Global DNA Methylation: a Common Early Event in Oral Cancer Cases with Exposure to Environmental Carcinogens or Viral Agents

R. GUERRERO-PRESTON, DrPH, MPH*†; A. BÁEZ, Ph D‡; A. BLANCO, MD**; M. BERDASCO, Ph D§; M. FRAGA, Ph D§; M. ESTELLER, MD, Ph D§

Introduction: Two separate molecular pathways have been proposed for the early carcinogenic events observed in the oral cavity and pharynx: one is associated with chemical etiological factors such as smoking and drinking, and the other one is associated with HPV insertion.

Objective: A proof-of-principle study was performed to ascertain if global DNA methylation could be used to distinguish between the early molecular changes in premalignant oral lesions.

Methods: Personal histories of tobacco and alcohol use were obtained by questionnaire. HPV insertion in tumor tissue was detected by polymerase chain reaction (PCR). Global DNA methylation levels were obtained using HPLC for fraction separation and mass spectrometry for quantification. Predictive simulations were performed to explore potential associations between different etiological factors and the global DNA methylation index. Significance of results was ascertained using Pearson’s Chi-squared test.

Results: The global methylation index was found to be 4.28 (95% CI, 4.1, 4.4) in an oral cancer case series. Pearson’s chi squared test showed no statistically significant difference between cases that had smoking (p=0.21), drinking (p=0.31) or HPV insertion (p=0.34) as etiologic risk factors, when compared to cases that did not. An inverse significant association between smoking and DNA methylation was observed. As the smoking effect increases, the global methylation index decreases. In addition, no associations between the probability of DNA methylation and drinking, or DNA methylation and HPV insertion were observed in simulations.

Conclusions: The global DNA methylation index was shown to vary for oral cancer cases with different etiologies. Smoking was inversely correlated with DNA methylation levels when generalized linear model simulations were performed. Future studies should look at global DNA methylation alterations associated to the progression from normal to premalignant oral epithelium tissue in a cohort of smokers and non-smokers.

Key words: Epigenomics, DNA methylation, Early detection and cancer prevention, Oral cancer screening, Smoking-related molecular alterations

The primary environmental risk factors in approximately 80% of Head and Neck Squamous Cell Carcinomas (HNSCC) (oral, oropharyngeal and laryngeal) are tobacco smoking and the consumption of alcoholic beverages; and their joint effect appears to be multiplicative (1-2). Human papilloma virus (HPV) infection is now also recognized as another risk factor for HNSCC (3-6). Consumption of vegetables and fruit may modulate the carcinogenic effects of tobacco and alcohol, whereas low body-mass index increases the risk of oral cancer (7). The foods and drinks that people consume are not merely because of personal choice; likewise opportunities for physical activity can be constrained. Therefore, identifying the underlying factors that determine etiology in different populations may impact the burden of HNSCC. Each year, almost 650,000 patients worldwide are diagnosed with HNSCC and some 350,000 die from this disease (8). It is estimated that 24,180 men and 10,180 women in the United States (US) will be diagnosed with and 7,550 people will die in 2007 due to one of the most common types of HNSCC, oral cavity and pharyngeal carcinomas (OCP) (9).

An estimated 35% of overall cancer mortality has been attributed to only nine modifiable risk factors leading to cancer disparities (10). OCP disparities have been shown among the Latino population in the US. The age-standardized incidence per 100,000 people with OCP (excluding lip and nasopharynx), has been reported higher in Puerto Rican men (17.5) and women (4.5) than among...
Latino populations in the United States (men, 8.9; women, 2.7) (11). A recent analysis found that OCP carcinoma incidence and mortality rates among Latinos in New York is higher than among Latinos in the US, which may be due to ethnic and regional differences among Latinos in the US and Puerto Rico (12).

HNSCC arises by the accumulation of genetic and epigenetic changes in oncogenes, tumor suppressor genes, and/or DNA stability genes (13). Two separate molecular pathways, one driven by exposure to environmental carcinogens (modifiable life-style risk factors such as tobacco and alcohol use) without HPV involvement and the other only involving infection with oncogene-expressing HPV16, have been proposed to explain the early carcinogenic events observed in the oral cavity and pharynx. Subsequently, these two separate pathways converge into common late genetic and epigenetic events (14).

