Stratification of HPV-associated and HPV-negative oropharyngeal squamous cell carcinomas based on DNA methylation epigenotypes

Takuya Nakagawa¹,², Keisuke Matsusaka²,³, Kiyoshi Misawa⁴, Satoshi Ota⁵, Masaki Fukuyo⁷, Bahiytar Rahmutulla², Naoki Kunii³, Daiju Sakurai¹, Toyoyuki Hanazawa¹, Hisahiro Matsubara⁵, Yoshitaka Okamoto¹ and Atsushi Kaneda²

¹Department of Otolaryngology, Head and Neck Surgery, School of Medicine, Chiba University, Chiba, Japan
²Department of Molecular Oncology, School of Medicine, Chiba University, Chiba, Japan
³Department of Pathology, Chiba University Hospital, Chiba, Japan
⁴Department of Otologyngology/Head and Neck Surgery, Hamamatsu University School of Medicine, Hamamatsu, Japan
⁵Department of Otorhinolaryngology, Head and Neck Surgery, School of Medicine, Chiba University, Chiba, Japan
⁶Department of Frontier Surgery, Graduate School of Medicine, Chiba University, Chiba, Japan

While the incidence of oropharyngeal squamous cell carcinoma (OPSCC) has been increasing in these two decades, primarily due to human papillomavirus (HPV), stratification of OPSCC into molecular subgroups showing different clinicopathological features has not been fully investigated. We performed DNA methylome analysis using Infinium 450k for 170 OPSCC cases, including 89 cases in our cohort and 81 cases reported by The Cancer Genome Atlas, together with targeted exon sequencing analysis. We stratified OPSCC by hierarchical clustering analysis using methylome data. Methylation levels of classifier markers were validated quantitatively using pyrosequencing, and area under the curve (AUC) values of receiver operating characteristics (ROC) curves were calculated. OPSCC was stratified into four epigenotypes: HPV(+) high-methylation (OP1), HPV(+) intermediate-methylation (OP2), HPV(−) intermediate-methylation (OP3) and HPV(−) low-methylation (OP4). Ten methylation marker genes were generated: five to classify HPV(+) cases into OP1 and OP2, and five to classify HPV(−) cases into OP3 and OP4. AUC values of ROC curves were 0.969 and 0.952 for the two marker panels, respectively. While significantly higher TP53 mutation and CCND1 copy number gains were observed in HPV(−) than in HPV(+) groups (p < 0.01), no significant difference of genomic aberrations was observed between OP1 and OP2, or OP3 and OP4. The four epigenotypes showed significantly different prognosis (p = 0.0006), distinguishing the most favorable OPSCC subgroup (OP1) among generally favorable HPV(+) cases, and the most unfavorable OPSCC subgroup (OP3) among generally unfavorable HPV(−) cases. HPV(+) and HPV(−) OPSCC are further divided into distinct DNA methylation epigenotypes, showing significantly different prognosis.

Additional Supporting Information may be found in the online version of this article.

Key words: DNA methylation, human papillomavirus, head and neck squamous cell cancer

Abbreviations: AUC: area under the curve; COSMIC: Catalogue of Somatic Mutations in Cancer; GO: gene ontology; HME: high methylation epigenotype; HNSCC: head and neck squamous cell cancer; HPV: human papillomavirus; IHC: immunohistochemistry; LME: intermediate methylation epigenotype; Indels: insertions and deletions; LME: low methylation epigenotype; OPSCC: oropharyngeal squamous cell carcinoma; OS: overall survival; PBC: peripheral blood cells; PRC: polycomb repressive complex; ROC: receiver operating characteristics; SNVs: single nucleotide variants; SRA: sequence read archive; TCGA: The Cancer Genome Atlas; TSS: transcription start site

DOI: 10.1002/ijc.32890

History: Received 12 Nov 2019; Accepted 15 Jan 2020; Online 30 Jan 2020

Correspondence to: Atsushi Kaneda, E-mail: kaneda@chiba-u.jp

Introduction

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common cancer worldwide, with an annual incidence of 600,000 cases, accounting for 3–5% of all malignancies. Despite significant advancement in treatment strategy, HNSCC remains lethal with a 5-year survival rate of 40–50%.

HNSCC has shown a paradigmatic change in characteristics in the last two decades due to human papillomavirus (HPV) infection. Although the incidence of cervical cancer is decreasing due to improvement in screening systems and vaccination, the incidence of oropharyngeal squamous cell carcinoma (OPSCC) has been on the rise over the last two decades, primarily due to the dramatic increase in the number of HPV-associated OPSCC cases and the non-availability of screening instruments for HPV-associated OPSCC, unlike cervical cancer.

HPV-associated OPSCC has distinct clinical and biological characteristics compared to HPV-negative cancer. A meta-analysis confirmed that patients with HPV-associated OPSCC have a 28% lower risk of death (meta HR 0.72 [95% CI 0.5–1]).
Infection with human papilloma virus (HPV) generally lowers risk of death for patients with oropharyngeal squamous cell carcinoma but 20% of patients do not “benefit” from HPV infection in this sense. Here the authors used DNA methylation analysis to further subdivide HPV+ and HPV- cancer patients, successfully recapitulating the distinct prognosis. They also generated classifier markers to accurately distinguish between subtypes based on epigenome analysis, but not genomic mutations, thus potentially aiding with clinical treatment decisions in the future.
samples were obtained at Chiba University, and 37 OPSCC samples were obtained at Hamamatsu University. Written informed consent was obtained from all patients. Noncancerous mucosa samples were obtained from OPSCC patients via biopsy from the side opposite to the cancerous side. These biopsy specimens were immediately frozen in liquid nitrogen and stored at −80°C. The cancer cell content of the frozen materials was determined to be ≤10, 20, 30, 40, 50, 60, 70, 80 or ≥90% through microscopic examination by two independent pathologists, and was enriched by dissection when necessary. Among 96 samples, 89 samples, of which the cancer cell content was determined ≥50% by both the two pathologists, were used for subsequent molecular analyses. The remaining seven samples were excluded because two contained no cancer cell, three were determined as ≤50% cancer cell content by both pathologists, and two were determined as ≤50% by one pathologist. DNA was extracted using the QIAquick DNA mini kit (Qiagen, Hilden, Germany). DNA of normal peripheral blood cells (PBC) was purchased from the Coriell Cell Repositories. The protocol was approved by the institutional review board at Chiba University and Hamamatsu University.

