Persistent HPV infection is causative for the majority of cervical cancer (CC) cases; however, current guidelines do not require HPV testing for newly diagnosed CC. Using an institutional cohort of 88 CC patients treated uniformly with standard-of-care chemoradiation (CRT) with prospectively collected clinical outcome data, we observed that patients with cervical tumors containing HPV genotypes other than HPV 16 have worse survival outcomes after CRT compared to patients with HPV 16 positive tumors, consistent with previously published studies. Using RNAseq analysis we quantified viral transcription efficiency and found higher levels of E6 and the alternative transcript E6*I in cervical tumors with HPV genotypes other than HPV 16. These findings were validated using whole transcriptome data from The Cancer Genome Atlas (n=304). For the first time transcript expression level of HPV E6*I was identified as a predictive biomarker of CRT outcome in our complete institutional dataset (n=88), and within the HPV 16 positive subset (n=36). In vitro characterization of HPV E6*I and E6 overexpression revealed that both induce CRT resistance through distinct p53-p21 dependent mechanisms. Our findings suggest that high expression of E6*I and E6 may represent novel biomarkers of CRT efficacy and these patients may benefit from alternative treatment strategies.
HPV transcript expression affects cervical cancer response to chemoradiation

Fiona J Ruiz¹,²
Matthew Inkman¹,³
Rashmi Ramachandran¹
Naoshad Muhammad¹
Nishanth Gabriel¹
Christopher A Miller⁴
Michael D McLellan⁴
Michael Goldstein¹
Stephanie Markovina¹,⁵
Perry Grigsby¹,⁵,⁶
Jin Zhang¹,³,⁵
Julie K Schwarz¹,⁵,⁷

¹ Department of Radiation Oncology, Washington University School of Medicine, St. Louis, MO, USA.
² Division of Biological and Biomedical Sciences Molecular Cell Biology, Washington University School of Medicine, St. Louis, MO, USA.
³ Institute for Informatics, Washington University School of Medicine, St. Louis, MO, USA.
⁴ McDonnell Genome Institute, Washington University School of Medicine, St. Louis, MO, USA.
⁵ Alvin J. Siteman Cancer Center, Washington University School of Medicine, St. Louis, MO, USA.
⁶ Division of Nuclear Medicine, Mallinckrodt Institute, Washington University School of Medicine, St. Louis, MO, USA.
⁷ Department of Cell Biology and Physiology, Washington University School of Medicine, St. Louis, MO, USA.

Running Title: HPV transcript expression affects CC response to CRT

Keywords: HPV, cervix, radiation

Corresponding Author: Julie Schwarz, MD, PhD
Address: Department of Radiation Oncology; Washington University School of Medicine, 4511 Forest Park Avenue, Saint Louis, MO 63108
Phone: 314-273-0275
FAX: 314-747-5498
Email: jschwarz@wustl.edu

Conflicts of Interest: The authors declare no potential conflicts of interest.

Text Word Count: 5793 words    Figures: 4    Tables: 2
Supplemental Files: 1    Supplemental Figures: 6    Supplemental Tables: 1
Abstract

Persistent HPV infection is causative for the majority of cervical cancer (CC) cases; however, current guidelines do not require HPV testing for newly diagnosed CC. Using an institutional cohort of 88 CC patients treated uniformly with standard-of-care chemoradiation (CRT) with prospectively collected clinical outcome data, we observed that patients with cervical tumors containing HPV genotypes other than HPV 16 have worse survival outcomes after CRT compared to patients with HPV 16 positive tumors, consistent with previously published studies. Using RNAseq analysis we quantified viral transcription efficiency and found higher levels of E6 and the alternative transcript E6*I in cervical tumors with HPV genotypes other than HPV 16. These findings were validated using whole transcriptome data from The Cancer Genome Atlas (n=304). For the first time transcript expression level of HPV E6*I was identified as a predictive biomarker of CRT outcome in our complete institutional dataset (n=88), and within the HPV 16 positive subset (n=36). In vitro characterization of HPV E6*I and E6 overexpression revealed that both induce CRT resistance through distinct p53-p21 dependent mechanisms. Our findings suggest that high expression of E6*I and E6 may represent novel biomarkers of CRT efficacy and these patients may benefit from alternative treatment strategies.
Introduction

Cervical cancer is the fourth leading cause of cancer incidence and mortality in women worldwide (1). In 2021 14,480 new cervical cancer diagnoses and 4,290 cervical cancer deaths are expected in the United States alone (2). Persistent HPV infection is causative for more than 91% of cervical cancer cases (3). Encoded within the HPV genome are oncogenes E6 and E7, which target p53, pRb and p130, an integral component of the pRb – related transcriptional repressor DREAM complex, respectively for degradation (4–7). Through E6 and E7, the virus induces cell immortalization resulting in a hyperproliferative state that facilitates viral replication. High-risk HPV genotypes use alternative splicing to switch between transcription of HPV E6 to E7 (8). Although a few of the high-risk HPV genotypes are able to encode multiple E6 splice variants, all of the high-risk HPV genotypes encode the E6*I splice variant (9). This primary alternatively spliced variant uses the first 5’ alternative splice site in the E6 open reading frame to generate a truncated E6 splice variant, E6*I (hereinafter referred to as E6*), and the full length E7 transcript (8). E6* has been previously reported to encode a functional protein that acts as an antagonist to full length HPV E6 (10, 11). HPV infection itself, however, is not sufficient to induce cervical cancer (12). The hyperproliferative state associated with E6 and E7 expression results in the development of cervical interepithelial neoplasms (CIN) 1 – 3 which can spontaneously regress. Over time, if HPV infection persists, the uncontrolled replication of the infected cell leads to the accumulation of mutations which can further deregulate and reprogram the cell leading to tumorigenesis.

