Genome-wide gene expression profiles of HPV-positive and HPV-negative oropharyngeal cancer: potential implications for treatment choices

Pawadee Lohavanichbutr[1], John Houck[1], Wenhong Fan[2], Bevan Yueh[3], Eduardo Mendez[1,4,5], Neal Futran[5], David R. Doody[1], Melissa P. Upton[6], D. Gregory Farwell[7], Stephen M. Schwartz[1,8], Lue Ping Zhao[1,2], Chu Chen[1,5,8]

Authors' Affiliations:

Programs in [1]Epidemiology and [2]Biostatistics and Biomathematics, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington; [3]Department of Otolaryngology: Head and Neck Surgery, University of Minnesota, Minneapolis, Minnesota. [4]Surgery and Perioperative Care Service, VA Puget Sound Health Care System, Seattle, Washington; Departments of [5]Otolaryngology: Head and Neck Surgery, and [6]Pathology, University of Washington, Seattle, Washington; [7]Department of Otolaryngology: Head and Neck Surgery, University of California, Davis, Davis, California; and [8] Department of Epidemiology, University of Washington, Seattle, Washington

Running Title: Gene expression in HPV+ and HPV- oropharyngeal cancer

Key Words: Human papillomavirus, oral cancer, gene expression

Grant support: NIH NCI R01CA095419 (PI: Chu Chen)

Requests for reprints: Chu Chen, Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Mail stop M5-C800, 1100 Fairview Avenue North, Seattle, WA 98109-1024. Phone: 206-667-6644; Fax: 206-667-2537; E-mail: cchen@fhcrc.org

Disclaimer: The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

Word Count: 2,968

Abstract

Objective: To study the difference in gene expression between human papillomavirus (HPV)-positive and HPV-negative oral squamous cell carcinoma (OSCC).

Design: We used Affymetrix U133 plus 2.0 arrays to examine gene expression profiles of OSCC and normal oral tissue. HPV DNA was detected using PCR followed by the Roche Linear Array HPV Genotyping Test, and the differentially expressed genes were analyzed to examine their potential biological roles using the Ingenuity Pathway Analysis Software (IPA 5.0).

Subjects: Tumor tissue from 119 primary OSCC patients and normal oral tissue from 35 patients without cancer, all of whom were treated at three University of Washington-affiliated medical centers.

Results: HPV DNA was found in 41 of 119 (34.5%) tumors and 2 of 35 (5.7%) normal tissue samples, with 39 of 43 HPV being HPV type 16; there was a higher prevalence of HPV DNA in oropharyngeal cancer (23 of 31) than in oral cavity cancer (18 of 88). We found no significant difference in gene expression between HPV-positive and HPV-negative oral cavity cancer but found 446 probe sets (347 known genes) differentially expressed between HPV-positive and HPV-negative oropharyngeal cancer. The most prominent functions of these genes are DNA replication, DNA repair, and cell cycle. Some genes differentially expressed between HPV-positive and HPV-negative oropharyngeal cancer (e.g., *TYMS, STMN1, CCND1* and *RBBP4*) are involved in chemotherapy or radiation sensitivity.

Conclusion: These results suggest that differences in the biology of HPV-positive and HPV-negative oropharyngeal cancer may have implications for the management of patients with these different tumors.

Oral cavity and oropharyngeal squamous cell carcinomas (OSCC) constitute a major public health burden worldwide. Approximately 400,000 new cases of OSCC were diagnosed in 2002, and approximately 200,000 patients died of these cancers [1]. The major risk factors for OSCC are cigarette smoking, alcohol consumption, and betel quid chewing. The evidence for HPV as a potential etiologic agent in OSCC was first reported in 1983, when the presence of HPV antigen was shown in oral cancer specimens [2]. Subsequent publications showed an association between infection with high risk types of HPV and OSCC risk [3-6]. Molecular and epidemiologic studies strongly suggest that HPV-positive OSCC comprise a distinctive disease entity that differs from HPV-negative OSCC in molecular, histopathologic and prognostic characteristics [7-10], with HPV-positive OSCC less frequently associated with p53 mutations, primarily located in the oropharynx, tending to be poorly differentiated and basaloid subtypes, and having a more favorable disease outcome. [7]. To understand the molecular mechanisms underlying these two entities of OSCC, we examined genome-wide gene expression profiles of HPV-positive and HPV-negative OSCC.

Materials and Methods.

Study population

Eligible cases were patients with first incident primary OSCC scheduled for surgical resection or biopsy between December 2003 and May 2006 at one of three University of Washington-affiliated medical centers in Seattle, WA. We were able to recruit 135 patients from among 184 eligible patients,. Among 135 consented patients, tissue could not be obtained from seven patients, and two patients had a final pathological diagnosis of dysplasia. These nine patients were excluded from the study.

Eligible controls were patients who had oral surgery for treatment of diseases other than cancer, such as obstructive sleep apnea, at the same institutions and during the same time period in which the OSCC cases were treated. In that time period, there were 45 eligible controls approached for participation by study staff, of whom 37 were recruited.

Each patient was interviewed using a structured questionnaire regarding demographic, medical, and lifestyle history, including tobacco and alcohol use. Data on tumor characteristics (site, stage) were obtained from medical records. Two cancer patients who enrolled in the study but did not consent to having their medical records reviewed were excluded from analysis. Thus, 124 cancer patients were included in this study. This research was conducted with written informed consent and institutional review board approval.

