NEOPLASIA
www.neoplasia.com

Volume 21 Number 2 February 2019 pp. 197–205 197

Genome-Wide Analysis of Head and Neck Squamous Cell Carcinomas Reveals HPV, *TP53*, Smoking and Alcohol-Related Allele-Based *Acquired Uniparental Disomy* Genomic Alterations[1]

Musaffe Tuna[*,†], **Christopher I. Amos**[†] **and Gordon B. Mills**[‡,§,¶]

[*]Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX; [†]Department of Medicine, Baylor college of Medicine, Houston, TX; [‡]Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, TX; [§]Department of Cell, Developmental & Cancer Biology, Oregon Health & Science University, Portland, OR; [¶]Precision Oncology, Knight Cancer Institute, Oregon Health & Science University, Portland, OR

Abstract

Smoking and alcohol intake are major risk factors in head and neck squamous cell carcinomas (HNSCCs). Although the link between *TP53* mutation and smoking has been well established, very little is known about the link between acquired uniparental disomy (aUPD) and smoking and/or alcohol consumption or other clinical characteristics. We used TCGA genomic data to investigate whether smoking, alcohol intake, clinical and demographic variables, HPV status and *TP53* mutation are associated with aUPD at specific chromosomal regions. In multivariate analysis, we found association between aUPD regions and risk factors and clinical variables of disease. aUPD regions on chromosome 4q, 5q, 9p, 9q, 13q, 17p and *CDKN2A* occurred significantly more often in patients with *TP53*-mutated HNSCC than in those with wild-type HNSCC, while aUPD regions on chromosome 9p and at *CDKN2A* were significantly more frequent in females than in males. Besides, aUPD occurred more frequent in HPV-positive than in HPV-negative samples with all HNSCC and larynx cancers on chromosome 9q 15q and 17p. Moreover, aUPD on *CDKN2A* region occurred more often in alcohol drinkers than nondrinkers in patients with all HNSCC and oral cavity cancers, while aUPD region on chromosome 5q occurred less in alcohol drinkers than nondrinkers in patients with all HNSCC and oral cavity cancers. Similarly, aUPD region on chromosome 5q occurred less in smokers than nonsmokers in patients with all HNSCC and oral cavity cancers. In conclusion, aUPD regions are not random, and certain regions are associated with risk factors for disease, and with *TP53* mutation status.

Neoplasia (2019) 21, 197–205

Introduction

Head and neck squamous cell carcinoma (HNSCC) is the seventh most common type of cancer and the second most common smoking-related cancer worldwide [1–3]. It presents in multiple sites, including the oral cavity, larynx, oropharynx, and hypopharynx. Oral cancers account almost one-third of HNSCCs [1]. Smoking is a major risk factor for HNSCCs (85%–90%), and alcohol is a risk factor for pharyngeal and laryngeal cancers. Thus, smoking either alone or combination with alcohol consumption increases the risk of HNSCCs.

Nonetheless, 10%–15% of HNSCCs are diagnosed in never-smokers and never-drinkers. Human papillomavirus (HPV) infection is another risk factor for HNSCCs, especially for oropharyngeal cancers (20%–72%); in fact, the incidence of oropharyngeal cancer in young men is higher among nonsmokers than among smokers in the United States [4–7]. Moreover, the incidence of oral cavity and oral tongue squamous cell carcinomas is increasing among white females aged 18 to 44 years [8].

Address all correspondence to: Musaffe Tuna, Department of Medicine, Baylor College of Medicine, Houston, TX.
E-mail: mtuna9@gmail.com
[1] Funding: This study is supported by The University of Texas MD Anderson's Cancer Center Support Grant (CCSG) NIH/NCI P30CA016672.
Received 4 November 2018; Revised 10 December 2018; Accepted 11 December 2018

1476-5586
https://doi.org/10.1016/j.neo.2018.12.002

TP53 mutations occur more frequently in smokers than in nonsmokers with HNSCC [9]; this mutation is also more common in HPV-negative than in HPV-positive patients with HNSCC [10]. However, it is not known whether smoking, alcohol intake, and HPV infections are linked to specific acquired uniparental disomy (aUPD) regions or to the frequency of aUPD in HNSCCs. aUPD, an allele-based alteration in the genome, represent commonly hidden alterations because in most cases, there is no change in the DNA copy numbers. It is not always as after one allele is lost, the remaining allele must be duplicated.

The concept of UPD was first introduced by Engel in 1980 as germline event [11]. aUPD can occur either as segmental or whole-chromosome aUPD. Germline and somatic UPD as well as whole-chromosome and segmental aUPD can arise through different mechanisms. A whole-chromosome can be lost in anaphase due to lagging and replicated in subsequent cell division resulting in whole-chromosome aUPD. The underlying mechanism for segmental aUPD has been proposed to occur as the consequence of mitotic recombination [12–15]. In addition, breakage-fusion-bridge (BFB) events in cancers can result in loss of chromosomal segments or whole-chromosomes that are frequently replicated in subsequent duplication [16]. Thus, BFB and subsequently replication of the retained chromosome or chromosome region is another mechanism underlying aUPD.

aUPD has been associated with the generation of homozygous aberrations that alter pathophysiology across cancer lineages affecting tumor aggressiveness, survival and therapeutic sensitivity. For example, aUPD encompassing the *BRCA1* region have been associated with shorter overall survival in ovarian cancers [17]. Similarly, in AML aUPD at chromosome 13q can result in homozygosity of mutations at *FLT3* and *JAK2* resulting in an aggressive phenotype [18]. *FLT3* and *JAK2* mutations increase ROS and HR capacity resulting in increased inter-chromosomal homologous recombination (iHR) events and extensive aUPD in AML [19]. Interestingly, treatment of *FLT3* and *JAK2* mutant AML cells with the antioxidant N-acetylcysteine decreased reactive oxygen species, and restored homologous recombination activity [19]. Of note, tumors with aUPD at chromosome 17p associated with mutant *TP53* showed increased RAD51 expression and HR activity resulting in increased resistance to traditional chemotherapy [20]. Association between aUPD at chromosome 4q and 9p resulting in homozygosity at *TET2* and *JAK2* genes, respectively, was found in myeloproliferative disorders, leading to increased oncogenic stress [21,22]. In fact, 33% of homozygous and 8% of heterozygous *TET2* mutant mice developed lethal myeloid malignancies in the first year of life [23]. Silencing of the wild type gene through aUPD regions can contribute to loss of tumor suppressors altering cellular transformation [24]. Studies, in particularly, comprehensive analysis of aUPD have been limited in HNSCC. In studies of a small number of tumors, recurrent aUPD at chromosome 3p21.31-p21.1 and 17p13.3-p13.1 regions were identified in HNSCCs [25], and at chromosome16p11.2 in oral cancers [26]. However, association between aUPD regions and clinical, and etiologic factors have not yet been presented.