Previous studies have demonstrated that HPV status and HPV genotype are both prognostic indicators in cervical cancer patients (13–18). Cervical cancer patients with undetectable lev-
els of HPV in their tumors have poor progression-free (PFS) and overall survival (OS) outcomes compared to patients with HPV positive tumors (13, 14). However, this HPV positive survival benefit is not binary, and patients with HPV genotypes other than HPV 16 in their tumors have worse PFS and OS after treatment than patients with HPV\textsuperscript{16+} tumors (15–18). This observation is also seen when HPV genotypes are grouped by viral clade, where patients with alpha 7 clade positive tumors, which includes HPV 18, have worse PFS and OS compared to patients with alpha 9 positive tumors, which includes the HPV 16 genotype (16, 18). Since HPV status has been shown to correlate with patient outcomes, it could be used as a diagnostic biomarker to identify patients who are likely to fail the current standard of care treatment strategy. However, the molecular basis of non-HPV 16 genotypes conferring worse prognosis is not well understood. Additionally, even within previously reported studies the HPV 16 patient subsets have PFS and OS below 75% indicating that within this clinically benefiting group there may be some patients who are likely to fail their standard of care treatment and would benefit from a more personalized treatment plan (15, 17). Therefore, identifying a viral diagnostic marker independent of HPV genotype may be beneficial. Currently it is unknown whether expression of HPV alternative transcripts affects tumor response to chemoradiation. The primary purpose of this study was to investigate whether variation in the viral transcriptome between HPV\textsuperscript{16+} and HPV\textsuperscript{Other} cervical cancer tumors correlated with poor prognosis after standard-of-care CRT. A secondary goal was to determine how HPV transcript expression, including the expression of alternatively spliced forms of HPV E6, affects cellular response to standard of care chemotherapy and radiation treatment.
Results

HPV genotypes detected in cohort

In our 88-patient cohort, 87.5% of patients had detectable HPV DNA and patients with undetectable HPV DNA were excluded from the study. Patients with non-HPV 16 genotypes, or multiple HPV genotypes detected were grouped together as HPV_{Other}. Forty-eight patients had HPV_{16+} tumors and 29 were HPV_{Other}; 9 HPV-18, 4 HPV-33, 3 HPV-45, 3 HPV-52, 3 HPV-59, 1 HPV-31, 1 HPV-56, 1 HPV-58, 1-HPV 68a, 1 HPV-82 and two patients with multiple HPV genotypes detected (Table 1). Initial analysis focused on confirming whether prognostic differences based upon HPV genotype, clade and viral integration observed in our institutional cohort was consistent with previously published studies. Patients with HPV_{16+} tumors had better PFS (p = 0.049) and OS (p = 0.004) compared to HPV_{Other} patients (Figure 1). Similarly, patients with alpha 9 positive tumors (n = 57) had better PFS (p = 0.012) and OS (p = 0.084) compared to those with alpha 7 clade (n =16) (Supplementary Figure 1 A-B). Viral integration state was assessed using RNAseq analysis and integration events occurred in 70.3% of HPV_{16+} and 85.7% of HPV_{Other} tumors (Supplementary Figure 1E). Patients with integrated HPV had less favorable PFS (p = 0.067) and OS (p = 0.057) compared to patients with episomal HPV, although these results were not statistically significant (Supplementary Figure 1C-D).

There were no significant differences in patient age at time of diagnosis, tumor staging, lymph node involvement, and metastasis at the time of diagnosis between HPV_{16+} and HPV_{Other} genotype patient groups (Table 2). Additionally, both patient groups had comparable radiation treatment intent and treatment completion (Table 2). Despite there being no difference in clinical characteristics or treatment strategies, patients with HPV_{16+} tumors had more favorable outcomes.
after chemoradiation compared to patients in the HPV\textsuperscript{Other} group. Patients with HPV\textsuperscript{Other} cervix tumors experienced more disease recurrence (p = 0.032) and increased mortality (p = 0.018) after chemoradiation treatment (Table 2).

**Comparison of viral transcription efficiency between HPV 16 and HPV other genotypes**

RNA from 68 patient biopsies was isolated and whole transcriptome sequencing, including all host and viral transcripts, was performed. Viral expression of E6, E6* and E7 was quantified and compared between HPV\textsuperscript{16+} (n = 36) and HPV\textsuperscript{Other} (n = 17) patient groups. The HPV E6* and E7 transcripts were more highly expressed than HPV E6 in both HPV\textsuperscript{16+} and HPV\textsuperscript{Other} groups (Figure 2A). HPV\textsuperscript{Other} tumors had an average of 2.9-fold increase in HPV E6 (p = 0.0085) and 2.0-fold increase in E6* (p = 0.025) transcript expression compared to HPV\textsuperscript{16+} tumors. There was an average 1.5-fold increase in HPV E7 transcript expression of HPV\textsuperscript{Other} compared to HPV\textsuperscript{16+} tumors, however this was not statistically significant. (Figure 2A).

We used publicly available TCGA cervical cancer data as a validation cohort for HPV transcript expression. Similar to our cohort, 56% of tumor samples were HPV\textsuperscript{16+} and 43% were HPV\textsuperscript{Other} in the cervix TCGA cohort (Supplementary Figure 2A). There was no significant difference in viral integration state between the HPV genotype groups (Supp. Fig. 2B). The HPV\textsuperscript{Other} patient group had an average fold increase of 3.0 for HPV E6 (p = 8.1e-12), 2.2 for E6* (p = 2.4e-9 ), and 1.4 for E7 (p = 0.0012)(Supp. Fig 2C).

The relative expression levels of HPV viral transcripts were tested as predictive biomarkers for CRT outcome using prospectively collected clinical outcome data from our study’s uniformly treated patient cohort. HPV E6, E6* and E7 transcript expression levels were grouped as
high (RPKM > upper quartile) and low (RPKM < upper quartile). High HPV E6* expression was a poor prognostic indicator for PFS (p = 0.047) and OS (p = 0.033) after CRT (Figure 2B-C). Interestingly, expression of HPV E6 or E7 alone was insufficient to significantly stratify patient survival outcomes (Supplementary Figure 3). To determine whether increased expression of HPV E6* could be used as an independent prognostic biomarker, we also analyzed the HPV16+ subset of patients alone (n = 36). HPV16+ patients with high HPV E6* expression again exhibited worse PFS (p = 0.38) and OS (p = 0.33) (Figure 2D-E) although this did not reach the level of statistical significance.