Tissue Collection

Tumor tissue was obtained at the time of resection from patients with primary OSCC prior to chemo/radiation therapy. Normal oral or oropharyngeal tissue was obtained from controls. One control provided two normal tissues and one cancer case had a large tumor that was divided into five pieces. Immediately after surgical removal, each tumor or normal tissue was soaked in RNALater (Applied Biosystems, Foster City, CA) for a minimum of 12 hours at 4 ° C and transferred to long term storage at – 80 ° C prior to use.

DNA Microarray

The DNA and RNA from each specimen were simultaneously extracted using the TRIzol method (Invitrogen, Carlsbad, CA). To increase DNA purity, we modified the DNA extraction protocol to include the use of a “back extraction buffer” (4 M guanidine thiocyanate, 50 mM sodium citrate, and 1 M Tris, pH 8.0). RNA was further purified with the use of an RNeasy mini kit (Qiagen, Valencia, CA) as per Affymetrix recommendations. For expression array analysis, 1.0 to 2.5 μg of total RNA was converted to double stranded cDNA using a GeneChip Expression 3’-Amplification One-cycle DNA Synthesis Kit (Affymetrix, Santa Clara, CA). The cDNA was purified and used in an *in vitro* transcription reaction to produce cRNA using the GeneChip Expression 3’-Amplification Reagents Kit (Affymetrix). The newly synthesized and biotin labeled cRNA was hybridized to a U133 2.0 Plus GeneChip (Affymetrix) and scanned using an Affymetrix GeneChip Scanner 3000 7G in the Fred Hutchinson Cancer Research Center’s Genomics Shared Resources as per Affymetrix protocols.

Quality Control (QC) of Microarray Results

We used the QC criteria specified by Affymetrix (http://www.affymetrix.com/support/downloads/manuals/data_analysis_fundamentals_manual.pdf) followed by the "affyQCReport" and "affyPLM" packages in Bioconductor (http://www.bioconductor.org) to search for poor quality GeneChips. These procedures identified seven GeneChips that did not pass QC tests (five from cancer patients and two from controls), and which were eliminated from further analyses. Thus 123 GeneChips from 119 cancer cases and 36 GeneChips from 35 controls were included in this analysis.

HPV Genotyping

We screened all samples for the presence of HPV DNA using a nested PCR based protocol [11]. All samples that had a positive PCR result and about 40% of the samples that had a negative result were tested for HPV DNA presence using the LINEAR ARRAY HPV Genotyping Test (Roche, Indianapolis, IN) under a research use only agreement. In brief, 50 to 75 ng of DNA were PCR-amplified as per kit protocols, followed by hybridization to strip arrays containing complementary sequences to the PCR products for 37 HPV genotypes (including the 13 "high risk" genotypes 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68) (Fig.1) and human β-globin (as an assay control). The results from the nested PCR protocol and the LINEAR ARRAY HPV Genotyping Test showed 100% corroboration. To further verify genotype calls by the linear array, we amplified a small subset of the samples using HPV type 16 specific primers, sequenced the amplified products, and compared them against a known HPV type 16 sequence (Genbank gi: 333031).

Data Analysis

Tumors were classified according to site as follows: oral cavity (including tongue, buccal mucosa, gingival, hard palate, retromolar trigone and floor of mouth) vs. oropharynx (including tonsils, soft palate, uvula, oropharynx and base of tongue).

Gene expression values for the ~54,000 probe sets were first extracted from probe intensity values (CEL files) using the gcRMA algorithm. We then eliminated the probe sets that either showed

no variation across the samples (inter-quartile range less than 0.1 on log2 scale) or that were expressed at very low magnitude (the maximum of the expression value across the samples is less than 3 on log 2 scale). These exclusions helped to limit the number of statistical tests applied when detecting differences between HPV-positive and HPV-negative tumors. After these two filtering processes, ~21,000 probe sets remained for further analysis.

Statistical tests were carried out to compare HPV-positive and HPV-negative OSCC using a regression framework implemented in GenePlus software (http://www.enodar.com/). To control for the type I error rate, we chose to declare a particular group of genes either "upregulated/overexpressed" or "downregulated/underexpressed" based on a pre-specified Number of False Discoveries (NFD)[12]. The choice of NFD, with an appropriate account for the number of genes under investigation (*J*), dictates the threshold for individual gene-specific p-values as NFD/*J*.

To determine whether the probe sets identified in the above analysis were up- or downregulated when compared to normal oral tissue, for each probe set we compared the mean expression values of each cancer group with those of controls using linear regression, calculating a robust estimator of variance, and accounting for the fact that multiple samples were tested for some subjects. The probe sets were then placed in order by ascending p-value, and a cutoff of 0.05 was chosen to indicate significant differences in expression.

The functional roles of the genes differentially expressed between HPV-positive and HPV-negative OSCC were assessed through the use of Ingenuity Pathways Analysis, IPA 5.0 (Ingenuity® Systems, www.ingenuity.com). The function analysis identified the biological functions by performing Fischer's exact tests to test the null hypothesis that the set of differentially expressed genes were not representative of each biological function.

Results

Patient characteristics

The characteristics of the study population overall and by HPV status are presented in Table 1. The cases were more likely to be older and to be current smokers, when compared with controls. More cases presented with oral cavity tumors than oropharyngeal tumors.

