defined in paraffin-embedded tissues,
which remains more difficult for DNA methylation markers that will require the development and evaluation of assays for tissues processed in every day practice.

In conclusion, AGTR1, FOXE1, and PENK promoter methylation may be associated with an increased risk of OSCC development in patients with OPLs. They may be associated with an early CIMP during oral tumorigenesis. LINE1 hypomethylation may also be associated with OSCC risk. Further studies are needed to validate these new biomarkers. The role of ARBs in head and neck and lung cancer warrants further investigation.

Disclosure of Potential Conflicts of Interest

V.A. Papadimitrakopoulou reports receiving commercial research grant from Merck & Company, Bayer, Novartis, Celgene, Astellas, AstraZeneca, Medimmune, Clovis Oncology, and Bristol-Myers Squibb and is a consultant/advisory board member for Amgen, Janssen Global Services, Clovis Oncology, and Bristol Myers Squibb. S.H. Lin reports receiving commercial research grant from STCube Pharmaceuticals Inc. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions


Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J.-P. Foy, C.R. Pickering, V. Papadimitrakopoulou, F.T. Mitchell, J. Wang, E.S. Kim, Y.H. Fan, J.N. Myers, S.M. Lippman, L. Mao, P. Saintigny


Other (provided materials for the study): A.K. El-Naggar

References

New DNA Methylation Markers and Global DNA Hypomethylation Are Associated with Oral Cancer Development


Access the most recent version of this article at:
doi:10.1158/1940-6207.CAPR-14-0179

This article cites 43 articles, 25 of which you can access for free at:
http://cancerpreventionresearch.aacrjournals.org/content/8/11/1027.full#ref-list-1

This article has been cited by 2 HighWire-hosted articles. Access the articles at:
http://cancerpreventionresearch.aacrjournals.org/content/8/11/1027.full#related-urls

Sign up to receive free email-alerts related to this article or journal.
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.