**HPV transcript expression correlates in vitro with RT sensitivity**

To test our hypothesis in vitro, cervical cancer cell lines, CaSki, SiHa and SW756 were evaluated for their relative transcript expression of HPV E6, E6* and E7 by qRT-PCR. CaSki (HPV 16) had higher levels of E6, E6* and E7 transcript expression compared to SiHa (HPV 16) and SW756 (HPV 18) cells (Supplementary Fig 4A). Next, we determined the sensitivity to radiation in these cervical cancer lines. The CaSki cell line with high viral transcript levels was more resistant to increasing doses of radiation compared to SiHa (4Gy: p = 0.032) and SW756 (4Gy: p = 0.017) cell lines with low transcript levels (Supplementary Fig 4B). Taken together, these results suggest that high viral transcript expression correlates with resistance to radiation in cervical cancer cells, consistent with our clinical observations.

Next we wanted to remove any potentially confounding effects of HPV genotype and test whether alterations to HPV viral transcript expressions alone could induce resistance to CRT. The low HPV transcript expressing SiHa cell line was selected to evaluate the effect of increased
E6 and E6* expression on cell sensitivity to CRT. HPV 16 E6 and E6* protein coding sequences were ligated into pCMV-6 mammalian expression vectors and transfected in SiHa cells, and primers specific to the vector HPV 16 E6 and E6* sequences were used to quantify vector specific transcript expression by qRT-PCR (Supplementary Fig 5A-B). E6 overexpression in SiHa cells resulted in a 26% reduction in baseline p53 levels as expected (p=0.0107). E6* overexpression in SiHa cells increased p53 protein expression compared to WT SiHa (FC = 1.35, p=0.0107) (Figure 3A-B), consistent with previously published findings that E6* is an antagonist of E6-mediated p53 degradation (10).

**E6 and E6* expression affects cellular response to chemoradiation treatment**

We assessed the effect of E6 and E6* overexpression in SiHa cells on survival following CRT treatment. Both, E6* and E6 overexpression decreased cell sensitivity to increasing doses of radiation treatment (Figure 3 C-D). Increased expression of E6* or E6 did not affect cell cycle progression following CRT treatment (Supplementary Figure 6 A-F). We found that apoptosis was not the primary mechanism of cell death in the cervical cancer lines following CRT, although all cell lines were able to induce apoptotic cell death following etoposide treatment (Supplementary Figure 6 G-J). Kinetics of DNA damage induction and repair were determined by monitoring the formation and resolution of γH2AX foci after radiation. WT as well as E6 and E6* overexpressing SiHa cells showed similar levels of γH2AX foci 10min after exposure to radiation. 24h following irradiation γH2AX were resolved in E6 expressing SiHa cells (p = 0.017) cells whereas WT cells and cells expressing E6* had residual γH2AX foci suggestive of unrepaired DNA breaks in the latter two lines (Figure 3 E-F).

**E6* overexpression induces p21 mediated cellular senescence**
To test whether the function of p53 as a transcription factor is affected by viral transcript expression, qRT-PCR was performed for six p53 target genes. The SiHa E6* overexpressing cell line had higher *CDKN1A* (*p* = 0.021) and *NOXA* (*p* = 0.0089) transcript expression, and the SiHa E6 overexpressing cell line had lower *PUMA* (*p* = 0.045) transcript expression compared to the WT SiHa cell line (Figure 4A, Supplementary Figure 5C-D). *BAX, GADD45A* and *MDM2* transcript expression was not affected by either E6 or E6* overexpression (Supplementary Fig 5 E-G). *CDKN1A* encodes the p21 protein which is a direct target of p53 regulation. SiHa E6* cells had 3.9-fold higher p21 protein expression compared to WT SiHa (*p* = 0.0028) (Figure 3A). The functionality of p21 was assessed in the three cell lines including cellular proliferation and induction of cellular senescence. The p53 and p21 low SiHa E6 cell line had the fastest rate of proliferation followed by WT SiHa. SiHa E6* overexpressing had the lowest proliferation rate (Figure 4B). High p53 and p21 expression can induce cellular senescence (19–21). To test whether increased p21 was associated with increased senescence in E6* expressing cervical cancer, SiHa (parent), SiHa E6* and SiHa E6 overexpressing cells were plated and grown for 8 days and then stained for SA-β-galactosidase expression and senescent (SA-β-gal positive) cells were quantified. SiHa cells overexpressing E6* had more SA-β-gal positive cells compared to WT SiHa (*p* = 0.029) and SiHa E6 cell lines (*p* = 0.057) (Figure 4 C-D. There was no difference in SA-β-gal positivity between SiHa E6 and WT SiHa cells (*p* = 0.2). Increase in p21 is known to induce G1 cell cycle arrest (22), however neither E6* or E6 overexpression induced a G1 arrest (Supplementary Figure 5H-I). Additionally previous reports have implicated that an increase in p21 levels leads to transcriptional repression of cell cycle genes due to an interaction with the RB and
DREAM complex (23). To test this 5 DREAM target genes were selected, and their expression was evaluated across the cell lines, however no significant alteration was found except for a decrease in KIF23 which was observed in both the p21 high E6* (p = 0.007) and p21 low E6 (p = 0.01) expressing cell lines (Supplementary Figure 5J). Both the parental SiHa cell line and the engineered SiHa E6* and SiHa E6 cell lines all have HPV E7 expression intact, and E7 expression has been previously reported to abrogate both p21 dependent G1 cell cycle arrest and the p53-p21-DREAM pathway (24, 25).