HPV detection in OSCC and control

HPV DNA was found in 41 of 119 (34.5%) tumor tissues from cases, and in 2 of 35 (5.7%) normal oral tissues from controls. Twenty-three of 31 oropharyngeal tumor tissues (74.2%) were HPV positive, whereas only 18 of 88 oral cavity tumor tissues (20.5%) contained HPV DNA. The great majority (39/43) of the HPV-positive samples contained only HPV 16. The remaining HPV-positive samples contained HPV 32, 35, 45, and 53. HPV types 32 and 53 (low risk type) were found in oral cavity cancers whereas HPV type 35 and 45 (high risk type) were found in oropharyngeal cancers. The tumor sample that contained HPV 32 was determined to be positive for HPV by the nested PCR test but negative by the Roche kit (which does not test for HPV 32). Sequencing demonstrated homology to HPV 32.

Genome-wide comparison between HPV-positive and HPV-negative OSCC

We used NFD = 1 as a selection criteria. This means that we control the number of false positive gene in the discover gene list to be fewer than 1. We did not find a significant difference in gene expression between HPV-positive and HPV-negative OSCC. When we analyzed oral cavity cancers and oropharyngeal cancers separately, we found no significant difference in gene expression between HPV-positive oral cavity cancers and HPV-negative oral cavity cancers, but we found 446 probe sets (Supplement 1, please see online version of manuscript) differentially expressed in HPV-positive oropharyngeal cancers compared to HPV-negative oropharyngeal cancers. This means that one probe set among 446 probe sets could be a false positive finding, corresponding to a false discovery rate (FDR) of 0.2%. The molecular and cellular functions of these genes that had the lowest p-values from the Ingenuity Pathways Analysis, IPA 5.0 were DNA replication, DNA recombination, DNA repair, and cell cycle (Table 2).

Comparisons to controls

In order to determine which, if any, of these 446 probe sets were up- or downregulated when compared to normal oral tissue, we compared HPV-positive oropharyngeal cancers and HPV-negative oropharyngeal cancers to HPV-negative normal oropharyngeal tissues from controls. Among 446 probe sets, 299 were significantly different between HPV-positive oropharyngeal cancers and HPV-negative oropharyngeal controls ($p < 0.05$), with 222 probe sets upregulated and 77 probe sets downregulated. Many of the genes involved in DNA replication, cell cycle and cell proliferation, such as *RPA2, LIG1, POLD1, POLH, MCM2, MCM3, MCM7, NASP, CDC7, CCNE2, CDKN2A, CDK2, RBBP4, PCNA*, and *Ki67,* were upregulated in HPV-positive oropharyngeal cancers. We also found upregulation of genes involved in DNA repair, such as *XRCC1, DDB2, FANCG* and *TOPBP1*. Cell cycle genes that were downregulated were *CCND1, APC* and *HIPK2*. The top 50 probe sets for upregulated and downregulated genes are presented in Table 3. A complete list of the 299 differentially expressed probe sets is in Supplement 2 (please see online version of manuscript). When comparing HPV-negative oropharyngeal cancers to HPV-negative oropharyngeal controls using these 446 probe sets, we found 79 upregulated probe sets and 122 downregulated probe sets. Table 4 lists the top 50 probe sets from this analysis. The list of 201 probe sets is shown in Supplement 3 (please see online version of manuscript).

There were 21 probe sets that were upregulated in HPV-positive oropharyngeal cancers but downregulated in HPV-negative oropharyngeal cancers, and four were downregulated in HPV-positive oropharyngeal cancers but upregulated in HPV-negative oropharyngeal cancers (Table 5).

Comment

Human Papillomaviruses (HPV) are small DNA viruses that are known to be associated with a subset of OSCC. We found significant differences in gene expression on the genome-wide level between HPV-positive and HPV-negative oropharyngeal cancers, but not in oral cavity cancers. That our results depend on tumor site is consistent with the large body of evidence that HPV are more frequently detected in, and more strongly associated with, the development of oropharyngeal cancers

than cancers in other head and neck sites [13]. Our results confirm the difference between HPV-positive and HPV-negative oropharyngeal cancers on the molecular level.

The top functions of genes that were differentially expressed between HPV-positive and HPV-negative oropharyngeal cancers based on the Ingenuity Pathway Analysis were DNA replication, DNA recombination, DNA repair, and cell cycle. HPV do not encode DNA or RNA polymerase but depend on the host cell's cell cycle control and replication machinery to enter S phase and they replicate along with host cell DNA. HPV drive cells into S phase through binding and inactivating Rb by their E7 oncoprotein and by displacing E2F [14]. We found upregulation of many cell cycle genes in HPV-positive oropharyngeal cancers, such as *CCNE2, E2F, CDC7* and *CDKN2A.* We also found upregulation of *PCNA* and *Ki67*, markers of cell proliferation, in HPV-positive oropharyngeal cancers. This finding is consistent with observations that HPV 16 enhances proliferation of an OSCC cell line [15].