Identification of aUPD regions can also pinpoint monoallelically expressed genes in the region and/or homozygous alterations in genes in the region that can affect tumor pathophysiology. Thus integrating aUPD with identification of mutations or genomic methylations can provide important mechanistic, etiologic and therapeutic information.

In this study, we investigated whether risk factors such as smoking, alcohol intake, HPV infection, and clinical factors are associated with the acquisition or relative frequency of specific aUPD regions in the genome in HNSCC.

Materials and Methods

Samples

In this study, we included HNSCCs from The Cancer Genome Atlas (TCGA) (http://tcga-portal.nci.nih.gov/tcga-portal, currently in https://portal.gdc.cancer.gov/). Clinical and patient characteristics were acquired from TCGA. HPV status is as reported by TCGA and by Nulton et al. [10,27]. HNSCCs are a heterogeneous group of cancers that include oral, oropharyngeal, larynx, and hypopharynx cancers. Cancers in the buccal mucosa, floor of mouth, hard palate, lip, oral cavity, oral tongue, and alveolar ridge are considered oral cavity cancers [10]. Cancers in the tonsil, soft palate, base of tongue and oropharynx are considered oropharyngeal. Analysis was performed in six groups separately: 1) all HNSCC samples, regardless of HPV status and organ site, 2) HPV-negative HNSCC, 3) all oral cavity cancers, regardless of HPV status, 4) HPV-negative oral cavity cancers, 5) all laryngeal cancers, and 6) HPV-negative laryngeal cancer samples. All clinical and demographic data are summarized in Table S1. Patients who were reported to be a smoker for 15 years or less and are not current smokers were considered *reformed smokers*. Patients who currently smoke were considered *current smokers*. Subjects who reported being lifelong nonsmokers were considered *nonsmokers*. Alcohol intake was described as *life-long nondrinker* when patients had <12 alcoholic beverages during their lifetime or >12 beverages during their lifetime but with a consumption per day equal to "0." The remaining subjects were considered *drinkers*.

Genomic Data and Analysis

We used SNP genotyping data from TCGA. QC was performed by using Genotyping Console software (GCOS), and CHP files were generated with the same software. In this study, only data that passed QC were included (n = 448). aUPD (Figure S1) analysis was performed by using Copy Number Analyzer for GeneChip (CNAG) v3.4 (http://www.genome.umin.jp) as described earlier [17]. The smallest overlapping regions of aUPD were defined by comparing aUPD endpoints (3′ and 5′) as described before [17]. The Dec 2013 human genome browser (NCBI Build 38/hg38; http://genome.ucsc.edu) was used for identification of gene localization. If only one breakpoint was occurred, this was described as *telomeric aUPD*. When at least two breakpoints occurred, this was described as *centromeric aUPD*. Telomeric and centromeric aUPD together was considered *segmental*. If aUPD presented as a whole-chromosome, this was considered *whole-chromosome aUPD*; if aUPD presented in only one arm of a chromosome, this was considered to be either *p* or *q arm aUPD*. *Total aUPD* was a summary of segmental and whole-chromosome.

Statistical Analyses

A Pearson's chi-squared test and multivariate logistic regression model were used for association between aUPD regions and stage, grade, age, gender, HPV status, smoking, alcohol intake, and *TP53* mutation status. A nonparametric Kruskal-Wallis test and multivariate linear regression analysis were performed to identify the association between total, telomeric, centromeric, segmental, and

whole-chromosome; p and q arms of chromosome aUPD; and smoking, alcohol intake, and *TP53* mutation status. The *TP53* mutation was selected because it is the most frequently mutated gene in HNSCCs. Statistical analyses were performed with use of STATA v10.0 software (Stata Corp., College Station, TX).

Results

In this study, we investigated whether the frequency of aUPD (total, telomeric, centromeric, segmental, whole, p and q arms of chromosome) was associated with smoking, alcohol intake, and *TP53* mutation status in all HNSCC, only HPV-negative HNSCC, all oral cavity, only HPV-negative oral cavity, all laryngeal, and only HPV-negative laryngeal cancer samples. We excluded only HPV-positive HNSCC, oropharyngeal, and hypopharyngeal cases due to small sample size.

We found that the frequency of aUPD was associated with smoking status in laryngeal cancers. The frequencies of total aUPD (P = .0008), telomeric (P = .0006), centromeric (P = .0363), segmental (P = .0007), p arm (P = .0370), and q arm of chromosomes (P = .0006) were significantly higher, except whole-chromosome aUPD (P = .4433) in reformed smokers than in current smokers in laryngeal cancer samples (Table S2 and Figure S2). Indeed, the frequency of aUPD (total, telomeric, centromeric, segmental, and q arm of chromosome) was significantly higher in reformed smokers (P = .0079, P = .0085, P = .0434, P = .0064, and P = .0071, respectively) than in current smokers in HPV-negative patients with laryngeal cancers; no significant associations were found between whole (P = .5434) and p arm (P = .0722) of chromosomes, and smoking. These results indicated that smoking may be a more prevalent risk factor in laryngeal cancers for aUPD than in oral cavity cancers. Indeed, 91.2% of patients with laryngeal cancers were smokers, either current or reformed, whereas only 69.26% of patients with oral cavity cancers were smokers in this cohort. In contrast, the frequency of aUPD (total, telomeric, centromeric, segmental, whole-chromosome, p and q arms of chromosomes) was not associated with smoking status in all HNSCC, HPV-negative HNSCC, all oral cavity, and only HPV-negative oral cavity cancer samples (Table S2 and Figure S2). In multivariate analysis, no association was found between frequency of aUPD and smoking in patients with HNSCC, oral cavity and larynx cancers (Table S2 and Figure S2).

Moreover, we found that the frequency of centromeric aUPD was significantly higher in alcohol drinkers than in nondrinkers in all HNSCCs (P = .0440), in all oral cavity cancers (P = .0120), and in only HPV-negative oral cavity cancers (P = .0381), but no difference was found in all laryngeal (P = .8418) or HPV-negative laryngeal (P = .9702) cancers. In contrast, the frequency of whole-chromosome aUPD was significantly higher in nondrinkers than in drinkers in all HNSCC (P = .0090), in HPV-negative HNSCC (P = .0011), in all oral cavity (P = .0010), and in HPV-negative oral cavity cancers (P =.0004), whereas no difference was found between drinkers and nondrinkers in all laryngeal (P = .7467) and HPV-negative laryngeal cancer (P = .6832) samples (Table S2 and Figure S3). In multivariate analysis, only frequency of whole-chromosome was associated with alcohol intake in patients with HPV-negative HNSCC (P = .018), with all oral cavity (P = .016), and with HPV-negative oral cavity cancers (P = .024) (Table S2).