**Immunohistochemistry**

IHC for p53 was performed using a DO-7 anti-mouse monoclonal antibody (anti-p53 DO-7 primary antibody, Roche, Basel, Switzerland) and the BenchMark ULTRA automated staining system (Roche). Samples with diffuse strong positive nuclear staining were considered p53-IHC(+) and designated as TP53-mutation(+), while those with negative nuclear staining were considered p53-IHC(−) and also designated as TP53-mutation(+). Those with sporadic positive nuclear staining were considered p53-IHC wild-type and designated as TP53-mutation(−). For p16INK4A, IHC was performed using an antimouse monoclonal antibody (CINtec p16 Histology, Roche) and BenchMark ULTRA, and samples with positive nuclear staining (nuclear and cytoplasmic immunoreactivity in >75% of tumor cells) were considered p16-IHC(+) and designated as p16(+).

**Amplification of L1 DNA region in high-risk HPV**

HPV infection status was evaluated by amplifying a portion of the L1 region in high-risk HPV using GP5 and GP6 primers, as previously reported.\(^{28}\) Samples with the PCR amplicon were designated as HPV-L1(+).

**Infinium assays**

The Infinium HumanMethylation450 BeadChip (Illumina) contains approximately 485,000 individual CpG sites covering 99% RefSeq genes with an average of 17 CpG sites per gene. For each CpG site, β-value, ranging from 0.00 to 1.00, was measured using a methylated probe relative to the sum of both methylated and unmethylated probes. Bisulfite conversion was performed using the Zymo EZ DNA methylation kit (Zymo Research, Irvine, CA) with 500 ng of genomic DNA for each sample. Whole genome amplification, labeling, hybridization and scanning were performed according to the manufacturer’s protocols. Infinium analysis was performed against 89 clinical OPSCC and five noncancerous mucosa samples.

For analyzing DNA methylation at promoter regions based on Infinium data, the probe nearest to the transcription start site (TSS) was selected when multiple probes were designed for one promoter region. The CpG score for each probe was calculated based on a previous report.\(^{29,30}\) For analysis of CpG island shore regions, N_Shore and S_Shore, the probe nearest to TSS was selected among probes at the CpG island shore (annotated by Illumina). For analysis of gene body, the probe showing the highest CpG score was selected among probes at gene body, that is, +4 kb downstream from the TSS through the transcription termination site based on a previous report.\(^{31}\)

**TCGA samples and their HPV status**

Independent Infinium datasets of 81 OPSCC samples supplied by TCGA (https://gdac.broadinstitute.org/) were used for additional and combined analysis of DNA methylation status. Their HPV status was determined by either p16-IHC or in situ hybridization test, and adapted to our study.

**Targeted exon-sequencing**

Targeted exon-sequencing against 86 clinical OPSCC and seven noncancerous mucosa samples was performed using GeneRead DNAseq Targeted Panels V2 and the Human Comprehensive Cancer Panel (Qiagen), which targets 160 candidate driver genes, suits short amplicon design and enables detection of low-level mutations, reducing the amount of DNA required. For each of the somatic mutations identified using this analysis, variants with an allele frequency <0.1 and >0.9 were excluded. The known variants previously reported in the dbSNP (http://www.ncbi.nlm.nih.gov/SNP) were filtered out, and synonymous mutations were also excluded. Subsequently, we used MutationTaster classification tools that can predict the functional consequences of amino acid changes or frameshift mutation, and considered the observed mutations significant when MutationTaster evaluated “disease-causing automatically.” Finally, we excluded the mutations that were also found in seven noncancerous samples.

**Mutation signatures analysis**

*De novo* extractions of mutation signatures and quantification of the contribution of Catalogue of Somatic Mutations in Cancer (COSMIC) signatures to the mutational profiles were performed, using R-3.5.1 (https://cran.r-project.org), MutationPatterns (https://github.com/UMCUgenetics/MutationalPatterns) and NMF packages (https://cran.r-project.org/web/packages/NMF/index.html).
Figure 1. Infinium 450k analysis of OPSCC and noncancerous samples from our study and TCGA. (a) HPV infection status. The presence of the HPV-L1 region was analyzed using PCR (142-bp PCR product), followed by gel electrophoresis. Most p16(+) cases were HPV-L1(+) (Case 1), whereas p16(−) cases were generally HPV-L1(−). (b) Immunostaining of p16 and p53. Top, hematoxylin and eosin; middle, p16; bottom, p53. p16(+) correlated inversely with TP53-mutation; that is, most p16(+) OPSCC samples showed TP53-mutation(−) status (Case 1). Most p16(−) OPSCC cases showed TP53-mutation(+), either positive staining of p53 (Case 2) or negative staining of p53 (Case 3). Scale bar, 100 μm. (c) Genome-wide DNA methylation analyzed for clinical OPSCC (n = 89) and noncancerous samples (n = 5). A probe nearest to the transcriptional start site (TSS) was selected for analysis of each promoter region. Probes with standard deviation of β-values > 0.10 in 89 clinical OPSCC and five noncancerous mucosa samples (n = 2,404) were used for clustering (top). Genes with CpG island shore regions upstream (N_Shore) and downstream (S_Shore) from TSS, and gene body regions are also shown (middle). The methylation data of PBMCs and fibroblasts are additionally shown as the normal control (right). White: HPV-L1(−), p16(−), TP53-mutation(−) and alive. Black: HPV-L1(+), p16(+), TP53-mutation(+) and dead. Color codes indicate race (bottom). (d) Genome-wide DNA methylation analysis for 81 clinical OPSCC samples obtained from TCGA. The same 2,404 probes were used for clustering (top). Genes with CpG island shore regions upstream (N_Shore) and downstream (S_Shore) from TSS, and gene body regions are also shown (middle). [Color figure can be viewed at wileyonlinelibrary.com]

Int. J. Cancer: 146, 2460–2474 (2020) © 2020 UICC
**Gene ontology analysis**

Gene annotation enrichment analysis was conducted based on GO (biological process, cellular component and molecular function) using the Functional Annotation tool of Metascape (http://metascape.org/gp/index.html#/main/step1).\(^{32}\)

**Pyrosequencing analysis**

Methylation status in 77 OPSCC cases and two normal samples was quantitatively validated using pyrosequencing with PyroMark Q96 (Qiagen) as previously reported.\(^{30}\) Primers were designed using the Pyro Q-CpG software (Qiagen) to...
amplify bisulfite-treated DNA (Supporting Information Table S1). Methylation control samples (0, 25, 50, 75 and 100%) were prepared as described previously\textsuperscript{33} and used to confirm the high quantification ability of pyrosequencing assays (Supporting Information Fig. S2).