**p21 knockdown sufficient to sensitize SiHa E6* to radiation treatment**

Lastly we tested whether knockdown of p21 is sufficient to rescue treatment sensitivity in the p21 high E6* overexpressing cell line (Supplementary Figure 5 K-L). Transient knockdown of p21 in the parental SiHa and E6 overexpressing cell lines, which are low in p21 expression, did not have a significant impact on sensitivity to radiation treatment. In the p21 high E6* overexpressing cell line, transient knockdown of p21 led to an increase in cell sensitivity to increasing doses of radiation treatment (Figure 4E). Altogether these findings indicate that overexpression of the p53 agonist E6* stabilizes p53 expression, leading to upregulation of p21 and subsequent induction of cellular senescence and overall reduced sensitivity to chemoradiation treatment. This phenotype can be rescued by knockdown of p21 prior to radiation treatment.
Discussion

Treatment for cervical cancer is currently guided by clinical staging, which includes the results of physical examination and whole-body imaging. In general, early-stage cervical cancers are managed by surgical resection while locally advanced cases are treated with pelvic irradiation and the concurrent administration of cisplatin chemotherapy. While this treatment strategy is effective, 30-50% of patients with locally advanced cervical cancer will fail CRT, necessitating the need to develop predictive biomarkers to assist in treatment planning. Current clinical guidelines do not require HPV testing of newly diagnosed cervical cancer. In this study we demonstrate that expression of an HPV genotype other than HPV 16 is a predictive biomarker for a poor disease response to the standard of care chemoradiation treatment using a population of patients that were uniformly treated with standard of care CRT with prospectively collected clinical outcomes. Furthermore, using RNASeq, we demonstrated that cervical tumors with non-HPV 16 genotypes had higher viral transcript expression of HPV E6 and E6*I. High HPV E6*I expression was significantly associated with patient risk of developing disease recurrence and was a predictive biomarker of poor disease response following CRT treatment. Interestingly, the prognostic value of high HPV E6*I expression was also observed when restricting analysis to the HPV 16 patient subset, indicating that it may be a valuable genotype independent predictive biomarker. These results suggest that more detailed analysis of HPV genotype and HPV related gene expression could be used to identify cervical cancer patients at risk for treatment failure from standard of care chemoradiation.

These survival differences affected by the HPV types are not exclusive to cervical cancer. Approximately 20% of oropharyngeal squamous cell cancer (OPSCCs) are HPV related, and pa-
tients with HPV related OPSCC have favorable outcomes after chemoradiation (26, 27). More recently, HPV genotype differences are being appreciated in HPV positive HNSCC where HPV 18 genotype has been identified as an independent poor prognostic factor of overall survival (28). In addition, expression of alternative transcripts such as HPV E1-E4 has been identified as a candidate biomarker for treatment resistance in HPV positive head and neck cancer (28). Based upon the results found in our study with prospectively collected tumor specimens and patient outcomes, as well as other previous studies, HPV genotype should be considered a new predictive biomarker for outcome after chemoradiation. Furthermore, the association between HPV genotype and the expression of HPV related genes, including full length E6 and the alternatively spliced variant E6*I merit further study.

In this study we focused on interrogating the relationship between HPV genotypes and viral transcript levels. Additionally, previous studies have shown that integrated HPV virus is a poor prognostic indicator in cervical cancer patients, although the biological mechanism to explain this clinical observation is poorly understood (29, 30). In our institutional cohort 77% of HPV\textsuperscript{16+} patients had integrated whereas 83% of patients with HPV\textsuperscript{Other} tumors had viral integration (Supplementary Figure 1A). Similarly, in the cervix TCGA cohort 70% of HPV\textsuperscript{16+} patient tumors had viral integration compared to 86% of HPV\textsuperscript{Other} tumors (Supplementary Figure 2A). For both independent patient cohorts, HPV\textsuperscript{Other} tumors tended to have a higher incidence of viral integration compared to patients with HPV\textsuperscript{16+} tumors. Additionally, in our institutional cohort patients with HPV integrated tumors had worse PFS (p = 0.067) and OS (p = 0.057) after CRT; however, perhaps due to the small cohort size these survival differences did not meet the threshold for significance (Supplementary Figure 1 B-C). Factors regulating HPV integration into the
host genome are poorly understood. Host genome sites where HPV integration frequently occurs have been reported throughout the genome and are found within fragile sequence sites and areas of open chromatin (31). HPV integration into the host genome can result in transcription of viral-host fusion proteins and influence transcriptional activity of neighboring host genes via a pseudo-promoter effect (32). Using both our institutional cohort and validating with the cervix TCGA cohort we found that viral integration state, also correlated with HPV transcript expression (Supplementary Figure 1D and 2D). As previously mentioned, for both our institutional and TCGA cervix cohorts HPV\textsuperscript{Other} tumors had a higher incidence of viral integration and the episomal group is primarily composed of patients with HPV\textsuperscript{16+} tumors (Supplementary Figure 1A and 2B). Taken together this raises the question as to whether non-HPV 16 genotypes are more efficient in integrating their viral genome into the host genome, and perhaps this is why these patient tumors have higher viral transcript efficiency; however, additional in vitro work will need to be done to determine the relationship between HPV genotype, viral state and transcript expression.

In vitro characterization of high HPV E6* and E6 transcript expression revealed that increasing expression of either transcript was sufficient to induce CRT treatment resistance. Each transcript induced treatment resistance through different mechanisms, but both involved differences in E6 related regulation of p53 and downstream p21 activation. Increasing full length E6 expression inhibited p53 and p21 activity, resulting in enhanced efficiency of DNA repair of RT induced lesions as evidenced by Figure 3E-F. In contrast, high E6* expression inhibited E6 induced degradation of p53, resulting in p53 stabilization and downstream p21 activation leading to induction of cellular senescence as shown in Figure 3A and 4C-D. Furthermore, we demonstrated that the decrease in treatment sensitivity observed in the high E6* expression cell line
was dependent on the increase in p21 expression and could be sensitized to treatment following p21 knockdown (Figure 4E). Work is ongoing in our lab to use this model in vivo and investigate the pairing of novel drug therapies to sensitize high viral transcript expressing tumors to CRT.