In analysis by gender, we found that the frequency of aUPD in males was higher than in females in all and in HPV-negative HNSCC, in all and in HPV-negative oral cavity cancers, but not in laryngeal cancers (Figure 1). The frequencies of centromeric (P = .0007), segmental (P = .0386), and q arm (P = .0013) aUPD were significantly higher in males than in females, but total (P = .1279), telomeric (P = .4399), whole-chromosome (P = .1199), and p arm (P = .7477) aUPD frequency did not differ by gender in all HNSCC (Figure 1). The frequencies of centromeric (P = .0127) and q arm (P = .0108) aUPD were significantly higher in male than in female patients in all oral cavity. In multivariate analysis also association was found between gender and frequency of total aUPD (P = .039), centromeric (P = .002), segmental (P = .29), and q arm (P = .001) in patients with all HNSCC, and frequency of q arm aUPD (P = .015) in all oral cavity cancers (Table S2).

Next, we determined whether the frequency of aUPD was associated with *TP53* mutation status. We found a significant association between the frequency of aUPD and *TP53* mutation in all groups. The frequency of total aUPD (P = .0001), telomeric (P = .0001), centromeric (P = .0001), segmental (P = .0001), whole (P = .0019), p arm (P = .0001), and q arm (P = .0001) of chromosomes was significantly higher in *TP53*-mutated samples than in wild types in all HNSCC samples (Figure 2). Moreover, the prevalence of total aUPD (P = .0001), telomeric (P = .0001), centromeric (P = .0001), segmental (P = .0001), whole (P = .0240), p arm (P = .0002), and q arm (P = .0001) of chromosomes was significantly higher in *TP53*-mutated HPV-negative HNSCC than in wild types (Figure 2). Similarly, in all laryngeal cancer samples, the frequency of total aUPD (P = .0004), telomeric (P = .0067), centromeric (P = .0048), segmental (P = .0012), whole-chromosome (P = .0195), p arm (P = .0115), and q arm (P = .0012) of chromosomes was significantly more common in *TP53*-mutated than in non-*TP53* mutated cases (Figure 2 and Table S2). Moreover, the frequencies of total aUPD (P = .0015), telomeric (P = .0103), centromeric (P = .0054), segmental (P = .0016), p arm (P = .0051), and q arm (P = .0121) were significantly higher in *TP53*-mutated samples than in wild types in HPV-negative laryngeal cancer samples, except in whole-chromosome aUPD (P = .0943) (Figure 2). In contrast, the frequencies of only total aUPD (P = .0173), centromeric (P = .0014), and q arm (P = .0194) of chromosomes was significantly higher in *TP53*-mutated than in wild types in patients with HPV-negative oral cavity cancers. In all oral cavity cancers, the frequencies of total aUPD (P = .0034), centromeric (P = .0001), segmental (P = .0132), whole-chromosome (P = .0293) and q arm of chromosomes (P = .0053) was significantly higher in *TP53*-mutated than in wild types (Figure 2 and Table S2).

Next we tested for differences in the frequency of aUPD among patient samples by age. The frequency of p arm aUPD was significantly higher in all HNSCC (P = .0403) and in HPV-negative HNSCC (P = .0478) patients with age over 50 than in under and equal to 50. However, the frequency of aUPD was not associated with age in oral cavity and larynx samples (Table S2). Multivariate analysis also revealed that frequency of aUPD significantly associated with *TP53* mutations, gender and age in patients with HNSCC, while frequency of aUPD associated with alcohol intake, *TP53* mutation and gender in oral cavity cancers, and with *TP53* mutation in larynx cancers (Table S2).

Association of Smallest Overlapping Regions (SOR) of aUPD with Clinical Characteristics

We then identified 34 of the SORs of aUPD including *CDKN2A*. We further investigated whether stage, grade, gender, age, HPV status, smoking, alcohol intake, and *TP53* mutations were associated

