Human Papillomavirus (HPV) Viral Proteins Substitute for the Impact of Somatic Mutations by Affecting Cancer-Related Genes: Meta-analysis and Perspectives

Zisheng Shang¹, Valentina L. Kouznetsova²,³, Igor F Tsigelny²,⁵*

¹MAP Program UC San Diego, La Jolla, California, USA
²Moores Cancer Center UC San Diego, La Jolla, California, USA
³San Diego Supercomputer Center, UC San Diego, La Jolla, California, USA
⁴Department of Neurosciences, UC San Diego, La Jolla, California, USA
⁵CureMatch Inc. San Diego, California, USA

Article Info

Article Notes
Received: April 24, 2020
Accepted: June 11, 2020

*Correspondence:
Igor F. Tsigelny, UC San Diego, San Diego, California, USA; ORCID- 0000-0002-7155-8947; Email: itsigeln@ucsd.edu.

© 2020 Tsigelny IF. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License.

Keywords
HPV
Cervical cancer
Head and neck cancer
Somatic mutations
Viral proteins
Gene networks

Abstract

Purpose: Although a strong association between human papillomavirus (HPV) and a variety of cancers has long been established, infection by HPV alone has been shown to be insufficient for the induction of cancer, with a large number of HPV infections regression without causing cancer. Additionally, HPV-negative cases have worse prognosis rates than HPV-positive ones across a multitude of cancer types. The reasons behind these phenomena are poorly understood. We try to explain it.

Methods: In this meta-analysis study, we first compared the mutation sets of HPV-positive and HPV-negative cancers using data from the Cancer Genome Atlas (TCGA) and from open published sources. Then, focusing on the genes that are exclusively mutated in HPV-negative head and neck cancer, we used KEGG pathways and reviewed literature to investigate how HPV can mimic the effects of mutations in those genes and thereby trigger carcinogenesis despite the absence of those mutations.

Results: Using TCGA mutation frequencies of significantly mutated genes stratified by HPV status, we extracted 17 genes exclusively mutated in HPV-negative cancers and detailed the HPV infection mechanisms that are largely able to mimic the effects of mutations in 13 of these genes. Through our literature searches and pathway analyses, we also presented antitumor effects caused by HPV infection that reduce carcinogenic efficiency and could contribute to the prognosis difference.

Conclusion: Our findings confirm that HPV infection substitutes for some, but not all, somatic driver mutations required for the induction of cancer by mimicking the effects of those mutations through various virus–host interaction mechanisms, thereby accelerating the cancer-related pathways.

Introduction

The human papillomavirus (HPV) is a DNA virus that infects the skin and mucus membranes and has long been identified as a common etiological agent for many types of cancers in those types of tissues¹. In fact, it is associated with almost 100% of cervical cancers, 90% of anal cancers, a significant percentage of head-and-neck squamous cell carcinomas, 40% of vulva, vagina, and penile cancers, 12% of pharyngeal cancers, and 3% of oral cancers¹. It has also recently been shown to be associated with bladder carcinomas²,³. Among all the subsites of head-and-neck cancer (which include the larynx, oral cavity, hypopharynx, oropharynx and the sinonasal tract), cancers in the oropharynx presently show the most pronounced link to HPV infection⁴,⁵.
At least 200 types of HPV have been identified, but only a few are truly oncogenic and are classified as high-risk types of HPV, which include the well-known HPV-16 and HPV-18 because of their high association with squamous cell carcinomas. Additionally, among the multiple viral proteins that HPV encodes, E6 and E7 stand out as the main viral proteins that support carcinogenesis, as they drive cellular proliferation, block apoptosis and inhibit the cell cycle arrest mechanisms. These mechanisms either accelerate the rate of cell division or inhibit the normal regulations on the cell-cycle-progression system, thus leading to unchecked cell growth, which may lead to cancer.

However, despite HPV’s high prevalence in squamous cell carcinomas, study results have shown that HPV infection alone is not sufficient for the induction of cancer. In fact, the majority of HPV infections are asymptomatic and regress: about 70% of new HPV infections resolve spontaneously in one year and 90% naturally resolve in two years. Progression to cancer is a rare event: only in a small percentage of cases do these viral infections persist and cause epithelial lesions, which increase the risks of cancer development but are still not definite indicators of cancer. So, why are these outcomes of HPV infection drastically different from each other? What factors besides HPV infection contribute to these differences?

Another interesting phenomenon exists in the difference in prognosis between HPV-positive and HPV-negative cancers. For oropharyngeal cancers, compared to HPV-negative patients, HPV-positive patients tend to be younger, with less exposure to tobacco and alcohol, and there is a higher incidence of HPV-positivity for males than females. Across a multitude of tumor sites, HPV-positive cancers are associated with younger patient ages and are shown to confer better survival rates. Many previous studies have correlated the better survival rates with younger ages and less tobacco and alcohol exposure, dismissing the role of HPV as an important prognostic factor. However, there are studies that contradict this assumption, finding that the prognosis differences still exist after adjusting for the age, smoking and drinking prognostic factors.