Some of the genes that we found differentially expressed between HPV-positive and HPV-negative tumors have been reported to be associated with chemosensitivity to cisplatin, 5-FU, and paclitaxel, common chemotherapeutic agents used for treatment of head and neck squamous cell carcinoma (HNSCC). A study in breast cancer cell lines demonstrated that the cell line which was most sensitive to cisplatin expressed low levels of cyclin D1, and that cell lines transfected with cyclin D1 siRNAs exhibited enhanced sensitivity to cisplatin [16]. Furthermore, cisplatin can suppress E6 mRNA, restore p53 function and enhance radiosensitivity in HPV 16 E6 containing SiHa cells [17]. Since we found lower expression of *CCND1* in HPV-positive oropharyngeal cancers, it would be important to determine whether HPV-positive oropharyngeal cancers are more sensitive to cisplatin than HPV-negative oropharyngeal cancers. If so, cisplatin may be a good choice of adjuvant therapy for an HPV-positive oropharyngeal cancer patient.

Thymidine synthase (TS) is a target enzyme for 5-FU, and high expression of TS is related to poor response to 5-FU based chemotherapy [18]. In the present study, we found higher expression of *TS* (*TYMS*) in HPV-positive than in HPV-negative oropharyngeal cancers. Thus, HPV-positive oropharyngeal cancers may be more resistant to 5-FU chemotherapy.

STMN1, encodes stathmin, a protein involved in the regulation of microtubules. Two studies have shown that stathmin overexpression decreases sensitivity to paclitaxel *in vitro* [19, 20]. P53 regulates the G2/M check point by reducing expression of stathmin [21]. We found upregulation of *STMN1* in HPV-positive oropharyngeal cancers. We speculate that inactivation of p53 by high-risk HPV E6 may cause increased expression of stathmin, resulting in greater resistance to paclitaxel compared to HPV-negative oropharyngeal cancers. Further study is needed to confirm this.

Radiation is a common treatment choice for oropharyngeal cancer, because of the morbidity associated with surgical resection. Torres-Roca *et al*. demonstrated that radiosensitive cell lines had higher expression of RBBP4 than radioresistant cell lines, and transfection of RBBP4 into cell lines induced radiosensitization of these cell lines [22]. We found upregulation of *RBBP4* in HPV-positive oropharyngeal cancers. It would be interesting for future study to examine whether HPV-positive oropharyngeal cancers are more sensitive to radiation than HPV-negative oropharyngeal cancers.

Although the association between HPV infection and oropharyngeal cancer is well established, the clinical benefit of testing HPV in oropharyngeal cancer patients has not been established. Our results suggest the possibility of using HPV status for selecting personalized therapy for these patients. To translate these findings to patient management, clinical trials to evaluate the efficacy of cisplatin, 5-FU, paclitaxel, or radiation in the treatment of oropharyngeal cancers based on their HPV status are clearly warranted.

To the best of our knowledge, two previous studies have compared genome-wide gene expression between HPV-negative and HPV-positive HNSCC directly, using the same Affymetrix chip as the current study [23, 24]. Martinez, *et al*. were the only ones to specifically examine gene expression in oropharyngeal tissue [24] using three HPV-positive and four HPV-negative oropharyngeal cancers, and four normal oral mucosa tissues. They identified 124 upregulated and 42 downregulated genes in HPV-positive oropharyngeal cancers as compared to HPV-negative oropharyngeal cancers. Only three genes (*TYMS, TUBGCP3* and *SEC24D*) from their list overlapped with ours. A greater overlap was seen between the respective lists of genes that were differentially expressed between HPV-positive oropharyngeal cancers and normal tissues from controls. These genes included *CDKN2A, PCNA, RFC4, MCM2, MCM3, CDC7, TYMS, CCNE2, USP1*, and *ACTL6A*.

Slebos, *et al*. studied 36 HNSCC, including 15 oral cavity cancers, nine oropharyngeal cancers, nine laryngeal cancers and three hypopharyngeal cancers [23]. They detected HPV DNA in seven of nine oropharyngeal cancers but in none of the oral cavity cancers. They found 91 genes differentially expressed between HPV-positive and HPV-negative HNSCC. We found 20 genes overlapping among our results and theirs. These included *ACTL6A, ALG6, ASS1, CCDC52, CDC7, CDKN2A, CENPK, DHFR, EIF2B5, EZH2, LIG1, MCM2, MCM6, OPA1, RFC4, RPA2, STMN1, TOPBP1, USP1* and *WEE1*.

The difference between the results of these studies could be due to study design, different tumor sites and different approaches used for statistical analyses. Our analysis only included oropharyngeal cancers, whereas in the study by Slebos *et al*. one-third of the tumors were from other head and neck sites. To the extent that gene expression associations with HPV status differed in subsets of cases defined according to site, different results would be expected between our studies. Although both our and Martinez's studies focused on oropharyngeal tumors, our much larger sample size would be expected to detect larger numbers of differentially expressed genes.

Our study was limited by the small number of study subjects, particularly for comparisons among oropharyngeal cancer patients. With only eight HPV-negative oropharyngeal cancers, it was not feasible to include cigarette smoking, sex, and age as covariates when we compared genome-wide gene expression between the HPV-positive and HPV-negative oropharyngeal tumors.

In conclusion, we found differences in gene expression in HPV-positive and HPV-negative oropharyngeal cancers. These differences suggest that 1) HPV-positive oropharyngeal cancers may be more resistant to 5-FU and paclitaxel than HPV-negative oropharyngeal cancers, 2) cisplatin may be a better choice in treatment of HPV-positive oropharyngeal cancers, and 3) HPV-positive oropharyngeal cancers may also be more sensitive to radiation. Further study in cell lines and clinical trials are needed to investigate the possibility of using HPV as a guide for the management of oropharyngeal cancer.

References

(1) Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005 March;55(2):74-108.