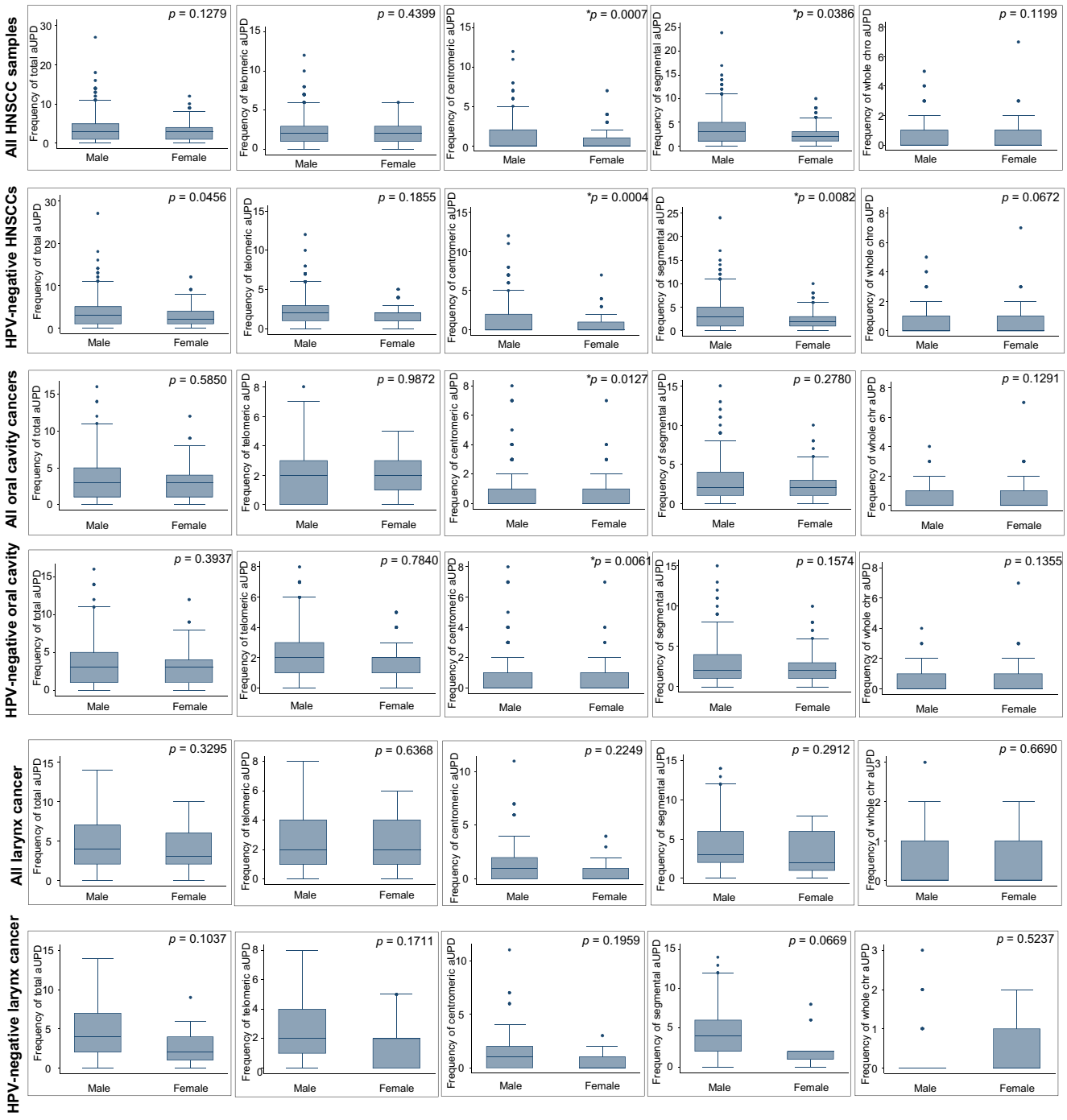

Figure 1. Frequency of total, telomeric, centromeric, segmental, and whole-chromosome aUPD by gender in all HNSCC, HPV-negative HNSCC, all oral cavity cancers, and HPV-negative oral cavity cancers.

with the SORs of aUPD in the whole genome. In addition, we tested deletion of the *CDKN2A* region for association with clinical characteristics.

Association Between Smoking and aUPD Regions

We found that aUPD at the 17p13.3 region (P = .032 and P = .018) and deletion at *CDKN2A* (P = .039 and P = .039) were significantly more common in current smokers than in never smokers and reformed smokers in all (Table S3) and in HPV-negative HNSCCs (Table S4). When we tested only patients with all oral cavity cancers, three independent regions at chr 9q22.33 (P = .020), chr 9q33.2 (P = .042), and 9q34.13 (P = .009) and one at chr 17p13.3 (P = .050) were significantly more common in current smokers than in never and reformed smokers (Table S5). Similarly, one aUPD region at chr 9p21.1 (P = .032) and two at chr 9q (9q22.33; P = .025 and 9q34.13; P = .005) were more frequent in current smokers than in never and reformed smokers in HPV-negative oral cavity cancers (Table S6). In contrast, aUPD at chr 5q11.2 region was more frequent in never smokers than in smokers (current and reformed smokers) in all (P = .042) (Table S3) and in HPV-negative (P = .020) HNSCCs (Table S4). Moreover, in patients with laryngeal cancer, three independent aUPD regions

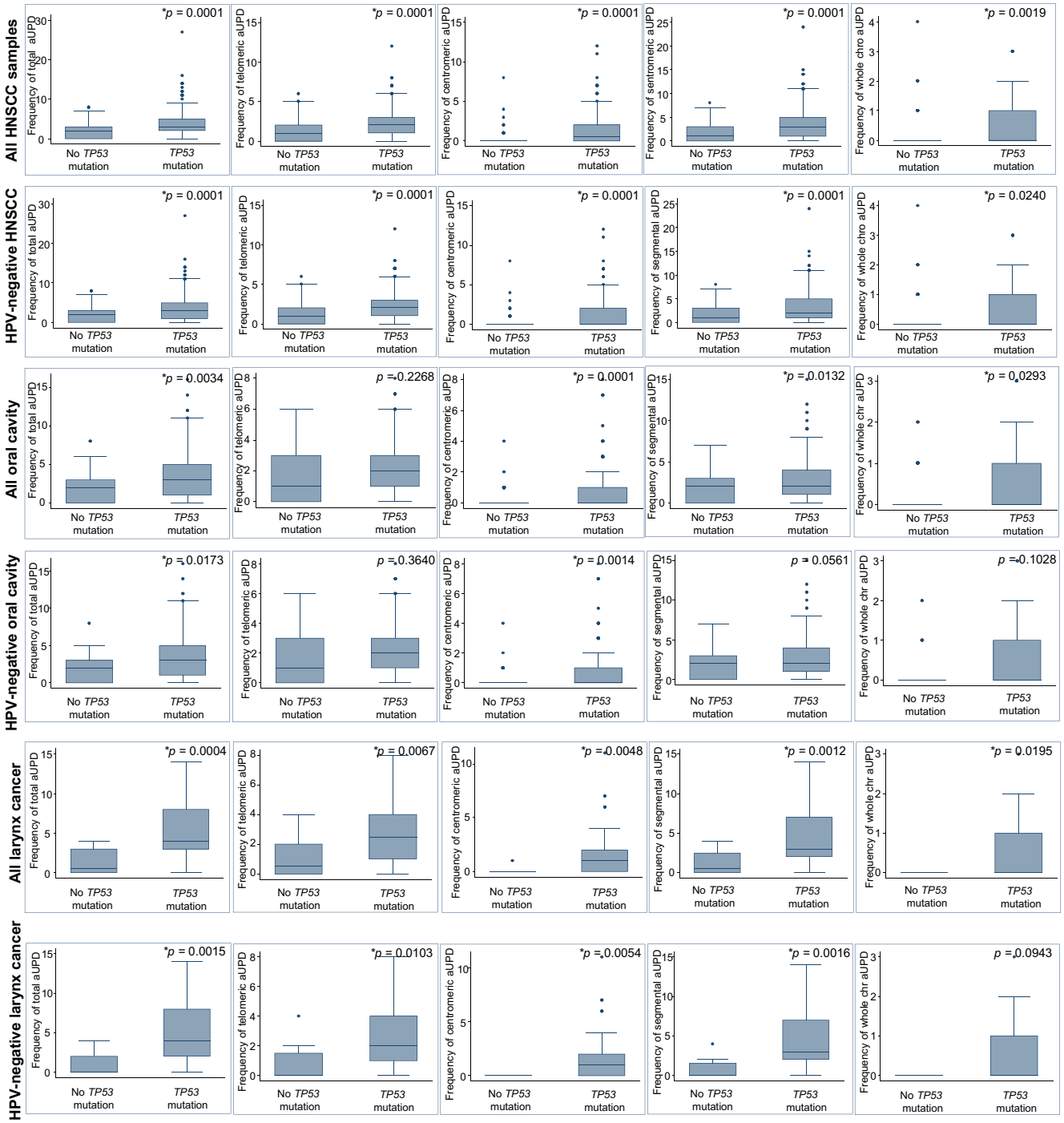

Figure 2. Frequency of total, telomeric, centromeric, segmental, and whole-chromosome aUPD by *TP53* status in all HNSCC, HPV-negative HNSCC, all oral cavity cancers, and HPV-negative oral cavity cancers.

were significantly more common in never smokers than in current and reformed smokers: regions at chr 4q32.1–32.3 (P = .004), at chr 5q11.2 (P = .014), and at chr 5q23.3-q31.1 (P = .029) (Table S7). Similarly, in HPV-negative samples from patients with laryngeal cancer, two aUPD regions at chr 5q11.2 (P = .0050) and 5q23.3-q31.1 (P = .0100) were more frequent in never smokers than in smokers, and the region at chr 13q22.1 was more frequent in never and reformed smokers (P = .0370) than in current smokers (Table S8). In multivariate analysis also, aUPD region at chr 5q23.3-q31.1 was more frequent in never smokers than current and reformed smokers in all and HPV-negative HNSCC, and in all oral cavity cancers (Tables S3–S6).