Cellular senescence is a state in which cells no longer proliferate but are still viable. These cells alter the tumor microenvironment because of their senescence associated secretory profile (SASP) and can reprogram cells of the microenvironment to a tumor favorable phenotype (33). Senescent cells are non-proliferative and therefore have an intrinsic resistance to radiation treatment. However, an emerging group of chemotherapies are being developed to target and kill senescent cells (34). This new chemotherapy class should be tested for its efficacy in treating E6* overexpressing HPV transformed cell lines to determine whether targeting senescent cells will lead to better treatment efficacy following CRT treatment (35, 36).

The enhanced DNA damage repair we observed in high E6 expressing cells should be further investigated to determine the mechanism of repair being affected by HPV E6 expression. In Figure 3 E-F, we show that HPV E6 overexpressing cells had enhanced p-γH2AX foci clearance after radiation treatment compared to wild type and E6* overexpressing cells. HPV E6 targets p53 for degradation and, as we show in Figures 3A and 4A, this decrease in p53 degradation coincides with a decrease in p21 transcript and protein expression. One of the functions of p21 is to inhibit PCNA, which is an essential component of DNA damage repair (37, 38). HPV E7 has been previously shown to upregulate PCNA expression (39). In our in vitro system, the SiHa E6 overexpressing cells express endogenous HPV 16 E7, therefore the enhanced DNA repair that we find in this cell line following radiation treatment may be due to the synergy between inhibition of p21 by HPV E6 and the upregulation of PCNA by HPV E7, however future studies will be
required to identify the DNA damage repair pathway being affected by this alteration. Patients with high E6 expression could benefit from pairing radiation treatment with DNA damage response inhibitor targeting the pathway being affected by E6 overexpression.

In conclusion, using a population of patients uniformly treated with standard of care CRT with prospectively collected clinical outcome data, this study confirms HPV genotype and identifies for the first time HPV transcriptional efficiency as predictive biomarkers of poor patient outcomes after standard of care chemoradiation in cervical cancer. We validated the association between HPV genotype and viral transcript expression efficiency using the cervix TCGA cohort which also demonstrated that HPV 16 positive tumors indeed have lower transcriptional efficiency of HPV E6, E6*I and E7. Using cervical cancer cell lines engineered to overexpress E6 and E6*I from HPV 16, we demonstrate that transcript variants of HPV E6 are associated with treatment resistance to standard of care CRT. While overexpression of full-length HPV E6 is associated with enhanced DNA repair, overexpression of HPV E6*I is associated with tumor cell senescence. These results support more detailed examination of HPV genotype and viral transcript expression in patients with locally advanced cervical cancer and suggest that this information may be used in the future to identify patients at risk of CRT failure. In addition, these results provide important insights into treatment resistance mechanisms for HPV positive tumors.
Materials and Methods

Patients

Patient tumor biopsies and blood were obtained prior to the initiation of therapy and stored at the Tissue Procurement Facility at Washington University in St. Louis School of Medicine. Radiotherapy (RT) consisted of external beam irradiation and intracavitary brachytherapy per institutional guidelines, and concurrent cisplatin or carboplatin chemotherapy was also administered. Median follow-up time for patients alive at the time of last follow-up was 62 months (range: 25-92 months). Inclusion in downstream data analysis required detectable HPV in patient tumors. In the reporting of this data we have adhered to REMARK guidelines (40).

Clinical follow-up

Patients were followed at the following timepoints after the completion of CRT: 6 weeks, 3 months, 6 months, 12 months, 18 months, 2 years, 3 years and 5 years. Physical examination including a pelvic exam was performed at each visit. Patients received post treatment $^{18}$F-Flourodeoxy-glucose positron emission tomography (FDG-PET) scans at 3 months post-therapy to assess treatment response.

Statistics

Tumor recurrence and overall survival (OS) were the primary endpoints of the study. Survival outcomes were measured from the completion of treatment. Recurrence was defined as new or progressive disease both locally within the pelvis and distantly outside the RT field. The Kaplan-Meier method and log rank test were used to determine differences in progression-free and overall survival. R version 3.5.2 and the packages survminer and survival were used for the analysis. Wilcoxon signed-rank test was used to compare patient age between HPV cohorts. Chi-
squared and Fisher exact tests were used for all other patient categorical variables. The Wilcoxon signed-rank test and student t-tests were used for in vitro experiments. One-way ANOVA was used when comparing more than 2 groups. $P$ less than 0.05 was set as the threshold for significance for all study outcomes.

**Study Approvals**

The study population included 88 patients prospectively enrolled into a tumor-banking protocol (January 2008 through July 2011). This study was approved by the Institutional Review Board and all patients provided written informed consent for sequencing (201105374). RNAseq data can be accessed at Gene Expression Omnibus (GEO) with accession number GSE151666.

**HPV genotyping**

DNA isolated from 88 pre-treatment tumors were sequenced using next-generation sequencing, and probes for the following HPV genotypes were used to identify the HPV genotypes present within the tumor: high-risk HPV-16, -18, -31, -33-, 35, -39, -45, -51, -52 -56, -58, and -59 genotypes and low-risk HPV-6b, -11, -26, -30, -53, -54, -57, -61, -66, -68a, -70, -72, -73, -82, -98, -99, -100, -104, -105, and -113. SAMtools idxstats (41) was used to extract number of reads mapping to each HPV accession, and HPV positivity was defined as any sample with at least 100 HPV reads detected for one of the 32 genotypes. Patient groups were defined as patients with HPV 16 detected (HPV$^{16+}$), non-HPV 16 genotypes detected (HPV$^{\text{Other}}$) or no HPV detected in the sample (HPV undetected). Patients with undetectable HPV were excluded from study analysis.