Smoking is associated at the early carcinogenic stage with allelic loss at 3p11, 5q11, 9p21, 17p13, 18q12, gain at 11q13, and amplification of CCND1 gene, loss of p16 and TP53 mutations (15-17). HPV16 initially drives carcinogenesis by inactivating p53 and pRb with the viral oncoproteins E6 and E7, while showing gain at 18q12 (14, 18). Global DNA methylation, an epigenetic marker of early carcinogenesis (19), could be useful as a cancer prevention and control tool if it can distinguish early molecular changes associated with the two carcinogenic routes proposed in oral cancer.

A proof-of-principle study was performed in a case-series obtained from the Tumor Biology Laboratory tissue bank of the University of Puerto Rico School of Medicine to ascertain if global DNA methylation, an epigenetic marker of early carcinogenesis, could be a useful tool in distinguishing early molecular changes associated with the two carcinogenic routes proposed in OCP. An early detection biomarker could be used in population based studies for the differential impact of environmental and lifestyle risk factors in populations with confounding ethnic, regional and environmental etiologic factors.

Methods

Tissue samples from fifteen oral cavity cancer cases were collected from surgical specimens of HNSCC tissue banked at the Tumor Biology Laboratory of the University of Puerto Rico School of Medicine for this proof-of-principle study. Personal histories of tobacco and alcohol use were ascertained by questionnaire. HPV infection was determined by detecting HPV DNA in tumor tissue by polymerase chain reaction (PCR). DNA was extracted using standard methods. Genomic DNA samples were boiled and treated with nuclease P1 and alkaline phosphatase. Global DNA methylation levels were obtained using HPLC for fraction separation and Mass Spectrometry for quantification. Fifty (50) µl of the hydrolyzed-DNA solution were injected onto a reversed phase dc18 column. Two buffers, 0.1% formic acid in water and 0.1% formic acid in 50% water /50% methanol, were used.

Identification of 2’-deoxyctydine (dC) and 5-methyl-2’-deoxycytidine (5mCdC) was done by UV detection at A254 and A280. Quantification of global DNA methylation was obtained from integration peak areas of 5mCdC relative to global cytidine (5mCdC + dC). The significance of associations between the methylation index and the predictor variables, age, gender, smoking, alcohol and HPV insertion was ascertained in Stata 9.0 with a bivariable Pearson’s Chi squared test. Predictive simulations were performed to explore associations between etiological factors and global DNA methylation. Generalized linear models were fitted, predictive simulations were implemented, and scatterplots were made in R 2.6. We can use the sim() function in R to create simulations that represent our uncertainty in the estimated regression coefficients. The following code was used to implement the utilization of the sim() function that is defined in the arm package of R:

```
R code

n.sims <- 1000
fit.1 <- lm (methylation index ~ hpv insertion + smoking + drinking)
sim.1 <- sim (fit.1, n.sims)
```

where sim.1$beta is a matrix with 1000 rows and 4 columns (representing 1000 independent simulations of the vector (β₀, β₁, β₂, β₃)). Three scatter plots draw the correlation between β₀, the y-intercept (methylation index), and the parameters of the three predictors : β₁ (hpv insertion positive); β₂ (smoking); and β₃ (drinking).

Results

The global methylation index, measuring methylated cytosine over total cytosine in the genome, was found to be 4.28 (95% CI, 4.1, 4.4) in an oral cancer case series. The Pearson’s chi squared test showed no statistically significant differences in the association between the global DNA methylation levels of cases that had smoking (p=0.21), drinking (p=0.31) or HPV insertion (p=0.34) as etiologic factors, when compared to cases that did not. An inverse association between smoking and DNA methylation was observed after 1,000 simulations of the glm linear model (y = α + β X). As the probability of
smoking increases, the probability of DNA methylation decreases (Figure 1). No associations were observed between the probability of DNA methylation and drinking or HPV insertion after 1,000 simulations.

**Discussion**

Tissue specific global methylation was shown for oral cancer cases with different etiologies, with a mean and standard deviation different from those previously found by us in liver cancer tissue using the same methodology. No difference in global DNA methylation levels between cases with different etiologies was observed, although smoking was correlated to DNA methylation levels when continuous predictive simulations utilizing a generalized linear model were performed. These preliminary in-vivo and in-silico results suggest that global DNA methylation may precede genetic alterations and molecular changes associated with exposure to viral and environmental carcinogens in HNSCC, as our conceptual model depicts (Figure 2). Many

![Figure 1](image-url)

**Figure 1.** Scatterplot of modeled simulations ($y = a + \beta X$) for HPV infection, smoking and drinking. The methylation index values (the intercept of the glm equation) are plotted on the x axis. Parameter estimates are plotted on the y-axis for each covariate: $\beta_1$ (hpv insertion positive); $\beta_2$ (smoking); and $\beta_3$ (drinking).
methylated cytosines have been found in retrotransposons, endogenous retroviruses and repetitive sequences, which may have evolved as a host defense mechanism to prevent the mobilization of these parasitic elements and reduce the occurrence of chromosomal rearrangements and the gain or loss of whole chromosomes (aneuploidy) (19-21).