Association Between aUPD Regions and Alcohol Intake

Next we determined whether any of these SORs of aUPD were associated with alcohol intake. We found that the aUPD region at chr 5q23.3-q31.1 was more common in nondrinkers than in drinkers in all HNSCC (P = .042) (Table S3), in HPV-negative HNSCC (P = .026) (Table S4), in all oral cavity cancer (P = .024) (Table S5), and in HPV-negative oral cavity cancer (P = .026) (Table S6) samples. No

association was found between any SORs of aUPD and alcohol intake in larynx cancers (Tables S7 and S8). Similar result was found in multivariate analysis; aUPD region at chr 5q23.3-q31.1 was more common in nondrinkers than in drinkers in all and HPV-negative HNSCC, and in patients with all and HPV-negative oral cavity cancers (Tables S3–S6). In contrast, aUPD was significantly more common on *CDKN2A* region in drinkers than nondrinkers in all HNSCC (OR = 2.85, $P = .010$, 95%CI = 1.28–6.32), in HPV-negative HNSCC (OR = 3.31, $P = .007$, 95%CI = 1.38–7.93), in all oral cavity cancers (OR = 3.36, $P = .015$, 95%CI = 1.27–8.87), and in patients with HPV-negative oral cavity cancers (OR = 3.63, $P = .014$, 95%CI = 1.29–10.19) (Table S3–S6).

Association Between aUPD Regions and Gender

Next, we determined whether any of the SORs of aUPD that we identified were associated with gender. Of interest, even though frequency of aUPD was higher in males (mean: 3.78 in males, 2.96 in female), nine SORs of aUPD at chr 9p (9p24.3; $P < .0001$, 9p24.1; $P < .0001$, 9p23-p22.3; $P = .007$, 9p22.3-p22.2; $P = .005$, 9p21.3_1; $P = .0050$, 9p21.3_2; $P = .009$, 9p21.3-p21.2; $P = .023$, 9p21.7; $P = .044$, and 9p13.3; $P = .043$) and at *CDKN2A* ($P = .004$) were significantly more common in females than in males in all HNSCC samples (Table S3). Moreover, two aUPD regions at chr 6p22.1-p21.33 ($P = .039$) and at chr 6p12.3 ($P = .038$), seven regions at chr 9p (9p24.3; $P = .001$, 9p24.1; $P = .001$, 9p23-p22.3; $P = .018$, 9p22.3-p22.2; $P = .015$, 9p21.3_1; $P = .018$, 9p21.3_2; $P = .022$, and 9p21.3-p21.2; $P = .042$), and *CDKN2A* ($P = .007$) were significantly higher in female than in male patients with oral cavity cancers (Table S5). Similarly, five aUPD regions at chr 9p (9p24.3; $P = .003$, 9p24.1; $P = .003$, 9p23-p22.3; $P = .050$, 9p22.3-p22.2; $P = .040$, and 9p21.3_1; $P = .050$), and one at *CDKN2A* ($P = .022$) were significantly higher in female than in male patients with HPV-negative oral cavity cancers (Table S6). In contrast, no association was found between aUPD regions and gender in all and HPV-negative laryngeal samples (Tables S7 and S8). In multivariate analysis, ten aUPD regions at chr 9p and *CDKN2A* associated with gender in all and HPV-negative HNSCC, and in all and HPV-negative oral cavity cancers (Tables S3–S6).

Association Between aUPD and TP53 Mutation

We then determined whether any of the SORs were associated with *TP53* mutations. We found that regions at chr 4q32.1–32.3 ($P = .047$), 5q23.3-q31.1 ($P = .034$), 6p12.3 ($P = .025$), 9p13.3 ($P = .021$), 9q22.33 ($P < .0001$), 9q31.3 ($P < .0001$), 9q33.2 ($P < .0001$), 9q34.13 ($P < .0001$), *CDKN2A* ($P < .0001$), 11q25 ($P = .039$), 13q22.1 ($P = .022$), 15q26.3 ($P = .009$), 17p13.3 ($P = .001$), and 17p12 ($P = .005$) and deletion on *CDKN2A* regions were significantly more frequent in *TP53*-mutated samples than in wild types in all HNSCC samples (Table S3). Similar results were observed in HPV-negative HNSCCs (Table S4). Moreover, in patient with oral cavity cancers, aUPD regions at chr 9q22.33 ($P = .009$), chr 9q31.3 ($P = .012$), chr 9q33.2 ($P = .007$), chr 9q34.13 ($P = .005$), and chr 13q22.1 ($P = .020$) and deletion on *CDKN2A* ($P < .0001$) were significantly more common in *TP53*-mutated than in wild types (Table S5). Similarly, in patients with HPV-negative oral cavity cancers, aUPD regions at chr 9q22.33 ($P = .021$), 9q31.3 ($P = .028$), 9q33.2 ($P = .022$), 9q34.13 ($P = .015$), 13q22.1 ($P = .037$), and in deletion on *CDKN2A* ($P < .0001$) were significantly more frequent in *TP53*-mutated than in wild types (Table S6). In contrast, in all and HPV-negative laryngeal cancers, only deletion on *CDKN2A* region ($P = .01$ and $P = .047$) was more common in *TP53*-mutated samples than in wild types (Tables S7 and S8). In multivariate analysis, 11 independent aUPD regions [(chr 4q32.1–32.3, chr 5q23.3-q31.1, chr 9p21.1, 9p13.3, four at chr 9q (9q22.33, 9q31.3, 9q33.2, 9q34.13), chr 13q22.1, two at chr 17p (17p13.3 and 17p12), and *CDKN2A*)], and deletion on *CDKN2A* region were associated with *TP53* mutation status in all HNSCC, while only one aUPD regions at 9p21.3, four at chr 9q (9q22.33, 9q31.3, 9q33.2 and 9q34.13), and aUPD and deletion on *CDKN2A*, and one at 13q22.1 in oral cavity cancers, and three independent aUPD regions at chr 9q (q31.3, 9q33.2 and 9q34.13), and one at 13q22.1, and deletion on *CDKN2A* were associated with *TP53* mutation in HPV-negative oral cavity cancers (Tables S3, S5 and S6).