**Viral transcriptome and integration state**
RNA was extracted, with 68 patients yielding at least the 0.5µg of RNA with RIN > 7 sufficient for library preparation. RNAseq was performed for these patient samples using the Illumina HiSeq 3000. Raw sequencing reads were aligned to reference human and HPV viral genomes in GRCh38.d1.vd1 using STAR 2.7.0f with chimSegmentMin set to 18 (42). STAR-generated chimeric junction files were parsed for reads that were chimeric between human and HPV references and the corresponding alignment information and nucleotide sequences were extracted and manually inspected using BLAT and BLAST to verify viral integration. To obtain read counts for HPV E6, E6* and E7, sequences were aligned against HPV reference genomes for both the E6 and E6*I alternative splice using BWA 0.7.17-r1188. Numbers of junction spanning reads for wildtype E6 ($N_{E6,j}$) and E6* ($N_{E6*,j}$) were counted, as well as the number of reads aligned to E6 between the splice sites ($N_{E6,b}$). The remaining $N_a$ reads align to the ranges over which E6 and E6* are identical. Total counts of E6 ($N_{E6}$) and E6* ($N_{E6*}$) reads were therefore obtained by assigning these reads to wildtype or E6* transcripts based on the proportions of each type of junction spanning read normalized by the size of the nt ranges from which such a read could arise, as in the following formulas:

$$N_{E6} = \frac{N_{E6,j}}{N_{E6,j} + 2N_{E6*,j}} N_a + N_{E6,j} + N_{E6,b}$$

$$N_{E6*} = \frac{N_{E6*,j}}{2N_{E6*,j}} N_a + N_{E6*,j}$$

Counts of reads aligned to HPV E6, E6*I and E7 were then normalized by reads per kilobase of transcript per million mapped reads (RPKM).

**TCGA viral transcriptome and integration state**
Viral gene expression and integration status was obtained for 304 TCGA-CESC primary tumor samples using a similar method to that employed for our institutional data. RNAseq reads were extracted and aligned to reference human and HPV viral genomes in GRCh38.d1.vd1 using STAR 2.7.0f with chimSegmentMin set to 18. Reads that were chimeric between human and HPV references were extracted, and samples with greater than one such chimera per million total reads were designated as positive for HPV viral integration. Normalized HPV E6, E6* and E7 expression levels were also obtained as above.

**Engineering E6 and E6* overexpressing cell lines**

Cervical cancer cell lines were obtained from the American Type Culture collection (ATCC), and maintained in IMDM media (Life Technologies, CA) with 10% heat inactivated FBS and incubated at 37°C in 5% CO₂. Mycoplasma testing was performed periodically to verify no infection. Last date of mycoplasma testing for cell lines used in this study was July 17th, 2019 for SiHa and January 22nd, 2020 for CaSki and SW756. Experiments were performed on cell lines under passage 30. CaSki, SiHa and SW756 were used to evaluate endogenous expression of HPV E6, E6* and E7 transcripts and determine cell sensitivity to radiation treatment by alamarBlue (Thermo Fisher Scientific, Waltham, MA, USA). SiHa had the lowest transcript expression for all HPV transcripts evaluated and was selected to engineer cell lines for the overexpression of HPV 16 E6* and E6. The mammalian expression vector pCMV-6 was used as the vector for exogenous HPV viral transcript expression. Origene VC101902 HPV 16 E6 ORF was used as the codon optimized sequence for HPV 16 E6 minigene (Integrated DNA Technologies, Coralville, IA, USA) and customized for HPV 16 E6* with an early stop codon inserted at cDNA position 151 (G > T) (Origene, Rockville, MD, USA), sequences provided in Supplementary File 1. HPV
16 E6 and E6* ORFs were restriction enzyme digested with 10U/µl sgfl and 10U/µl mlul, ligated into pCMV-6 vectors using T4 DNA ligase, and transformed into 5-alpha competent E.coli (New England Biosciences, (NEB)) overnight and then selected using 25µg/ml kanamycin LB agar plates. Midiprep of plasmids was performed using ZymoPure II Plasmid Midiprep Kit (ZymoPure) according to manufacturer’s protocol. Purified vectors were transfected into SiHa cells using Lipofectamine 3000 (Thermo Fisher Scientific, Waltham, MA, USA). Transformed cells were G418 selected and resulting cell pools were validated using qRT-PCR for vector E6* and E6 sequence expression.

**Western blot**

Cells were lysed with Cell Lysis Buffer (Cell Signaling Technology, Danvers, MA, USA), proteinase/phosphatase inhibitors and PMSF on ice for 30min and then sonicated using the Bioruptor UCD-200 (Diagenode, Liege, Belgium) on high 5min. NuPAGE LDS Sample Buffer (4x) (Invitrogen) and SDS buffer (2x) were added, samples were boiled at 95°C for 10min and gel electrophoresed on 4-20% gradient gels (Mini-Protean TGX, Bio-Rad, Hercules, CA, USA), transferred to PVDF blot, and blocked with 5% TBST milk. Blots were incubated with primary antibodies at 4°C overnight; anti-p21 (1:1000 CST 2947) (Cell Signaling Technology, Danvers, MA, USA). HRP-conjugated secondary antibodies were incubated 1h at room temperature; anti-rabbit (1:4000, sc-2357), P53-HRP (1:1000, sc-126), and Actin-HRP (1:5000 sc-47778) (Santa Cruz Biotechnology, Dallas, TX, USA), were used with ECL chemiluminescent reagent (GE Healthcare Life Sciences, Pittsburgh, PA, USA). Blots were visualized and quantified using the Bio-Rad ChemiDoc MP imaging system and Image Lab software (Bio-Rad, Hercules, CA, USA).
**qRT-PCR**