**Statistical analysis**
Association between clinicopathological factors and DNA methylation or mutation status was analyzed using the Fisher’s exact test, \( \chi^2 \) test and Student’s \( t \)-test. Unsupervised two-way hierarchical clustering was performed using the Cluster 3.0 software. OS was measured from the date of registration to the date of OPSCC-related death, filtering out patients who were alive at last follow-up and non-OPSCC-related deaths. OS distribution was estimated using the Kaplan–Meier method. Differences between groups were determined using a log-rank test and Cox proportional hazard model using the \( R \) software (www.r-project.org/). Univariate and multivariate analysis for correlation of clinicopathological factors with prognosis was performed using the Cox proportional hazard model. \( p < 0.01 \) or <0.05 was considered statistically significant. False discovery rate was estimated by calculation of \( q \)-value by Benjamini–Hochberg method.

**Data availability**
Infinium data were submitted to the GEO database under the accession number GSE124633. Targeted exon-sequencing data were submitted to the Sequence Read Archive (SRA) database under the accession number: SRR8775783–SRR8775850.

**Results**

**Immunostaining analyses and HPV status**
HPV status was evaluated using PCR amplification of the L1 region in high-risk HPV with GP5/GP6 primers (Fig. 1a). Expression of \( \beta \)-NakA and p53 in HPV-associated OPSCC and HPV-negative OPSCC was analyzed using IHC (Fig. 1b). Whereas 49 of 51 HPV-associated OPSCC samples subjected to IHC were p16(+), only one HPV-negative OPSCC sample was p16(+). Most of the p16(+) cases, except two, were TP53-mutation(−). HPV status and p16(+) status correlated significantly with TP53-mutation(−) (\( p < 1 \times 10^{-5} \)), consistent with a previous study.\textsuperscript{31}

**Hierarchical clustering analysis of Infinium 450k data**
We performed Infinium 450k BeadArray analysis on 89 clinical OPSCC and five noncancerous mucosa samples (Fig. 1c). According to the previously proposed classification,\textsuperscript{29} Infinium probes were classified into three categories: high-CpG, intermediate-CpG and low-CpG probes, on the basis of CpG ratio and GC content within the 500-bp region (from −250 bp to +249 bp) around the probe site. In our study, we used high- and intermediate-CpG probes, and the probe closest to the TSS was selected for each gene for promoter analysis; 2,404 probes with standard deviation in \( \beta \)-values >0.10 in 89 clinical OPSCC and five noncancerous mucosa samples were selected.

Unsupervised 2-way hierarchical clustering analysis of 89 OPSCC and five normal mucosa samples clearly classified these samples into three clusters: high methylation epigenotype (HME), intermediate methylation epigenotype (IME) and low methylation epigenotype (LME). OPSCC samples in HME were highly methylated compared to those in the other epigenotypes (\( p = 2 \times 10^{-6} \)) and were all HPV-associated OPSCC (\( p < 1 \times 10^{-5} \); Fig. 1c). IME was the moderately methylated cluster including both HPV-associated and HPV-negative OPSCC samples. LME was the lowest methylation cluster, including the five noncancerous mucosa samples and mainly HPV-negative OPSCC samples. We also investigated CpG island shore regions (N_Shore: upstream from TSS and S_Shore: downstream from TSS) and gene body regions. All of these regions in each cluster were generally hypermethylated.

To validate these epigenotypes, we performed unsupervised two-way hierarchical clustering analysis of 81 OPSCC data uploaded by TCGA using the same 2,404 probes. OPSCC samples were clearly classified into three epigenotypes: HME, IME and LME (Fig. 1d). While almost all HME and IME samples were HPV-associated OPSCC, LME mainly included HPV-negative samples. Outlier samples mainly included HPV-negative OPSCC with intermediate level of methylation.

**Stratification of OPSCC into four subtypes**
For clearer stratification, we combined our methylation data of 89 OPSCC samples and TCGA’s methylation data of 81 OPSCC samples (Fig. 2a). We selected 1,315 high- and intermediate-CpG probes with standard deviation of \( \beta \)-values >0.10 in the 170 clinical OPSCC and five noncancerous

---

**Figure 2.** Infinium 450k analysis of OPSCC and noncancerous samples from our study and TCGA. (a) Genome-wide DNA methylation analyzed for 170 clinical OPSCC samples (89 from our study and 81 from TCGA). High- and intermediate-CpG probes with standard deviation of \( \beta \)-values >0.10 in 170 clinical OPSCC and five noncancerous mucosa samples (\( n = 1,315 \)) were selected. Then, 170 OPSCC cases were divided into mainly four subgroups (top). The methylation data of PBMC and fibroblast as the normal control and PRC-target genes in ES cells (black bars) are also shown. Genes with CpG island shore regions upstream (N_Shore) and downstream (S_Shore) from TSS, and gene body regions are shown (middle). Methylation levels in each epigenotype were compared using Student’s \( t \)-test (right panels). The difference in methylation levels was most apparent in high and intermediate CpG promoter regions, especially between HPV(+)HME versus HPV(−)LME (\( p = 2 \times 10^{-116}; q\text{-value} = 2 \times 10^{-115} \)). \( *p < 0.01 \) (\( t \)-test) and \( q < 0.1 \) (Benjamini–Hochberg method). **White:** HPV-L1(−), p16(−), TP53-mutation(−) and alive. **Black:** HPV-L1(+), p16(+), TP53-mutation(+). (b) Feature of each epigenotype. OP1: HPV(+) high-methylation epigenotype (HME). OP2: HPV(+) intermediate-methylation epigenotype (IME). OP3: HPV(−) intermediate-methylation epigenotype (IME). OP4: HPV(−) low-methylation epigenotype (LME). [Color figure can be viewed at wileyonlinelibrary.com]
mucosa samples for this analysis. The 170 OPSCC cases were classified into mainly four subgroups: OP1, HPV(+) HME; OP2, HPV(+) IME; OP3, HPV(−) IME; OP4, HPV(−) LME (Fig. 2b). An additional cluster including all the normal samples was considered as “Others (Normal-like).” Upon increasing the number of samples, IME was stratified into HPV(+) IME and HPV(−) IME. The OP1 subgroup (HPV(+) HME) showed markedly higher methylation levels in promoter region than the other subgroups when a cluster of 486 genes was extracted from the 1,315 high- and
intermediate-CpG genes (β-value 0.31 ± 0.00 in OP1, 0.22 ± 0.00 in OP2, 0.18 ± 0.00 in OP3 and 0.15 ± 0.00 in OP4; Fig. 2a).