Association Between aUPD Regions and HPV Status, Disease Stage, Grade, and Age

Next we found aUPD regions at chr 9q22.33 ($P = .034$) and 9q33.2 ($P = .032$), 15q26.3 ($P = .031$) in HNSCC (Table S3) and at chr 17p13.3 ($P = .036$) in oral cavity cancers (Table S5) were more common in HPV-positive than in HPV-negative cases. Similarly, in patients with laryngeal cancer, aUPD regions at chr 9q (9q22.33; $P = .025$, 9q31.3; $P = .035$, 9q33.2; $P = .025$, and 9q34.13; $P = .035$, respectively), chr 15q26.3 ($P = .027$), and chr 17p13.3 ($P = .026$) were more frequent in HPV-positive than in HPV-negative cases (Table S7). In multivariate analysis, we also found five aUPD regions (four at chr 9q and one at 15q) in patients with HNSCC, and six aUPD regions (four at chr 9q, one at chr 15q, and one at chr 17p) in patients with larynx cancers were more common in HPV-positive than in HPV-negative cases (Tables S3 and S7).

When we tested the association between aUPD regions and stage in all HNSCC samples, we found aUPD at chr 9q regions (9q22.33; $P = .046$, 9q31.3; $P = .043$, 9q33.2; $P = .043$, 9q34.13; $P = .032$), were more prevalent in stages III and IV than in stages I and II (Table S3). These data suggest that aUPD in those regions may occur in later stages of disease and that genes in these regions may be involved in disease progression. In multivariate analysis, no association was found between stage and aUPD regions (Tables S3-S8).

When we tested the association between aUPD regions and grade, we found that aUPD region at chr 3p24.3 was more common in grade 1 than in grades 2, 3, and 4 in all HNSCC samples ($P = .003$), in HPV-negative HNSCCs ($P = .041$) (Table S3), in all oral cavity cancers ($P = .032$) (Table S5), and in all larynx cancers ($P = .004$) (Table S7). These data indicate that aUPD at chr 3p24.3 may occur in lower grades of disease. In multivariate analysis, 13 independent aUPD regions including *CDKN2A* were associated with grade in all HNSCC, and four aUPD regions in all oral cavity cancers (Tables S3 and S5).

When we tested the association between aUPD regions and age, we found that aUPD region at chr 17p12 was more common in age older than 50 than in equal or under 50 in all larynx ($P = .016$) (Table S7) and HPV-negative larynx cancers ($P = .037$) (Table S8), when aUPD regions at chr 8p23.2 ($P = .030$ and $P = .036$), 8p23.1 ($P = .030$ and $P = .036$) were more common in younger age. In multivariate analysis, only aUPD region on *CDKN2A* region was significantly frequent in age older than 50 than in equal or under 50 in all and HPV-negative HNSCC, and oral cavity cancers (Tables S3-S6).

We integrated aUPD with TCGA mutation and methylation data from the most frequently altered genes *TP53* and *CDKN2A*. aUPD at *TP53* region was identified in 131 samples, and in 127 out 129 aUPD positive samples where sequencing data was available, *TP53* was

mutated. Thus only two samples were wild type for *TP53*. However both of the samples that were wild type for *TP53* carried mutations in other genes in the region, in one sample *DVL1* gene was mutated, and in the other *DNAH2* was mutated. Thus aUPD provides a mechanism to derive uniformly mutated p53 protein in the majority of samples with aUPD associated with *TP53*. *CDKN2A* is more complicated than *TP53*. aUPD at the *CDKN2A* region was found in 96 samples. *CDKN2A* was mutated in 28 of 96 samples, methylated in 61 samples, homozygously deleted due to aUPD in 27 samples, and exon 1α was lost in 11 samples. Some of the samples demonstrated more than one alteration. Twenty-one samples of the 96 with sUPD in the *CDKN2A* region did not demonstrate mutation/methylation or homozygous deletion at *CDKN2A*. There remains a possibility that undetected mutation, methylation, or imprinting of *CDKN2A* is associated with the aUPD. Alternatively, the aUPD at the *CDKN2A* region may be associated unknown alterations in other genes in the region. Nevertheless, aUPD appears to contribute to complete loss of *CDKN2A* in a number of HNSCC tumors.

Discussion

Previously, the *TP53* mutation (86%) and *CDKN2A* alterations (90%), including mutation, deletion, methylation, and abnormally spliced transcripts, were shown to occur significantly more frequently in HPV-negative HNSCCs [10]. This finding indicated the coexistence and association of two inactivated tumor suppressor genes with HPV-negative HNSCC. In this study, we identified 34 SORs of aUPD and tested whether the frequency of aUPD and any of the SORs was associated with clinical variables, smoking and alcohol intake, HPV, and *TP53* mutation status. As expected, we found that the deletion on *CDKN2A* was also more frequent in *TP53*-mutated cases in all groups. This result was consistent with an earlier report [28]. Furthermore, aUPD on *CDKN2A* was significantly more common with *TP53* mutation in all and HPV-negative HNSCC and in all oral cavity cancers. However, there was no difference between *TP53*-mutated and wild types in all and HPV-negative laryngeal cases. Our results indicate that aUPD on *CDKN2A* is associated with *TP53* mutation in all HNSCC and oral cavity cancers but not in laryngeal cases.

Moreover, *TP53* mutation has been linked to smoking with 3.5 times more *TP53* mutations have reported in smokers than in nonsmokers with HNSCCs [29,30]. However, it is still not well known whether smoking also increases the occurrence of allele-based alterations. We conducted this study to determine whether smoking is linked to allele-based alterations in specific regions of the genome. We found that aUPD at chr 9q regions were significantly higher in current smokers than in nonsmokers or reformed smokers in all and HPV-negative oral cavity cancers, but not in laryngeal cancers. Of more interest, the same regions were also significantly higher in *TP53*-mutated samples in all and HPV-negative oral cavity cancers. In contrast, no association was found between aUPD at 9q regions and *TP53* mutation status and smoking in laryngeal cancer patients. On the other hand, *TP53* mutation was higher in all and HPV-negative HNSCC, but no association was found with smoking. These data indicate that three aUPD regions at chromosome 9q link to smoking status in oral cavity cancer patients and that expression of genes in these regions may be deregulated by smoking. Moreover, these data indicated the coexistence of aUPD at 9q regions and of the *TP53* mutation in HNSCC and oral cavity cancers, but not in laryngeal cancers. Thus smoking may trigger mitotic recombination either secondary to smoking associated *TP53* mutation or via other mechanisms in HNSCC and in oral cavity cancers. In addition, it may be a more prevalent risk factor for oral cavity cancers than for laryngeal cancer. Our results support a previous report indicating that relative risk of death due to cancer among smokers compared with nonsmokers was 27.5 for oral cavity whereas it was 10.5 for laryngeal cancer [9]. Interestingly, the same aUPD regions at chromosome 9q were associated with HPV in all HNSCC and larynx cancers but not in oral cavity cancers. These data indicate that HPV may increase the mitotic recombination in these regions in larynx cancer.