Total RNA was isolated using the Trizol reagent according to manufacturer’s protocol (Sigma). 1\(\mu\)g of total RNA was used to synthesize cDNA using High Capacity RNA-to-cDNA (4387406) (Thermo Fisher Scientific, Waltham, MA, USA) kit per manufacturer’s protocol. Primers for HPV 16 and 18 native and vector sequences were synthesized by Integrated DNA Technologies, from reference sequences (8, 43–45) or de novo (Supplementary Table 1). Taqman probes were used for the p53 target genes; *CDKN1A, NOXA, PUMA, BAX, GADD45A*, and *MDM2* (Thermo Fisher Scientific, Waltham, MA, USA). Probes for the DREAM target genes; *BIRC5, CDC25c, KIF23, MYBL2*, and *PLK4* were obtained from (Thermo Fisher Scientific, Waltham, MA, USA). *GAPDH* and *ACTB* were used as reference genes. The qRT-PCR was run at 50°C for 2min, 95°C for 10min, 60°C for 1min using the 7900HT Fast Real-Time PCR System (Applied Biosystems, Carlsbad, CA, USA). The specificity of the reaction was verified by melt curve analysis. Data processing was performed using SDS 2.4 software (Applied Biosystems) and the relative quantitation of each mRNA was performed using the comparative C\(_T\) method as described earlier (46).

**Senescence Associated beta-galactosidase (SA-\(\beta\)-Gal) assay**

Cells were plated and grown in culture for 7-days, harvested and stained using the Senescence \(\beta\)-Galactosidase Staining Kit (#9860) (Cell Signaling Technology, Danvers, MA, USA) per manufacturers protocol. SA-\(\beta\)-gal positive cells were quantified as number of cells per image frame with brightfield microscopy, in non-overlapping regions.

**Proliferation, cell viability and clonogenic survival**
To assess cellular proliferation, cells were plated at a density of 0.4 x 10^6 cells per well and counted 3-, 7-, 9-, and 12-days post-plating on Vi-CELL cell counter (Beckman Coulter, Pasadena, CA, USA). For both cell viability and clonogenic survival, cells were treated with vehicle (0.01% DMSO) or 0.5µM carboplatin 1h prior to radiation treatment (2 and 4Gy). Cell viability was assessed 5 days post-treatment using alamarBlue (DAL1025)(Thermo Fisher Scientific, Waltham, MA, USA) per manufacturers protocol; viability was normalized to untreated samples. For colony formation assay, cells were treated as described, 48h post-treatment cells were detached with 0.05% Trypsin and re-plated at 500 cells/well in a 6-well dish. Cells grew for 10-12 days, until untreated control wells reached colonies of > 50 cells. Plating efficiency (PE) was calculated as (number of colonies in control) / (500 cells seeded) and surviving fraction (SF) was calculated as (number of colonies in treated wells) / PE*(500 cells seeded). Cell lines were normalized to their own untreated control wells. For p21 knockdown colony formation assay, subconfluent (60-70%) cells were transfected with 100 nM of p21 siRNA or control siRNA (Santa Cruz Biotechnology) using Lipofectamine RNAiMAX reagent (Invitrogen Life Technologies, CA, USA) in opti-MEM as per manufacturer’s protocol. After 48 h of siRNA transfection, transfected cells were irradiated and then processed as described.

**Immunofluorescence**

SiHa, E6 and E6* cells were seeded on chamber slides (Nunc Lab-Tek, Rochester, NY, USA) and treated with single-fraction 2Gy RT. Slides were fixed with 4% paraformaldehyde at 10m, 2h, 6h, 12h and 24h post-irradiation. Cells were permeabilized and processed for immunofluorescence according to standard protocol. Slides were incubated with primary antibody 1:300 anti-phospho-Histone H2AX (Ser139) (20E3) (Cell Signaling Technology, Danvers, MA, USA)
at 4°C overnight. Secondary antibody 1:500 anti-rabbit Alexa Fluor 488 (ThermoFisher Scientific, Waltham, MA, USA) incubated 2h at room temperature. Slides were mounted using ProLong Gold Antifade Reagent with DAPI (ThermoFisher Scientific, Waltham, MA, USA). Slides were imaged with the Axio (Carl Zeiss) microscope.

**Flow Cytometry**

Cell cycle and apoptosis were assayed using flow cytometry. For the following assays, cells were plated and then next day treated with 0.5µM carboplatin plus 4Gy radiation, cells were harvested at 24 and 48hrs following treatment. For cell cycle analysis, cells were washed and then stained with propidium iodide. For apoptosis, cells were washed then stained with Annexin-V and propidium iodide. Flow cytometry analysis was performed on MacQuant Analyzer and quantified on FlowJo. For each sample 5,000 events were collected.

**Author Contributions**

SM, PG, and JKS contributed by writing the human studies protocol, enrolling patients, obtaining cervical tumor biopsies, prescribing and performing CRT and prospectively collecting clinical outcome data. MM and CM performed HPV genotyping on the Wash U cohort tumor samples. FR, MI, JZ, SM, NG, MG, RR, NM and JKS conceptualized the study, designed experiments and performed analyses. FR, MI, and JZ performed data analysis of HPV genotypes, transcript expression and integration status using Wash U and TCGA patient data. FR, NM and NG performed in vitro experiments. FR and JKS prepared the manuscript. All authors contributed to final review of the manuscript.