We also investigated methylation levels in CpG island shore (N_Shore and S_Shore) and gene body regions. All these regions in each epigenotype generally showed higher level of DNA methylation, and the tendencies of average β-values in each epigenotype were similar in promoter region, except for the gene body between HPV(+) IME and HPV(−) IME. In summary, methylation levels were altered most apparently in promoter region, and we focused on promoter region for subsequent analyses.

**Classifier genes of the four subtypes**

Next, we extracted classifier marker genes to easily distinguish each epigenotype (Fig. 3). Initially, we divided HPV(+) OPSCC into OP1 and OP2 subgroups, and HPV(−) OPSCC into OP3 and OP4 subgroups. Exceptional samples, for example, HPV-negative OPSCC samples in OP1 or OP2 and HPV-associated OPSCC samples in OP3 and OP4, were excluded for this procedure. Among 15,392 high- and intermediate-CpG genes, 61 genes were extracted as OP1 markers showing hypermethylation in OP1 but not in OP2, while 68 genes extracted as OP1 + OP2 markers showing hypermethylation in both OP1 and OP2. While 1,513 of 15,392 genes (9.8%) were targeted by polycomb repressive complex (PRC) in ES cells,24 PRC-target genes were not significantly frequent in OP1 markers (11%, 7/61; p = 0.5, Fisher’s exact test), but significantly more frequent in OP2 markers (29%, 20/68; p < 1 × 10−5; Fig. 3a).

As for HPV(−) group, 81 genes were extracted as OP3 markers showing hypermethylation in OP3 but not in OP4, while 41 genes were extracted as OP3 + OP4 markers showing hypermethylation in both OP3 and OP4 (Fig. 3b). PRC-target genes in ES cells were not significantly frequent in OP3 markers (15%, 12/81; p = 0.1), but more frequently included in OP3 + OP4 markers significantly (24%, 10/41; p = 0.003; Fig. 3b). Marker genes specifically hypermethylated in a subgroup, such as OP1 and OP3, did not show enrichment of PRC-target genes in ES cells, while those hypermethylated commonly, such as OP1 + OP2 and OP3 + OP4, showed significant enrichment of PRC-target genes, as observed in other malignancies.30

Given these results, the five most valuable genes were chosen from OP1 and OP3 markers as HPV(+) and HPV(−) classifier markers, respectively (Figs. 3 and 4). Five HPV(+) classifier markers were MYOD1, CDH11, PLCB4, FGF5 and PIK3AP1 (Fig. 4a). The accuracy of classification was evaluated by ROC analysis based on true and false positive rates using quantitative DNA methylation levels of the classifier markers (Fig. 4b). When OPSCC sample showing hypermethylation (β > 0.2) in ≥2 out of 5 HPV(+) classifier markers was determined as OP1, the calculated AUC values of our cohort was 1.00 for our cohort, and as high as 0.933 if TCGA samples and our cohort were combined. The methylation levels of the markers were then validated by quantitative methylation analysis using pyrosequencing (Fig. 4a). We analyzed the methylation levels of CpG sites at the selected Infinium probe site as well as its surrounding CpG sites, using OPSCC and noncancerous samples that were previously used for Infinium analysis. Methylation levels were confirmed to be highly similar, not only for the selected probe site, but also the surrounding CpG sites (Fig. 4b). When OPSCC sample showing methylation level >20% in ≥2 out of 5 HPV(+) classifier markers was determined as OP1 by using pyrosequencing data, the calculated AUC value was as high as 0.969.

Five HPV(−) classifier markers were RAB3C, JAM3, PCDH20, FAM78A and FGFR3 (Fig. 4c). When OPSCC sample showing hypermethylation (β > 0.2) in ≥2 out of 5 HPV(−) classifier markers was determined as OP3, the calculated AUC values of our cohort was 1.00 for our cohort, and as high as 0.934 if TCGA samples and our cohort were combined (Fig. 4d). The methylation levels of the markers were analyzed by pyrosequencing, and were confirmed to be highly similar, not only for the selected probe site, but also the surrounding CpG sites. When OPSCC sample showing methylation level >20% in ≥2 out of 5 HPV(−) classifier markers was determined as OP3 by using pyrosequencing data, the calculated AUC value was as high as 0.952.

Figure 3. Classifier marker genes of the epigenotypes. (a) Extraction of markers to classify HPV(+) OPSCC into OP1 and OP2. Sixty-one OP1 markers were methylated in OP1 but not methylated in OP2 (p = 5 × 10−36, q-value = 9 × 10−29) or normal samples (p = 3 × 10−36, q-value = 8 × 10−34) and five markers were selected for further analysis (Fig. 4). Sixty-eight OP1 + OP2 markers were methylated in OP1 and OP2 but not methylated in normal samples (OP1 vs. normal samples, p = 1 × 10−6, q-value = 2 × 10−6); OP2 vs. normal samples, p = 6 × 10−42, q-value = 2 × 10−45). The ratio of the PRC-target genes is shown (middle). OP1 marker genes did not show significant enrichment of PRC-target genes in ES cells (p = 0.5, t-test), whereas OP1 + OP2 marker genes included significantly higher number of PRC-target genes (p < 1 × 10−5, t-test). OP1 markers showed significant enrichment of GO-terms, "regulation of synapse structure or activity" and "lipid phosphorylation", while no GO-term was significantly enriched in OP1 + OP2 markers (right). (b) Extraction of markers to classify HPV(−) groups into OP3 and OP4. Eighty-one OP3 markers were methylated in OP3, but not methylated in OP4 (p = 2 × 10−48, q-value = 6 × 10−48) or normal samples (p = 2 × 10−38, q-value = 1 × 10−35), and five markers were selected for further analysis (Fig. 4). Forty-one OP3 + OP4 markers were methylated in OP3 and OP4, but not methylated in normal samples (OP3 vs. normal samples, p = 2 × 10−25, q-value = 2 × 10−28); OP4 vs. normal samples, p = 3 × 10−27, q-value = 5 × 10−29). The ratio of the PRC-target genes is shown (middle). OP3 marker genes did not show significant enrichment of PRC-target genes in ES cells (p = 0.1, t-test), whereas OP3 + OP4 marker genes included significantly higher number of PRC-target genes (p = 0.003). No GO-term was significantly enriched in OP3 and OP3 + OP4 markers (right), *p < 0.01 (t-test) and q < 0.1 (Benjamini–Hochberg method). 1p < 1 x10−17 and q < 0.1 (GO analysis by Metascape). [Color figure can be viewed at wileyonlinelibrary.com]
Targeted exon sequencing analysis of 86 OPSCC samples