(2) Syrjanen S. Human papillomavirus (HPV) in head and neck cancer. *J Clin Virol* 2005 March;32(Suppl 1):S59-S66.

(3) Schwartz SM, Daling JR, Doody DR et al. Oral cancer risk in relation to sexual history and evidence of human papillomavirus infection. *J Natl Cancer Inst* 1998 November 4;90(21):1626-36.

(4) Mork J, Lie AK, Glattre E et al. Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck. *N Engl J Med* 2001;344(15):1125-9.

(5) Smith EM, Ritchie JM, Summersgill KF et al. Human papillomavirus in oral exfoliated cells and risk of head and neck cancer. *J Natl Cancer Inst* 2004 March 17;96(6):449-55.

(6) D'Souza G, Kreimer AR, Viscidi R et al. Case-control study of human papillomavirus and oropharyngeal cancer. *N Engl J Med* 2007 May 10;356(19):1944-56.

(7) Gillison ML, Koch WM, Capone RB et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J Natl Cancer Inst* 2000 May 3;92(9):709-20.

(8) Ritchie JM, Smith EM, Summersgill KF et al. Human papillomavirus infection as a prognostic factor in carcinomas of the oral cavity and oropharynx. *Int J Cancer* 2003 April 10;104(3):336-44.

(9) Schwartz SR, Yueh B, McDougall JK, Daling JR, Schwartz SM. Human papillomavirus infection and survival in oral squamous cell carcinoma: a population based study. *Otolaryngol Head Neck Surg* 2001;125(1):1-9.

(10) Dahlgren L, Dahlstrand HM, Lindquist D et al. Human papillomavirus is more common in base of tongue than in mobile tongue cancer and is a favorable prognostic factor in base of tongue cancer patients. *Int J Cancer* 2004 December 20;112(6):1015-9.

(11) Sotlar K, Diemer D, Dethleffs A et al. Detection and typing of human papillomavirus by e6 nested multiplex PCR. *J Clin Microbiol* 2004 July;42(7):3176-84.

(12) Xu XL, Olson JM, Zhao LP. A regression-based method to identify differentially expressed genes in microarray time course studies and its application in an inducible Huntington's disease transgenic model. *Hum Mol Genet* 2002 August 15;11(17):1977-85.

(13) Kreimer AR, Clifford GM, Boyle P, Franceschi S. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. *Cancer Epidemiol Biomarkers Prev* 2005 February;14(2):467-75.

(14) Zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer* 2002 May;2(5):342-50.

(15) Kingsley K, Johnson D, O'Malley S. Transfection of oral squamous cell carcinoma with human papillomavirus-16 induces proliferative and morphological changes in vitro. *Cancer Cell Int* 2006;6:14.

(16) Yde CW, Issinger OG. Enhancing cisplatin sensitivity in MCF-7 human breast cancer cells by down-regulation of Bcl-2 and cyclin D1. *Int J Oncol* 2006 December;29(6):1397-404.

(17) Huang H, Huang SY, Chen TT, Chen JC, Chiou CL, Huang TM. Cisplatin restores p53 function and enhances the radiosensitivity in HPV16 E6 containing SiHa cells. *J Cell Biochem* 2004 March 1;91(4):756-65.

(18) Longley DB, Harkin DP, Johnston PG. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer* 2003 May;3(5):330-8.

(19) Alli E, Bash-Babula J, Yang JM, Hait WN. Effect of stathmin on the sensitivity to antimicrotubule drugs in human breast cancer. *Cancer Res* 2002 December 1;62(23):6864-9.

(20) Alli E, Yang JM, Ford JM, Hait WN. Reversal of stathmin-mediated resistance to paclitaxel and vinblastine in human breast carcinoma cells. *Mol Pharmacol* 2007 May;71(5):1233-40.

(21) Johnsen JI, Aurelio ON, Kwaja Z et al. p53-mediated negative regulation of stathmin/Op18 expression is associated with G(2)/M cell-cycle arrest. *Int J Cancer* 2000 December 1;88(5):685-91.

(22) Torres-Roca JF, Eschrich S, Zhao H et al. Prediction of radiation sensitivity using a gene expression classifier. *Cancer Res* 2005 August 15;65(16):7169-76.

(23) Slebos RJ, Yi Y, Ely K et al. Gene expression differences associated with human papillomavirus status in head and neck squamous cell carcinoma. *Clin Cancer Res* 2006 February 1;12(3 Pt 1):701-9.

(24) Martinez I, Wang J, Hobson KF, Ferris RL, Khan SA. Identification of differentially expressed genes in HPV-positive and HPV-negative oropharyngeal squamous cell carcinomas. *Eur J Cancer* 2007 January;43(2):415-32.