Aneuploidy may be observed during chromosomal instability. DNA methylation has been associated with such instability. Loss of genomic integrity has been attributed to hypomethylation of repetitive elements, which can lead to inappropriate recombination resulting in defects in cell cycle monitoring check point genes as well as genes involved chromosome condensation, kinetochore structure and function, and centrosome and kinetochore formation (22). Chromosomal breakage and translocations in rare recessive genetic disorders are suggested to be due to mutations in the methyltransferase gene DNMT3b (23). Hypomethylation-induced translocations have been observed in multiple myeloma (24). DNA methylation seems to be a stabilizing agent in genomic structures comprising large amounts of repetitive elements by preventing recombination across these regions (19).

Global genomic DNA hypomethylation may precede and subsequently coexist with gene-specific promoter hypermethylation and hypomethylation in cancer (11, 25). Gene-specific hypomethylation has been associated with silencing of tumor suppressor genes (26-27). Gene specific hypomethylation has been associated with activation of oncogenes (28-29). The relationship between hypomethylation and hypermethylation in cancer is not well understood (30).

Global DNA hypomethylation may also be involved in other human diseases besides cancer working in concert with other epigenetic modifications (31). Histones post-translational modifications are also emerging as another important biological process in epigenetics to explain the behavior of cancer cells. These modifications are altered in tumor when compared to normal tissues. The loss of the monoacetylated Lys16 and trimethylated Lys20 residues of

---

**Figure 2.** Proposed mechanistic models of epigenetic/genetic alterations in oral and pharyngeal cancer according to etiology: chemical and viral. Smoking (top panel) is associated at the early carcinogenic stage with allelic loss at 3p11, 5q11, 9p21, 17p13, 18q12, gain at 11q13, and amplification of CCND1 gene, loss of p16 and TP53 mutations. HPV16 (bottom panel) initially drives carcinogenesis by inactivating p53 and pRb with the viral oncoproteins E6 and E7, while showing gain at 18q12.
histone H4 appears in the early phase of cell transformation and increases with tumor progression (32). These alterations may or may not be associated with promoter hypermethylation (33). The machinery that is responsible for the modification of chromatin and DNA work in a cooperative manner to silence genes in normal and malignant cells. The evidence of cross-talk between the histones and DNA methylation machineries makes this multi-protein complex a likely target for environmental carcinogens, by increasing the probability that genetic changes, when they occur, will lead to cancer initiation (13).

Global DNA hypomethylation has been associated with HNSCC diagnosis (34). Global DNA hypomethylation has also been found to be associated with smoking and alcohol use in HNSCC (34-35). Global DNA hypomethylation in non-target tissue (whole blood) has also been independently associated with HNSCC in a study that found smoking to have a significant differential effect on DNA methylation between cases and controls (3). In this paper, we have shown that global DNA methylation may also be an epigenetic marker in oral cancer. Global DNA methylation was not suggested as a useful marker to distinguish between the different signaling pathways in the early stages of oral cancer in this proof-of-principle study because it may be an early epigenetic event that precedes the differential carcinogenic alterations associated with exposure to HPV16, drinking or cigarette smoking in cancer of the oral cavity. The lack of association may also be due to a small sample size.

Screening high-risk populations for oral cancer in the primary care setting has been shown to be effective (36). Nevertheless, a systematic review of existing screening programs has not shown to be effective in impacting the burden of disease (37). A global DNA hypomethylation index, capturing loss of methylation at interspersed repeat sequences and genes, may well be a potential biomarker for the early detection of tumors and for prognostic use in monitoring disease progression (25). The sensitivity and specificity of this marker may be improved if it is combined with global histones H4 modification markers.

A surveillance program measuring global epigenetic biomarkers for OCP in saliva in high-risk populations in Puerto Rico can be utilized to predict future disease burden and establish preventive priorities. Reducing exposure to etiologic factors associated with high-risk behaviors in well designed preventive and health promotion initiatives may contribute to a reduction of existing cancer disparities as well as reducing future disease burden in Puerto Rico (38). Future studies should look at global epigenetic alterations associated to the progression from normal to premalignant tissue of oral cancer patients with different etiologies in a case control study.

References