To gain insight into genomic mutation profiles of the identified epigenotypes, we performed targeted exon sequencing using the Human Comprehensive Cancer Panel on 86 OPSCC samples and seven noncancerous samples. The total counts of somatic single nucleotide variants (SNVs) and insertions and
deletions (indels) were 34,988. We obtained an average coverage of 385× (ranging from 293× to 726×). The mutation frequency did not vary significantly between HPV(+) and HPV(−) cases (Supporting Information Fig. S3a) or among the four epigenotypes (Supporting Information Fig. S3b).

We reclassified the 86 OPSCC samples into OP1 - OP4 epigenotypes using HPV status and DNA methylation information of classifier markers, and compared gene mutations, copy number variations and clinicopathological data of OPSCC samples (Supporting Information Fig. S4). The representative carcinogenic pathways, namely, those involved in p53 signaling, cell cycle and survival, RTK/RAS signaling, PI3K signaling, epigenome and differentiation, were compared. For the HPV(−) group, OP3 and OP4 showed significantly higher number of somatic mutations in TP53 and CCND1 copy number gains \( (p < 0.01) \). For carcinogenic pathways, genes associated with p53 signaling and cell cycle and survival showed a significantly higher number of somatic mutations and copy number variations \( (p < 0.01) \). For the HPV(+) group, p16(+) and TP53-mutation(−) that were determined by IHC, as well as subtype of primary tumors (tonsils and base of tongue) were significantly more associated with OP1 and OP2 than OP3 and OP4 \( (p < 0.01) \). Notably, OP1 included remarkably lower number of dead patients, while OP3 included remarkably higher number of dead patients \( (p < 0.01) \). Any molecular markers other than DNA methylation did not significantly distinguish OP1 and OP2, or OP3 and OP4 (Supporting Information Fig. S4).

**Combined analysis of genomic alterations using our and TCGA’s data**

To further clarify the relationship among mutation status, clinicopathological factors and each epigenotype, we analyzed mutations and copy number variations by combining data of our 86 OPSCC samples and TCGA’s 74 OPSCC samples (Fig. 5). In this combined analysis, mutations that were classified as “silent” were excluded. OP3 and OP4 showed significantly higher number of somatic mutations in TP53 and CCND1 copy number gains \( (p < 0.01) \). The incidence of somatic mutations and copy number variations in genes associated with p53 signaling and cell cycle and survival were also significantly high in OP3 and OP4 \( (p < 0.01) \). The subsites of primary tumors (tonsils and base of tongue) were significantly highly represented in the OP1 and OP2 \( (p < 0.01) \). OP3 included remarkably higher number of dead patients \( (p < 0.01) \), whereas OP1 included remarkably lower number of dead patients \( (p < 0.01) \). Any molecular markers other than DNA methylation did not significantly distinguish OP1 and OP2, or OP3 and OP4 (Fig. 5).

**Comparison of epigenotype and mutation signatures**

To compare epigenotypes with mutation signatures, we performed mutation signature analysis for a total of 34,988 SNVs in OPSCC samples. Mutation signature analysis did not divide OPSCC samples into any subtypes, showing highly similar mutation signatures among epigenotypes (Supporting Information Fig. S5a). All the epigenotypes had frequent C>A (35–36%) and C>T mutations (41–42%; Supporting Information Fig. S5b). OPSCC samples were highly enriched in Signature 1, which correlated with age of cancer diagnosis (Supporting Information Fig. S6).

**Prognostic analysis of OPSCC samples with respect to HPV status and epigenotypes**

We finally investigated the association of epigenotype and HPV status with prognosis. HPV-associated OPSCC showed significantly better prognosis than HPV-negative OPSCC \( (p = 0.0001, HR 0.15 \ [95\%CI, 0.069–0.35]) \); Fig. 6a), in line with previous reports.\(^1\) TP53-mutation(−) samples also showed significantly better prognosis than TP53-mutation(+) samples \( (p = 0.02, HR 17 \ [95\%CI, 2.2–130]) \); Supporting Information Fig. S7). TNM stage did not significantly correlate with OPSCC prognosis \( (p = 0.4, HR 1 \ [95\%CI, 0.73–1.4]) \); Supporting Information Fig. S8). Intriguingly, epigenotype comparison revealed that the four epigenotypes showed significantly different prognosis \( (p = 0.0006, HR 2.0 \ [95\%CI, 1.4–2.7]) \); Fig. 6b). Among generally favorable HPV(+) OPSCC, OP1 showed the most favorable outcome, with OP2 being relatively worse \( (p = 0.02) \). Among generally unfavorable HPV(−) OPSCC, OP3 showed the most unfavorable outcome, with OP4 being relatively better \( (p = 0.03) \). These indicate that HPV-associated favorable OPSCC and HPV-negative
Figure 5. Comparison of each epigenotype, genomic alterations and clinicopathological factors. Then, 161 OPSCC samples (n = 86 from our study, n = 75 from TCGA) were classified into four subtypes based on HPV status and methylation status of classifier markers (top). Clinicopathological data (middle), mutations and copy number variations in individual genes and those associated with carcinogenic signaling pathways (bottom) were compared. OP1 included remarkably lower number of dead patients, whereas OP3 included remarkably higher number of dead patients (p < 0.01, χ² test). The subsite of the tonsils and the base of the tongue were significantly well-represented in the HPV(+) groups (OP1 and OP2), than in the HPV(−) groups (OP3 and OP4). The number of TP53 mutation and CCND1 copy number gains were significantly higher in the HPV(−) groups. Genes associated with p53 signaling, or cell cycle and survival, harbored significantly large number of somatic mutations and copy number variations in the HPV(−) groups (p < 0.01, Fisher’s exact test). *p < 0.01, Fisher’s exact test or χ² test. [Color figure can be viewed at wileyonlinelibrary.com]
unfavorable OPSCC can be further divided by DNA methylation epigenotypes.