Legend Fig. 1. Use of Roche Linear Array to detect HPV DNA (subject Samples in lanes 1, 2, 4-6). Lane 3 was a negative control (no DNA)

(Please see separate attachment "CHEN Figure 1, 1-3-08")

Table 1. Selected characteristics of study population

Characteristics	Case No. (%)				Control No*. (%)	
	Oral cavity		Oropharynx			
	HPV-pos (n=18)	HPV-neg (n=70)	HPV-pos (n=23)	HPV-neg (n=8)	HPV-pos (n=2)	HPV-neg (n=33)
Age						
< 50	2 (11.1)	17 (24.3)	3 (13.0)	1 (12.5)	1 (50.0)	21 (63.6)
50-59	3 (16.7)	19 (27.1)	14 (60.9)	4 (50.0)	1 (50.0)	3 (9.1)
60-69	8 (44.4)	19 (27.1)	4 (17.4)	2 (25.0)	0 (0.0)	7 (21.2)
70+	5 (27.8)	15 (21.4)	2 (8.7)	1 (12.5)	0 (0.0)	2 (6.1)
Gender						
Male	11 (61.1)	45 (64.3)	23(100.0)	5 (62.5)	2 (100.0)	23 (69.7)
Female	7 (38.9)	25 (35.7)	0 (0.0)	3 (37.5)	0 (0.0)	10 (30.0)
Race						
Caucasian	16 (94.1)	62 (92.5)	21 (91.3)	7 (100.0)	2 (100.0)	22 (66.7)
Non-Caucasian	1 (5.9)	5 (7.5)	2 (8.7)	0 (0.0)	0 (0.0)	11 (33.3)
Unknown	1	3	0	1	0	0
Smoking status						
Never	4 (22.2)	19 (27.1)	1 (4.3)	1 (12.5)	1 (50.0)	14 (42.4)
Former	6 (33.3)	17 (24.3)	10 (43.5)	1 (12.5)	0 (0.0)	10 (30.3)
Current	8 (44.4)	34 (48.6)	12 (52.2)	6 (75.0)	1 (50.0)	9 (27.3)
Drinking status						
Never	1 (5.6)	8 (11.6)	2 (9.1)	0 (0.0)	0 (0.0)	2 (6.1)
Former	5 (27.8)	13 (18.8)	4 (18.2)	3 (37.5)	0 (0.0)	7 (21.2)
Current	12 (66.7)	48 (69.6)	16 (72.7)	5 (62.5)	2(100.0)	24 (72.7)
Unknown	0	1	1	0	0	0

*34 of the 35 control tissues came from the oropharynx

Table 2. Top molecular and cellular functions of genes differentially expressed between HPV + and HPV – oropharyngeal cancers by Ingenuity Software Analysis.

Function	p-value	Gene
DNA Replication, Recombination, and Repair	6.47 E-12	**CDKN2A, LIG1, MCM6, CDT1, PTTG1, PARP2, SMC4, CDC7, DDB2, SASS6, XRCC1, SP1, UBE2B, STK24, ZWINT, KIFC1, E2F2, ASF1B, NUSAP1, CCNE2, ESPL1, TOPBP, POLE2, FANCG, MDC1,** DGUOK**, HMGN1, POLD1, APC, RPA2, MCM3, BCL2L1, PCNA,MCM2, SMC2, HELLS,** RBMS1**, UBTF, CDK2, BUB3, MCM7, POLH**
Cell Cycle	4.78 E-08	**LIG1, CDKN2A, MLLT6, CDT1, PTTG1, CUL4A, PARP2, SMC4, CDC7, MKI67, SASS6,** CCND1, CAMK2D**, TNFSF5IP1, UBE2B, SP1, CAST, ZWINT,** HIPK2**, MFN2, E2F2, KIFC1, TYMS, ASF1B, CCNE2, NUSAP1, ESPL1, DBI, UHRF1, TOPBP1, WEE1, MDC1, HMGN1, APC, BCL2L1, PCNA, SMC2, MCM2, HELLS, BHLHB2,** IRS1**, MPHOSPH1, TCF19, CENPH, CDK2, POLH, MCM7, BUB3**
RNA Post-Transcriptional Modification	2.42 E-07	**LSM6, CDKN2A, HNRPH1, SFRS10, SNRPB, SFRS12, PHF5A, SFRS3,** SRPK2**, PTBP1, BAT1, PABPN1, SNRP70, FUSIP1,** RBMS1**, LSM4**
Cellular Assembly and Organization	2.86 E-06	**CDKN2A**, **PTTG1**, **PARP2**, **SMC4**, **CDC7**, **SH3GLB1**, **NCK1**, CND1, **HDAC6**, **STMN1**, **EZH2**, **FUSIP1**, **ZWINT**, **CAST**, HIPK2, **KIFC1**, **DNM2**, **RANBP9**, **ASF1B**, **ESPL1**, PAM, **TUBGCP3**, FMOD, **APC**, **BCL2L1**, **EBP**, **SMC2**, **HELLS**, **HMGN2**, **PARD3**, **CDK2**, **BUB3**, RAB1A, **CDT1**, **SASS6**, PALLD, **OPA1**, **MFN2**, **NUSAP1**, **UHRF1**, **MAP3K1**, **HMGN1**, **RPA2**, SEC22B, PMP22, **SS18L1**, CDC42BPA, **NUP153**, **MCM2**, **MPHOSPH1**, **UBTF**, **CXADR**, **RYK**, **CTNND1**
Nucleic Acid Metabolism	3.86 E-06	**LIG1**, **TYMS**, **PCNA**, PHYH, **DDB2**, DGUOK, **AK2**, **HMGN1**, **POLH**

The bold fonts denote genes that were upregulated in HPV-positive compared to HPV-negative oropharyngeal cancer.

Table 3. Top 50 genes upregulated and downregulated in HPV -positive oropharyngeal cancers compared to HPV-negative oropharyngeal controls.