In addition, our analysis revealed that the frequency of centromeric aUPD was more common in alcohol drinkers than in nondrinkers in all HNSCC and oral cavity cancers, whereas whole-chromosome aUPD was more frequent in nondrinkers than in drinkers. aUPD regions at chr 5q and *CDKN2A*, and deletion on *CDKN2A* were more prevalent in drinkers than nondrinkers. These data indicate that genes in these regions may relate with alcohol drinking, and alcohol intake may increase mitotic recombination in HNSCC and oral cavity cancers. Besides, aUPD on *CDKN2A* region was associated with multiple factors including age, gender, grade, alcohol intake and *TP53* mutation in all HNSCC and oral cavity cancers, while it was not associated with any of factors in larynx cancers. Thus, the other factors may also increase aUPD in HNSCC and oral cavity cancers. Of interest, while the frequency of centromeric and segmental aUPD was higher in males, but aUPD at chr 9p regions was more common in females in all HNSCC and oral cavity cancers, but not in larynx cancers. Moreover, aUPD at chr 9q regions was significantly higher in stages III and IV than in stages I and II. aUPD at chr 9q regions was also higher in *TP53*-mutated cases than in wild types in all HNSCCs. This may due to 76.4% of *TP53*-mutated samples being stage III or IV. These data indicated that aUPD at chr 9q in HNSCCs may be associated with *TP53* mutation rather than with stage. Overall, *TP53* mutation may be involved in mitotic recombination in HNSCC cells.

Previous studies have demonstrated that alcohol intake and smoking have different impacts on copy number alterations; alcohol was associated with copy number changes in 3p, 3q, 9p, 11q, 17q, but not in 17p, while smoking was associated with mutations in *TP53* [9,31]. Alterations such as deletion, mutation and methylation at *CDKN2A* were correlated with smoking and HPV status in HNSCC [32]. Moreover, smoking deregulates PIWI-interacting RNAs (piRNAs) [33] and miRNAs [34] in HNSCCs. Alcohol intake [35], and smoking was associated with distinct DNA-methylation changes [10,36–40]. Of note, the methylation profile is also different between HPV-positive and negative groups in HNSCC [10]. Taken together, the data indicates that smoking, alcohol and HPV associated HNSCC may have different underlying etiologies that could lead to different therapeutic opportunities or needs for these patients. Our results support previous reports that different risk factors have distinct effect on DNA copy number changes and/or methylation by showing that they also alter aUPD. Homozygosity for an existing alteration mediated by aUPD could result in a more aggressive phenotype, decreased survival or resistance to therapy. Moreover, aUPD regions associated with specific risk factors may harbor genes that represent therapeutic opportunities. Thus, integrating aUPD with DNA sequencing and methylation provides an opportunity to better understand etiology, pathophysiology and potentially therapeutic options needed for HNSCC patients.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neo.2018.12.002.

Disclosure

G. B. Mills has sponsored research support from Pfizer, Abbvie, Critical Outcomes Technology, Immunomet, Karus Therapeutics, Nanostring, Millennium, Tesaro, and AstraZeneca; has ownership interests in Catena Pharmaceuticals, ImmunoMet, SignalChem, Spindle Top Ventures and Tarveda; has licensed technology to Nuevolution, and Myriad, and is a consultant/advisory board member of AstraZeneca, Critical Outcome Technologies, Immuno-Met, Ionis, Nuevolution, PDXbio, Signalchem Lifesciences, Symphogen, and Tarveda.

Acknowledgements

The authors thank Tamara K. Locke of Department of Scientific Publication (The University of Texas, MD Anderson Cancer Center) for editing the manuscript.

References

[1] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, and Bray F (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* **136**, E359-386.

[2] Jamal A, King BA, Neff LJ, Whitmill J, Babb SD, and Graffunder CM (2016). Current cigarette smoking among adults – United States, 2005-2015. *MMWR Morb Mortal Wkly Rep* **65**, 1205–1211.

[3] Siegel RL, Miller KD, and Jemal A (2015). Cancer statistics, 2015. *CA Cancer J Clin* **65**, 5–29.

[4] Forman D, de Martel C, Lacey CJ, Soerjomataram I, Lortet-Tieulent J, Bruni L, Vignat J, Ferlay J, Bray F, and Plummer M, et al (2012). Global burden of human papillomavirus and related diseases. *Vaccine* **30**(Suppl. 5), F12–F23.

[5] Sturgis EM and Ang KK (2011). The epidemic of HPV-associated oropharyngeal cancer is here: is it time to change our treatment paradigms? *J Natl Compr Canc Netw* **9**, 665–673.

[6] Gillison ML (2004). Human papillomavirus-associated head and neck cancer is a distinct epidemiologic, clinical, and molecular entity. *Semin Oncol* **31**, 744–754.

[7] Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, Jiang B, Goodman MT, Sibug-Saber M, and Cozen W, et al (2011). Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol* **29**, 4294–4301.

[8] Patel SC, Carpenter WR, Tyree S, Couch ME, Weissler M, Hackman T, Hayes DN, Shores C, and Chera BS (2011). Increasing incidence of oral tongue squamous cell carcinoma in young white women, age 18 to 44 years. *J Clin Oncol* **29**, 1488–1494.

[9] Brennan JA, Boyle JO, Koch WM, Goodman SN, Hruban RH, Eby YJ, Couch MJ, Forastiere AA, and Sidransky D (1995). Association between cigarette smoking and mutation of the p53 gene in squamous-cell carcinoma of the head and neck. *N Engl J Med* **332**, 712–717.

[10] TCGA Network (2015). Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, 576–582.

[11] Engel E (1980). A new genetic concept: uniparental disomy and its potential effect, isodisomy. *Am J Med Genet* **6**, 137–143.

[12] Stephens K, Weaver M, Leppig KA, Maruyama K, Emanuel PD, Le Beau MM, and Shannon KM (2006). Interstitial uniparental isodisomy at clustered breakpoint intervals is a frequent mechanism of NF1 inactivation in myeloid malignancies. *Blood* **108**, 1684–1689.

[13] Raghavan M, Lillington DM, Skoulakis S, Debernardi S, Chaplin T, Foot NJ, Lister TA, and Young BD (2005). Genome-wide single nucleotide polymorphism analysis reveals frequent partial uniparental disomy due to somatic recombination in acute myeloid leukemias. *Cancer Res* **65**, 375–378.

[14] Reliene R, Bishop AJ, and Schiestl RH (2007). Involvement of homologous recombination in carcinogenesis. *Adv Genet* **58**, 67–87.

[15] Blackburn AC, McLary SC, Naeem R, Luszcz J, Stockton DW, Donehower LA, Mohammed M, Mailhes JB, Soferr T, and Naber SP, et al (2004). Loss of heterozygosity occurs via mitotic recombination in Trp53+/– mice and associates with mammary tumor susceptibility of the BALB/c strain. *Cancer Res* **64**, 5140–5147.