We also analyzed the prognosis based on combination of HPV status and smoking or alcohol consumption status. The HPV(+) OPSCC showed significantly better prognosis than HPV(−) samples \((p = 0.0001, \text{HR} 0.15 [95\% \text{CI}, 0.069-0.35])\). Analysis of prognosis based on epigenotype. The four epigenotypes showed significantly different prognosis \((p = 0.0006, \text{HR} 2.0 [95\% \text{CI}, 1.4-2.7])\). OP1 showed significantly better prognosis than OP2 \((p = 0.02, \text{Cox proportional hazard model})\) in the HPV(+) group, and OP3 showed significantly worse prognosis than OP4 \((p = 0.03, \text{Cox proportional hazard model})\) in the HPV(−) group. \(*p < 0.01, \text{log-rank test or Cox proportional hazard model}. \) [Color figure can be viewed at wileyonlinelibrary.com]

**Discussion**

To elucidate the molecular mechanisms underlying intermediate risk HPV-associated OPSCC and generally unfavorable HPV-negative OPSCC, we here analyzed DNA methylome by
the Infinium 450K, genetic alterations by targeted exon sequencing, and their correlations with clinicopathological factors, using our cohort of 89 OPSCC in combination with TCGA’s 81 OPSCC samples. We found that OPSCC is stratified into four epigenotypes (OP1–OP4), reflecting different clinicopathological features. OP1 showed the most favorable outcome among generally favorable HPV(+) OPSCC, and OP3 showed the most unfavorable outcome among generally unfavorable HPV(−) OPSCC. These indicate that DNA methylation epigenotypes can further stratify HPV-associated and HPV-negative OPSCC, reflecting distinct prognosis, which could not be achieved by other features such as genomic mutations.

The correlation between DNA methylation and cancer has been observed in several cancer types, especially in virus-associated cancers such as gastric cancer (associated with Epstein–Barr virus, EBV), hepatocellular carcinoma (associated with hepatitis B and C viruses), cervical cancer, and HNSCC (associated with HPV), partially because aberrant DNA methylation can be induced by virus infection during carcinogenesis. Lleras et al. showed that HPV-associated HNSCC samples harbored approximately three times higher methylation than HPV-negative HNSCC samples. We have previously observed aberrant promoter hypermethylation in OPSCC samples, and showed that HPV-associated OPSCC samples had significantly higher methylation level.

Two previous studies based on comprehensive genome-wide data of HNSCC by TCGA, investigated distinct subtypes of HNSCC in association of HPV status, although these data included only 33 and 81 OPSCC samples, respectively, and mainly included other HNSCC samples. Lechner et al. also performed an DNA methylation analysis of 32 OPSCC samples on a genome-wide scale using formalin-fixed paraffin-embedded tissues, and showed hypermethylation of genes including CDH1 in HPV(+) OPSCC cases. Despite clinical requirements, there has been no study analyzing solely OPSCC samples in a sufficiently large sample size. To the best of our knowledge, this is the first study to stratify OPSCC cases using an adequate sample size, and successfully into four DNA methylation epigenotypes. Favorable HPV(+) OPSCC as well as unfavorable HPV(−) OPSCC are further divided into subgroups reflecting distinct prognosis, and two panels of classifier markers were established.

The complex mutation landscape of HNSCC has been elucidated by several studies, including TCGA projects. HPV-negative HNSCC is mainly characterized by genomic loss of TP53 and genetic alteration of CDKN2A/RB1, and is also associated with mutations in genes involved in carcinogenic signaling of differentiation (NOTCH1 and TP63), oxidative stress (NFE2L2 and KEAP1), WNT signaling (AUF1 and FAT1), immune evasion (HLA-A, B2M and TGFBR2) and chromatin remodeling (MLL2 and NSD1). The amplification of genes encoding receptor tyrosine kinases (RTKs) such as EGFR, ERBB2, MET and FGFR1 are also observed in this cancer type. In contrast, HPV-associated HNSCC are characterized by few genomic alterations in TP53 and CDKN2A, but harbor PIK3CA alterations, copy-number gains in TRAF3 and E2F1 and lack of CCND1 amplification, compared to HPV-negative cancer. While studies solely focusing on OPSCC are lacking, we compared our OPSCC data of DNA methyleme as well as targeted exon sequencing with TCGA data for OPSCC cases. Both our data and TCGA data showed that TP53 mutations, CCND1 copy number gains and alterations in genes associated with p53 signaling and cell cycle and survival were more frequent in HPV(−) than in HPV(+) OPSCC, consistent with the previous studies.

Although the relationship between genetic and epigenetic alterations has been previously reported in other cancer types, such as colorectal cancer exhibiting correlation of high methylation with BRAF mutation or intermediate methylation with KRAS mutation, reports on HNSCC are limited. Loss-of-function mutation in NSD1, which regulates the H3K36 methylation pathway, was previously reported in HNSCC, and the H3K36 alteration subgroup in our study showed marked DNA hypomethylation. However, only three OPSCC samples were included in this subgroup, which mainly constituted from laryngeal cancer and oral cancer samples. In our OPSCC cohort with 89 cases, only six samples harbored NSD1 mutation, and there was no significant correlation with any epigenotype. Although genetic mutation could stratify OPSCC samples into two types, HPV(+) OPSCC and HPV(−) OPSCC, it could not further divide HPV(+) or HPV(−) OPSCC into subtypes, indicating that, besides genetic alterations, the accumulation of aberrant epigenetic alterations might deeply influence OPSCC biology and contribute to additional subclassification of both HPV-associated and HPV-negative OPSCC.

Whereas HPV-associated OPSCC shows generally better prognosis than HPV-negative OPSCC, a fraction of HPV-associated OPSCC cases has reportedly exhibited therapeutic resistance or worse prognosis. Ang et al. reported the importance of smoking status combined with tumor stage, in stratifying HPV-associated OPSCC, which was subsequently validated by Granata et al. However, the molecular aberrations underlying this intermediate risk is yet to be clarified despite previous genome-wide and epigenome-wide studies in HNSCC. This could be because only a few epigenome analyses had focused solely on OPSCC, and even with these, the cohort size was small.