Upregulated Genes			Downregulated Genes		
Probe Set ID	Gene Symbol	p-value	Probe Set ID	Gene Symbol	p-value
202107_s_at	MCM2	< 0.001	218432_at	FBXO3	< 0.001
204023_at	RFC4	< 0.001	204047_s_at	PHACTR2	< 0.001
209644_x_at	CDKN2A	< 0.001	204048_s_at	PHACTR2	< 0.001
218039_at	NUSAP1	< 0.001	202582_s_at	RANBP9	< 0.001
201202_at	PCNA	< 0.001	225872_at	SLC35F5	< 0.001
201970_s_at	NASP	< 0.001	223028_s_at	SNX9	< 0.001
201555_at	MCM3	< 0.001	218313_s_at	GALNT7	< 0.001
204252_at	CDK2	< 0.001	226338_at	TMEM55A	< 0.001
225655_at	UHRF1	< 0.001	208158_s_at	OSBPL1A	< 0.001
208795_s_at	MCM7	< 0.001	223315_at	NTN4	< 0.001
222740_at	ATAD2	< 0.001	226780_s_at	HSPC268	< 0.001
201930_at	MCM6	< 0.001	225116_at	HIPK2	< 0.001
212300_at	TXLNA	< 0.001	202780_at	OXCT1	< 0.001
221428_s_at	TBL1XR1	< 0.001	38158_at	ESPL1	< 0.001
236381_s_at	YIPF1	< 0.001	200972_at	TSPAN3	< 0.001
202666_s_at	ACTL6A	< 0.001	200973_s_at	TSPAN3	< 0.001
229551_x_at	ZNF367	< 0.001	229663_at	LONP2	< 0.001
210371_s_at	RBBP4	< 0.001	212586_at	CAST	< 0.001
205339_at	STIL	< 0.001	207480_s_at	MEIS2	< 0.001
222843_at	FIGNL1	< 0.001	222587_s_at	GALNT7	< 0.001
231846_at	FOXRED2	< 0.001	225523_at	MRPL53	< 0.001
205909_at	POLE2	< 0.001	211763_s_at	UBE2B	< 0.001
218115_at	ASF1B	< 0.001	203335_at	PHYH	< 0.001
230464_at	EDG8	< 0.001	221024_s_at	SLC2A10	< 0.001
223274_at	TCF19	< 0.001	227962_at	ACOX1	< 0.001
203379_at	RPS6KA1	< 0.001	218694_at	ARMCX1	< 0.001
225017_at	CCDC14	< 0.001	217979_at	TSPAN13	< 0.001
200775_s_at	HNRPK	< 0.001	213247_at	SVEP1	< 0.001
200783_s_at	STMN1	< 0.001	218328_at	COQ4	< 0.001
204825_at	MELK	< 0.001	213897_s_at	MRPL23	< 0.001
235609_at	---	< 0.001	215000_s_at	FEZ2	< 0.001
213253_at	SMC2	< 0.001	203525_s_at	APC	< 0.001
222848_at	CENPK	< 0.001	226886_at	---	< 0.001
212945_s_at	MGA	< 0.001	208670_s_at	EID1	< 0.001
209053_s_at	WHSC1	< 0.001	227274_at	---	< 0.001
205034_at	CCNE2	< 0.001	241789_at	---	< 0.001
223570_at	MCM10	< 0.001	225927_at	MAP3K1	< 0.001
204026_s_at	ZWINT	< 0.001	209090_s_at	SH3GLB1	< 0.001
228401_at	ATAD2	< 0.001	218946_at	NFU1	< 0.001
219698_s_at	METTL4	< 0.001	203227_s_at	TSPAN31	0.001
208672_s_at	SFRS3	< 0.001	1566557_at	FLJ90757	0.001
202412_s_at	USP1	< 0.001	216620_s_at	ARHGEF10	0.001
212021_s_at	MKI67	< 0.001	212508_at	MOAP1	0.001
228868_x_at	CDT1	< 0.001	212333_at	FAM98A	0.001
202413_s_at	USP1	< 0.001	208407_s_at	CTNND1	0.001
227350_at	HELLS	< 0.001	208669_s_at	EID1	0.001
200893_at	SFRS10	< 0.001	243463_s_at	RIT1	0.002
214172_x_at	RYK	< 0.001	219826_at	ZNF419	0.002
224468_s_at	C19orf48	< 0.001	210788_s_at	DHRS7	0.002
200892_s_at	SFRS10	< 0.001	238447_at	RBMS3	0.002

Table 4. Top 50 genes upregulated and downregulated in HPV-negative oropharyngeal cancers compared to HPV-negative oropharyngeal controls.