[16] Alcaraz Silva B, Jones TJ, and Murnane JP (2017). Differences in the recruitment of DNA repair proteins at subtelomeric and interstitial I-SceI endonuclease-induced DNA double-strand breaks. *DNA Repair* **49**, 1–8.

[17] Tuna M, Ju Z, Smid M, Amos CI, and Mills GB (2015). Prognostic relevance of acquired uniparental disomy in serous ovarian cancer. *Mol Cancer* **14**, 29.

[18] Tiu RV, Gondek LP, O'Keefe CL, Huh J, Sekeres MA, Elson P, McDevitt MA, Wang XF, Levis MJ, and Karp JE, et al (2009). New lesions detected by single nucleotide polymorphism array-based chromosomal analysis have important clinical impact in acute myeloid leukemia. *J Clin Oncol* **27**, 5219–5226.

[19] Gaymes TJ, Mohamedali A, Eiliazadeh AL, Darling D, and Mufti GJ (2017). FLT3 and JAK2 mutations in acute myeloid leukemia promote interchromosomal homologous recombination and the potential for copy neutral loss of heterozygosity. *Cancer Res* **77**, 1697–1708.

[20] Klein HL (2008). The consequences of Rad51 overexpression for normal and tumor cells. *DNA Repair* **7**, 686–693.

[21] Yamamoto G, Nannya Y, Kato M, Sanada M, Levine RL, Kawamata N, Hangaishi A, Kurokawa M, Chiba S, and Gilliland DG, et al (2007). Highly sensitive method for genomewide detection of allelic composition in nonpaired, primary tumor specimens by use of affymetrix single-nucleotide-polymorphism genotyping microarrays. *Am J Human Genet* **81**, 114–126.

[22] Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, An J, Lamperti ED, Koh KP, and Ganetzky R, et al (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839–843.

[23] Ahn JS, Kim HJ, Kim YK, Jung SH, Yang DH, Lee JJ, Lee IK, Kim NY, Minden MD, and Jung CW, et al (2015). Adverse prognostic effect of homozygous TET2 mutation on the relapse risk of acute myeloid leukemia in patients of normal karyotype. *Haematologica* **100**, e351–e353.

[24] Martin-Trujillo A, Vidal E, Monteagudo-Sa Nchez A, Sanchez-Delgado M, Moran S, Hernandez Mora JR, Heyn H, Guitart M, Esteller M, and Monk D (2017). Copy number rather than epigenetic alterations are the major dictator of imprinted methylation in tumors. *Nat Commun* **8**, 467.

[25] Marescalco MS, Capizzi C, Condorelli DF, and Barresi V (2014). Genome-wide analysis of recurrent copy-number alterations and copy-neutral loss of heterozygosity in head and neck squamous cell carcinoma. *J Oral Pathol Med* **43**, 20–27.

[26] Morita T, Uzawa N, Mogushi K, Sumino J, Michikawa C, Takahashi KI, Myo K, Izumo T, and Harada K (2016). Characterizing genetic transitions of copy number alterations and allelic imbalances in oral tongue carcinoma metastasis. *Genes Chromosomes Cancer* **55**, 975–986.

[27] Nulton TJ, Olex AL, Dozmorov M, Morgan IM, and Windle B (2017). Analysis of The Cancer Genome Atlas sequencing data reveals novel properties of the human papillomavirus 16 genome in head and neck squamous cell carcinoma. *Oncotarget* **8**, 17684–17699.

[28] TCGA Network (2017). Integrated genomic and molecular characterization of cervical cancer. *Nature* **543**, 378–384.

[29] Harris CC and Hollstein M (1993). Clinical implications of the p53 tumor-suppressor gene. *N Engl J Med* **329**, 1318–1327.

[30] Lesmes GR and Donofrio KH (1992). Passive smoking: the medical and economic issues. *Am J Med* **93**, 38S–42S.

[31] Urashima M, Hama T, Suda T, Suzuki Y, Ikegami M, Sakanashi C, Akutsu T, Amagaya S, Horiuchi K, and Imai Y, et al (2013). Distinct effects of alcohol consumption and smoking on genetic alterations in head and neck carcinoma. *PLoS One* **8**,e80828.

[32] Ghosh A, Ghosh S, Maiti GP, Sabbir MG, Alam N, Sikdar N, Roy B, Roychoudhury S, and Panda CK (2009). SH3GL2 and CDKN2A/2B loci are independently altered in early dysplastic lesions of head and neck: correlation with HPV infection and tobacco habit. *J Pathol* **217**, 408–419.

[33] Krishnan AR, Korrapati A, Zou AE, Qu Y, Wang XQ, Califano JA, Wang-Rodriguez J, Lippman SM, Hovell MF, and Ongkeko WM (2017). Smoking status regulates a novel panel of PIWI-interacting RNAs in head and neck squamous cell carcinoma. *Oral Oncol* **65**, 68–75.

[34] Krishnan AR, Zheng H, Kwok JG, Qu Y, Zou AE, Korrapati A, Li PX, Califano JA, Hovell MF, and Wang-Rodriguez J, et al (2017). A comprehensive study of smoking-specific microRNA alterations in head and neck squamous cell carcinoma. *Oral Oncol* **72**, 56–64.

[35] Cervera-Juanes R, Wilhelm LJ, Park B, Grant KA, and Ferguson B (2017). Alcohol-dose-dependent DNA methylation and expression in the nucleus accumbens identifies coordinated regulation of synaptic genes. *Transl Psychiatry* **7**, e994.

[36] Guida F, Sandanger TM, Castagne R, Campanella G, Polidoro S, Palli D, Krogh V, Tumino R, Sacerdote C, and Panico S, et al (2015). Dynamics of smoking-induced genome-wide methylation changes with time since smoking cessation. *Hum Mol Genet* **24**, 2349–2359.

[37] Hasegawa M, Nelson HH, Peters E, Ringstrom E, Posner M, and Kelsey KT (2002). Patterns of gene promoter methylation in squamous cell cancer of the head and neck. *Oncogene* **21**, 4231–4236.

[38] Jamebozorgi I, Majidizadeh T, Pouryagoub G, and Mahjoubi F (2018). Aberrant DNA Methylation of Two Tumor Suppressor Genes, p14(ARF) and p15 (INK4b), after Chronic Occupational Exposure to Low Level of Benzene. *Int J Occup Environ Med* **9**, 145–151.

[39] Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, Guan W, Xu T, Elks CE, and Aslibekyan S, et al (2016). Epigenetic signatures of cigarette smoking, circulation. *Cardiovasc Genet* **9**, 436–447.

[40] Zeilinger S, Kuhnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, Weidinger S, Lattka E, Adamski J, and Peters A, et al (2013). Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One* **8**, e63812.