We here performed DNA methylation analysis of the largest OPSCC cohort on a genome-wide scale, and 170 OPSCC cases were successfully classified into four epigenotypes reflecting the prognostic status, while TNM stage did not correlate with prognosis. While there have been no molecular markers available that clinically distinguish low-risk patients from intermediate-risk patients in HPV-associated OPSCC, we here established highly accurate classifier marker panels to easily categorize OPSCC into distinct epigenotypes. While
smoking status was reportedly associated with subgroup of HPV-associated OPSCC with relatively worse prognosis, smoking status was not significantly associated with OP2 or any other epigenotypes (Fig. 5), and HPV(+) or HPV(−) OPSCC was not further stratified significantly by smoking status, suggesting that independent molecular markers that are more accurate than smoking status were identified in our study. In addition, multivariate analysis estimated our classification into four distinct epigenotypes as the only independent prognostic factor.

Although we successfully divided OPSCC into four epigenotypes, the cause of DNA methylation remains unclear. To explain two distinct patterns of aberrant DNA methylation accumulation in HME and IME, at least two different environmental factors might be involved. As HME is observed only in HPV(+) groups, HME might be related to HPV infection. The most direct explanation is that HPV might induce DNA methylation. E7, an oncoprotein of HPV, has been shown to directly interact with DNA methyltransferase 1 (DNMT1) in vitro, and increased expression of DNMT1 has been observed in HPV-associated OPSCC and cervical cancer, which is also an HPV-associated cancer. E6/E7 is reportedly upregulated in the case of E2-disrupted HPV integration or DNA methylation of E2-binding sites (E2BSs). Parfenov et al. reported in a relatively small 35 HNSCC cohort (including 22 OPSCC cases) that DNA methylation profiles of HPV-integration(+) tumors were similar to those of HPV-negative tumors, whereas HPV-integration(−) tumors showed relatively higher methylation. While no significant association was observed between HPV-integration status and clinical outcome in this 35 HNSCC cohort, Nulton et al. reported a significant relationship between HPV-integration(+) and poor survival in 56 HPV-associated HNSCC. In gastric cancer, while infected EBV genome exists in episomal form and EBV infection was shown to cause extensive DNA methylation induction, EBV(+) gastric cancer shows markedly high DNA methylation and favorable clinical outcome. Considering these, episomal HPV might be involved in DNA methylation induction and form a subgroup with high methylation and relatively good prognosis; further investigations, for example, DNA methylation status of E2BSs in episomal HPV genome, are however required to fully clarify this.

IME exists in both of HPV(+) and HPV(−) OPSCC (OP2 and OP3, respectively), and corresponds to poorer prognosis regardless of HPV status. A certain factor independent of HPV might be expected to involve DNA methylation induction to develop IME and poor prognosis. Our present study, however, showed no significant correlation of IME (OP2 and OP3) with any molecular or clinicopathological features including smoking, alcohol intake or genetic alterations. Further detailed investigations are also necessary to clarify the causes of IME and its association with poor prognosis.

In summary, our DNA methylation epigenotypes distinguished the most favorable OPSCC subgroup (OP1) from OP2 among generally favorable HPV(+) cases, and the most unfavorable OPSCC subgroup (OP3) from OP4 among generally unfavorable HPV(−) cases. We also established classifier marker panels, while such subclassification of OPSCC could not be achieved by other features such as genomic mutation. These findings might improve our understanding of the molecular basis of OPSCC and de-escalation treatment strategies.

Acknowledgements
We thank Chihiro Sato, Keiko Iida, Haruka Maruyama and Aki Komatsu for technical assistance, and Editage (www.editage.jp) for English language editing. Our study was supported by the Chiba University (2018-Y9, Global and Prominent Research grant), Japan Agency for Medical Research and Development (19ck0106253h0003) and Japan Society for the Promotion of Science (18K16831, 19H03726 and 24249081).

Conflict of interest
The authors have no conflict of interest to declare.

References
1. Ferlay J, Shin HR, Bray F, et al. Estimates of worldwide burden of cancer in 2008. GLOBOCAN 2008. Int J Cancer 2010;127: 2893–917.
2. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359–86.
3. Scudellari M. HPV: sex, cancer and a virus. Nature 2013;503:330–2.
4. Chaturvedi AK, Engels EA, Anderson WF, et al. Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. J Clin Oncol 2008;26:10026–9.
5. Ryerson AB, Peters ES, Coughlin SS, et al. Burden of potentially human papillomavirus-associated cancers of the oropharynx and oral cavity in the US, 1998-2003. Cancer 2008;113:2901–9.
6. Ndiaye C, Mena M, Alemany L, et al. HPV DNA, E6/E7 mRNA, and p16/INK4a detection in head and neck cancers: a systematic review and meta-analysis. Lancet Oncol 2014;15:1319–31.
7. Fakhry C, Westra WH, Li S, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. J Natl Cancer Inst 2008;100:261–9.
8. Grisor K, Dok R, Schoeneers J, et al. Differences in human papillomavirus-positive and -negative head and neck cancers in Belgium: an 8-year retrospective, comparative study. Oral Surg Oral Med Oral Pathol Oral Radiol 2016;121:456–60.
9. Rago CC, Taioi E. Survival of squamous cell carcinoma of the head and neck in relation to human papillomavirus infection: review and meta-analysis. Int J Cancer 2007;121:1813–20.
10. Taberna M, Mena M, Pavon MA, et al. Human papillomavirus-related oropharyngeal cancer. Ann Oncol 2017;28:2386–98.
11. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med 2010;363:24–35.
12. Seiwert TY, Zuo Z, Keck MK, et al. Integrative and comparative genomic analysis of HPV-positive and HPV-negative head and neck squamous cell carcinomas. Clin Cancer Res 2015;21:632–41.
13. Liu VW, Hedberg ML, Li H, et al. Frequent mutation of the PI3K pathway in head and neck cancer defines predictive biomarkers. Cancer Discov 2013;3:761–9.
14. Agrawal N, Frederick MJ, Pickering CR, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science 2011;333:1154–7.
15. Stranksy N, Egloff AM, Tward AD, et al. The mutational landscape of head and neck squamous cell carcinoma. Science 2011;333:1157–60.