Upregulated Genes			Downregulated Genes		
Probe Set ID	Gene Symbol	p value	Probe Set ID	Gene Symbol	p value
205542_at	STEAP1	<0.001	211070_x_at	DBI	<0.001
214247_s_at	DKK3	<0.001	209389_x_at	DBI	<0.001
227628_at	LOC493869	<0.001	202428_x_at	DBI	<0.001
218718_at	PDGFC	<0.001	218432_at	FBXO3	<0.001
204686_at	IRS1	<0.001	203525_s_at	APC	<0.001
223934_at	LOC93349	<0.001	202582_s_at	RANBP9	<0.001
225685_at	---	<0.001	212586_at	CAST	<0.001
200790_at	ODC1	<0.001	38158_at	ESPL1	<0.001
209797_at	TMEM4	<0.001	225179_at	---	<0.001
228284_at	TLE1	<0.001	212435_at	TRIM33	<0.001
202375_at	SEC24D	<0.001	212622_at	TMEM41B	<0.001
218113_at	TMEM2	<0.001	227455_at	C6orf136	<0.001
203181_x_at	SRPK2	<0.001	212145_at	MRPS27	<0.001
202864_s_at	SP100	<0.001	218170_at	ISOC1	<0.001
212958_x_at	PAM	<0.001	219990_at	E2F8	<0.001
214662_at	WDR43	<0.001	225585_at	RAP2A	<0.001
202972_s_at	FAM13A1	<0.001	228910_at	---	<0.001
238317_x_at	RBMS1	<0.001	225927_at	MAP3K1	<0.001
226921_at	UBR1	<0.001	211763_s_at	UBE2B	<0.001
202336_s_at	PAM	<0.001	212710_at	CAMSAP1	<0.001
224759_s_at	C12orf23	<0.001	219433_at	BCOR	<0.001
212539_at	CHD1L	<0.001	209090_s_at	SH3GLB1	<0.001
228158_at	LOC645166	<0.001	202780_at	OXCT1	<0.001
1569110_x_at	LOC728613	<0.001	208407_s_at	CTNND1	<0.001
1556988_s_at	CHD1L	0.001	201519_at	TOMM70A	<0.001
226453_at	RNASEH2C	0.001	212623_at	TMEM41B	<0.001
223741_s_at	TTYH2	0.001	218467_at	TNFSF5IP1	<0.001
238935_at	RPS27L	0.002	202681_at	USP4	<0.001
216483_s_at	C19orf10	0.002	218657_at	RAPGEFL1	<0.001
204017_at	KDELR3	0.002	200980_s_at	PDHA1	<0.001
214620_x_at	PAM	0.002	227433_at	KIAA2018	<0.001
231793_s_at	CAMK2D	0.002	227962_at	ACOX1	<0.001
212685_s_at	TBL2	0.002	207168_s_at	H2AFY	<0.001
225655_at	UHRF1	0.003	218328_at	COQ4	<0.001
224619_at	CASC4	0.003	214501_s_at	TLR4	<0.001
228345_at	---	0.003	228361_at	E2F2	<0.001
200652_at	SSR2	0.004	235531_at	IL17RB	<0.001
218062_x_at	CDC42EP4	0.004	210418_s_at	IDH3B	<0.001
210788_s_at	DHRS7	0.004	212721_at	SFRS12	<0.001
227611_at	TARSL2	0.004	215732_s_at	DTX2	<0.001
203379_at	RPS6KA1	0.004	213852_at	RBM8A	<0.001
221739_at	C19orf10	0.004	201903_at	UQCRC1	<0.001
202107_s_at	MCM2	0.005	204048_s_at	PHACTR2	<0.001
218008_at	C7orf42	0.006	201036_s_at	HADH	<0.001
222270_at	SMEK2	0.006	226257_x_at	MRPS22	<0.001
222108_at	AMIGO2	0.007	224784_at	MLLT6	<0.001
205339_at	STIL	0.008	208854_s_at	STK24	<0.001
227326_at	MXRA7	0.008	204047_s_at	PHACTR2	<0.001
204023_at	RFC4	0.009	226316_at	---	<0.001
209549_s_at	DGUOK	0.010	224801_at	NDFIP2	<0.001

Table 5. Genes that have expression in opposite direction in HPV + and HPV - oropharyngeal cancers when compared to controls

Probe Set ID	Gene Title	Gene Symbol
gene upregulated in HPV-positive but downregulated in HPV-negative oropharyngeal cancers		
1555878_at	Ribosomal protein S24	RPS24
1559006_at	CDNA clone IMAGE:4304686	---
201677_at	Chromosome 3 open reading frame 37	C3orf37
201687_s_at	apoptosis inhibitor 5	API5
203017_s_at	synovial sarcoma, X breakpoint 2 interacting protein	SSX2IP
203409_at	damage-specific DNA binding protein 2, 48kDa	DDB2
207231_at	zinc finger DAZ interacting protein 3	DZIP3
212533_at	WEE1 homolog (S. pombe)	WEE1
213140_s_at	synovial sarcoma translocation gene on chromosome 18-like 1	SS18L1
213573_at	Full-length cDNA clone CS0DH006YD11 of T cells (Jurkat cell line) of Homo sapiens	---
215792_s_at	DnaJ (Hsp40) homolog, subfamily C, member 11	DNAJC11
219649_at	asparagine-linked glycosylation 6 homolog	ALG6
224754_at	Sp1 transcription factor	SP1
225340_s_at	GPI-anchored membrane protein 1	GPIAP1
225396_at	Zinc finger and BTB domain containing 8 opposite strand	ZBTB8OS
225725_at	CDNA clone IMAGE:5261213	---
226265_at	glutamine and serine rich 1	QSER1
227451_s_at	Coiled-coil domain containing 90A	CCDC90A
227545_at	Transcribed locus	---
228380_at	Transcribed locus	---
242655_at	Transcribed locus	---
gene upregulated in HPV-negative but downregulated in HPV-positive oropharyngeal cancers		
202305_s_at	fasciculation and elongation protein zeta 2 (zygin II)	FEZ2
210788_s_at	dehydrogenase/reductase (SDR family) member 7	DHRS7
211698_at	EP300 interacting inhibitor of differentiation 1	EID1
217047_s_at	family with sequence similarity 13, member A1	FAM13A1