**Acknowledgements:**
This work was supported by NIH R01CA181745 to JK Schwarz. FR was supported by T32 CA113275.
References


19. Roninson IB. Tumor Cell Senescence in Cancer Treatment 1. In: 2003:


**Figure 1: HPV genotype significantly stratifies patient outcomes:**
A) Progression-free and B) overall survival curves stratified by HPV genotype (pink = HPV 16, brown = HPV other, grey = HPV undetected). Log rank test was used to determine statistical significance, calculated using the survminer package in R version 3.5.2.
Figure 2: HPV E6, E6*I and E7 transcript expression in WUSM cohort
A) Comparison of HPV transcript expression between HPV 16 (pink) and non-HPV 16 positive (brown) patient tumors (* p < 0.05, ** p < 0.01; non-parametric Wilcoxon- signed rank test). The full WUSM cohort (B-C) and the HPV 16 positive patient subset (D-E) were stratified based upon their relative HPV E6*I transcript expression, where high > Q3 and low < Q3, and assessed for progression-free (B,D) and overall survival (C,E). Significance calculated by log rank test using the survminer package in R version 3.5.2.
Figure 3: E6* and E6 overexpression effect on CRT response
SiHa E6 and E6* expressing cell lines were quantified for A) p53 and p21 protein expression by western blot (Representative blots, lanes run on the same gel but were not contiguous). Protein analysis was independently replicated 3 times. B) Band quantification of p53 protein expression was performed on 3 independent blots (* p < 0.05, Wilcoxon-ranked sign test). C-D) Clonogenic survival was assessed by colony formation assay. C) Representative images of colony formation assay shown. D) Normalized surviving fraction of SiHa parental (black) vs SiHa E6* (blue) and SiHa E6 (brown) mean ± SEM. (C-D) Representative experiment shown with technical replicates n = 3 for each condition, one-way ANOVA test used, Experiment independently repeated 3 times. E-F) DNA damage response was quantified using γH2AX p-S139 foci positive cells. E) Representative images of γH2AX foci (green) and nuclei stained by DAPI (blue) immunofluorescence of SiHa parental, SiHa E6* and SiHa E6 cell lines. F) Quantification of γH2AX foci positive cells (>3 foci/nuclei) at 10m, 2h, 6h, 12h, and 24h post-2 Gy radiation treatment (Each condition had 4-10 random images quantified, data representative of two independent experiments, one-way ANOVA).
Figure 4: HPV E6* and E6 overexpression stabilizes p53 and p21 stabilization leading to induction of cellular senescence

CDKN1A (p21) expression was quantified using A) qRT-PCR (normalized to GAPDH and ACTB transcript expression) (n = 6 biological replicates used for each, student's t-test). B) Proliferation assay of SiHa parental (black) vs SiHa E6* (blue) vs SiHa E6 (brown). (Data representative from 3 independent experiments, mean ± SEM, one-way ANOVA). C-D) Representative images of SA-β-Galactosidase staining (arrow = SA-β-Gal positive) and quantification of SA-β-Gal positive cells. (Representative experiment shown from 25 random field images obtained per sample, Wilcoxon-ranked sign test * p < 0.05, ** p < 0.01. Experiment was independently replicated 2 times.) E) Normalized surviving fraction of each SiHa cell line treated with control siRNA vs a p21 targeting siRNA mean ± SEM. (Representative experiment shown with technical replicates n = 3 for each condition, Student’s t-test used. Experiment independently repeated 3 times).
Table 1: HPV Genotypes

<table>
<thead>
<tr>
<th>HPV Genotype</th>
<th>(n = 88)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV 16</td>
<td>48</td>
</tr>
<tr>
<td>HPV 18</td>
<td>8</td>
</tr>
<tr>
<td>HPV 16 and 18</td>
<td>1</td>
</tr>
<tr>
<td>HPV Other</td>
<td>18</td>
</tr>
<tr>
<td>HPV 31</td>
<td>1</td>
</tr>
<tr>
<td>HPV 33</td>
<td>4</td>
</tr>
<tr>
<td>HPV 45</td>
<td>3</td>
</tr>
<tr>
<td>HPV 52</td>
<td>3</td>
</tr>
<tr>
<td>HPV 56</td>
<td>1</td>
</tr>
<tr>
<td>HPV 58</td>
<td>1</td>
</tr>
<tr>
<td>HPV 59</td>
<td>3</td>
</tr>
<tr>
<td>HPV 68a</td>
<td>1</td>
</tr>
<tr>
<td>HPV 82</td>
<td>1</td>
</tr>
<tr>
<td>&gt; 10 HPV types</td>
<td>2</td>
</tr>
<tr>
<td>HPV 16 + others</td>
<td>1</td>
</tr>
<tr>
<td>HPV 18 + others</td>
<td>1</td>
</tr>
<tr>
<td>HPV Negative</td>
<td>11</td>
</tr>
</tbody>
</table>
### Table 2: Patient baseline characteristics

<table>
<thead>
<tr>
<th></th>
<th>HPV 16 (n = 48)</th>
<th>HPV other (n = 29)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age at Diagnosis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(median, range)</td>
<td>51.5 (25-79)</td>
<td>53 (38-85)</td>
<td>0.235</td>
</tr>
<tr>
<td><strong>Stage</strong></td>
<td></td>
<td></td>
<td>0.468</td>
</tr>
<tr>
<td>Ib1</td>
<td>8</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Ib2</td>
<td>9</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IIA</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IIb</td>
<td>15</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>IIIa</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IIIb</td>
<td>12</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>IVa</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IVb</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Lymph Nodes at Diagnosis</strong></td>
<td>24</td>
<td>15</td>
<td>0.873</td>
</tr>
<tr>
<td>Aortic</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Pelvic</td>
<td>19</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>SCV</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Metastasis</strong></td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Bone</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Radiation Treatment Intent</strong></td>
<td></td>
<td></td>
<td>0.371</td>
</tr>
<tr>
<td>Curative</td>
<td>45</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Palliative</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Completed Treatment</strong></td>
<td>46</td>
<td>26</td>
<td>0.29</td>
</tr>
<tr>
<td>Recurrence</td>
<td>16</td>
<td>18</td>
<td>0.015</td>
</tr>
<tr>
<td>Pelvic</td>
<td>4</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Distant</td>
<td>8</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>No data</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Overall Survival</strong></td>
<td></td>
<td></td>
<td>0.035</td>
</tr>
<tr>
<td>Alive</td>
<td>32</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Dead</td>
<td>16</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>

Wilcoxon rank sum test used for Age  
Fisher exact test used for all other characteristics