16. van Kempen PM, Noorlag R, Braaninuis WW, et al. Differences in methylation profiles between HPV-positive and HPV-negative oropharynx squamous cell carcinoma: a systematic review. Epigenetics 2014;9:194–203.

17. Chung CH, Gillison ML. Human papillomavirus in head and neck cancer: its role in pathogenesis and clinical implications. Clin Cancer Res 2009;15:6758–62.

18. Lechner M, Frampton GM, Fenton T, et al. Targeted next-generation sequencing of head and neck squamous cell carcinoma identifies novel genetic alterations in HPV+ and HPV− tumors. Genome Med 2013;5:49.

19. Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 2015;517:576–82.

20. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. Carcinogenesis 2010;31:27–36.

21. Weiss D, Basel T, Sachse F, et al. Promoter methylation of cyclin A1 is associated with human papillomavirus 16 induced head and neck squamous cell carcinoma independently of p53 mutation. Mol Carcinog 2011;50:680–8.

22. Anayannis NV, Schlecht NF, Bellini TJ. Epigenetic mechanisms of human papillomavirus-associated head and neck cancer. Arch Pathol Lab Med 2015;139:1373–8.

23. Kaneda A, Matussaka K, Aburatani H, et al. Epstein-Barr virus infection as an epigenetic driver of tumorigenesis. Cancer Res 2012;72:3445–50.

24. Herceg Z, Paliwal A. Epigenetic mechanisms in hepatocellular carcinoma: how environmental factors influence the epigenome. Mutat Res Rev 2011;727:55–61.

25. Sartor MA, Dolinoy DC, Jones TR, et al. Genome-wide methylation and expression differences in HPV(+) and HPV(−) squamous cell carcinoma cell lines are consistent with divergent mechanisms of carcinogenesis. Epigenetics 2011;6:777–87.

26. Lleras RA, Smith RV, Adrien LR, et al. Unique DNA methylation loci distinguish anatomic site and HPV status in head and neck squamous cell carcinoma. Clin Cancer Res 2013;19:5444–55.

27. Papillon-Cavanagh S, Lu C, Gayden T, et al. Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. Nat Genet 2017;49:180–5.

28. Husnjak K, Grce M, Magdic L, et al. Comparison of five different polymerase chain reaction methods for detection of human papillomavirus in cervical cell specimens. J Virol Methods 2000;88:125–34.

29. Weber M, Hellmann I, Stadler MB, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 2007;39:547–66.

30. Matussaka K, Kaneda A, Nagae G, et al. Classification of Epstein-Barr virus-positive gastric cancers by definition of DNA methylation epigenotypes. Cancer Res 2011;71:1877–97.

31. Nakagawa T, Matussaka K, Miwa K, et al. Frequent promoter hypermethylation associated with human papillomavirus infection in pharyngeal cancer. Cancer Lett 2017;407:21–31.

32. Tripathi S, Pohl MO, Zhou Y, et al. Meta- and orthogonal integration of influenza “OMICS” data defines a role for UBR4 in virus budding. Cell Host Microbe 2015;18:723–35.

33. Yagi K, Akagi K, Hayashi H, et al. Three DNA methylation epigenotypes in human colorectal cancer. Clin Cancer Res 2010;16:21–33.

34. Ku M, Koche RP, Rheinbay E, et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet 2008;4:e1000242.

35. Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. Cell 2017;169:1327–41.e23.

36. Cancer Genome Atlas Research Network. Albert Einstein College of Medicine. Biological Services, et al. Integrated genomic and molecular characterization of cervical cancer. Nature 2017;543:378–84.

37. Lechner M, Fenton T, West J, et al. Identification and functional validation of HPV-mediated hypermethylation in head and neck squamous cell carcinoma. Genome Med 2013;5:15.

38. Hammerman PS, Hayes DN, Grandis JR. Therapeutic insights from genomic studies of head and neck squamous cell carcinomas. Cancer Discov 2015;5:239–44.

39. Yagi K, Takahashi H, Akagi K, et al. Intermediate methylation epigenotype and its correlation to KRAS mutation in conventional colorectal adenoma. Am J Pathol 2012;180:616–25.

40. Sakai E, Ohata K, Chiba H, et al. Methylation epigenotypes and genetic features in colorectal spreading tumors. Int J Cancer 2014;135:1586–95.

41. Boscolo-Rizzo P, Furlan C, Lupato V, et al. Novel insights into epigenetic drivers of oropharyngeal squamous cell carcinoma: role of HPV and lifestyle factors. Clin Epigenetics 2017;9:124.

42. Granata R, Miceli R, Orlando E, et al. Tumor stage, human papillomavirus and smoking affect the survival of patients with oropharyngeal cancer: an Italian validation study. Ann Oncol 2012;23:1832–7.

43. Kostareli E, Holzinger D, Bogatyrova O, et al. HPV-related methylation signature predicts survival in oropharyngeal squamous cell carcinomas. J Clin Invest 2013;123:2488–501.

44. Parfenov M, Pedamallus C, Gehlenborg N, et al. Characterization of HPV and host genome interactions in primary head and neck cancers. Proc Natl Acad Sci U S A 2014;111:15544–9.

45. Degli Esposti D, Sklias A, Lima SC, et al. Unique DNA methylation signature in HPV-positive head and neck squamous cell carcinomas. Genome Med 2017;9:33.

46. Burgers WA, Blanchon L, Pradhan S, et al. Viral oncoproteins target the DNA methyltransferases. Oncogene 2007;26:1650–5.

47. Schlecht NF, Ben-Dayan M, Anayannis N, et al. Epigenetic changes in the CDKN2A locus are associated with differential expression of P16INK4A and P14ARF in HPV-positive oropharyngeal squamous cell carcinoma. Cancer Med 2015;4:342–53.

48. Sawada M, Kanai Y, Arai E, et al. Increased expression of DNA methyltransferase 1 (DNMT1) protein in uterine cervix squamous cell carcinoma and its precursor lesion. Cancer Lett 2007;241:211–9.

49. Reuschbach M, Hubebers CU, Prigge ES, et al. Methylation status of HPV16 E2-binding sites classifies subtypes of HPV-associated oropharyngeal cancers. Cancer 2015;121:1966–76.

50. Nulton T, Kim NK, DiNardo LJ, et al. Patients with integrated HPV16 in head and neck cancer show poor survival. Oral Oncol 2018;80:52–5.