Overall, to explain the substantial difference between possible HPV infection outcomes and the observation that HPV-positive cancers often occur earlier in life than negative ones, we hypothesize that HPV infection mimics the effects of some somatic driver mutations by targeting a set of genes that play key roles in various carcinogenic pathways. Note that this set of genes might differ from the normally mutated driver genes in HPV-negative cases. This hypothesis implies that HPV infection has to be accompanied by the appropriate somatic driver mutations in order to support cancer development, while the mere infection itself is insufficient, because HPV doesn’t mimic the effects of all the mutations necessary to support cancer development. Such mimicking mechanism reduces the number of driver mutations required for the induction of cancer for HPV+ patients, because HPV infection can substitute the long somatic mutation accumulation process. We also hypothesize that this mimicking effect is not as powerful as the carcinogenic effects of a necessary set of random somatic driver mutations, possibly explaining the observation that HPV-positive status often confers better prognosis in HPV-associated cancers.

We explored the possibility of our hypothesis mainly by investigating the differences in necessary mutation sets between HPV-positive and HPV-negative cases and by analyzing the effects of HPV on genes that are exclusively mutated in HPV-negative cases.

Materials and Methods

KEGG Pathway Database

Short for the Kyoto Encyclopedia of Genes & Genomes, the KEGG database contains pathway diagrams that graphically illustrate many biochemical pathways. The pathways are all constructed based on existing reliable literature. In this study, we mainly used the KEGG pathways titled “Pathways in Cancer—Homo Sapiens” and “Human Papillomavirus Infection” to guide our investigation of mutation effects and HPV viral effects.

The Cancer Genome Atlas (TCGA)

TCGA, started in 2006, is a landmark cancer genomics program that catalogued the genetic mutations responsible for a variety of cancer types. In this study, we used head-and-neck squamous cell carcinoma mutation frequency data that is stratified by HPV status from The Cancer Genome Atlas. This data serves as the basis for our investigations, as it provides the exclusively mutated genes in HPV-negative cases.

Results

HPV-negative cases require more somatic driver mutations than HPV-positive cases

In cancers associated with HPV, HPV-negative cases generally occur with the presence of significantly more somatic driver mutations than HPV-positive ones. Using gene mutation frequency data stratified by HPV status from The Cancer Genome Atlas, we have constructed a table to compare the mutation frequencies of significantly mutated genes in HPV-positive and HPV-negative cases of head-and-neck cancer (Figure 1; data used is summarized in Table 1). The genes in the dataset play key roles in the cell death, immunity, differentiation, and oxidative stress signaling pathways including Ras/Raf and Akt/Pi3K. To extract the “validated” somatic mutations, a threshold line is placed at 3% mutation frequency, as suggested by Ma et
After taking this threshold, we noticed that there were 17 genes exclusively mutated in HPV-negative cases (and not mutated in HPV-positive cases), while only 4 genes were exclusively mutated in HPV-positive cases (with 0% mutation frequency in HPV-negative cases). This 17/4 ratio may demonstrate that more driver mutations are needed to induce cancer in HPV-negative cases than HPV-positive cases. As we previously proposed, this difference could be due to HPV proteins activating or inhibiting the specific genes, substituting for the random somatic mutation accumulation process.

In addition to the scenario in head-and-neck cancer mentioned above, we noticed a similar situation occurring in cervical cancer. According to a study conducted by Banister and co-authors that analyzes cervical cancer data from The Cancer Genome Atlas, "HPV-active tumors had on average 115 somatic mutations per tumor, whereas
HPV-inactive tumors had 228 somatic mutations per tumor. Furthermore, by comparing the somatic mutation frequencies of 26 key cervical cancer driver genes between HPV-positive and HPV-negative cases using two-tailed Fisher's Exact p-values and obtaining the odds ratios, Banister and colleagues showed that all 26 genes had different mutation rates across the different types of cases, and that all of the genes were more frequently mutated in HPV-negative cases (Figure 2), as the odds ratios of somatic mutations in HPV-negative vs. HPV-positive tumors are greater than 1 for all listed genes. These findings further support the notion that HPV-negative cancers require more somatic driver mutations than HPV-positive ones, across different HPV-associated cancer types.

We also note that in a study conducted by Iranzo and co-authors to estimate the average number of driver mutations needed for various types of cancers, the number of driver mutations needed for the induction of cervical cancer is only about 1, with a range from about 0.1 to 1.8. This surprisingly small number probably accounts for the extremely high HPV-prevalence in cervical cancers. Thus, corresponding to our proposed idea that the required number of driver mutations in HPV-positive patients are generally less than the number in HPV-negative patients, the fact that almost all cervical cancers are supported by HPV makes the average number of driver mutations surprisingly small.

**Prognosis Stratified by Age Group & HPV Status**

As mentioned above, HPV-positive status generally confers better prognosis than HPV-negative status in a wide range of HPV-associated cancers, including HNSCC, in which the presence of HPV generally leads to lower risks of dying, lower risks of recurrence and better response to therapy. Previous interesting hypothesis addressing this scenario has associated such prognosis outcome with the ages of HPV-positive and HPV-negative patients. As Larsson and colleagues corrected the overall and recurrence-free survival rates for age and tumor size, they discovered that the presence of HPV doesn't confer a survival advantage anymore and concluded that HPV isn't an independent prognostic factor in the diagnosis of cancer. This hypothesis seems reasonable, as many studies showed that HPV-positive cancers generally occur earlier in life than HPV-negative cancers, and this younger age for HPV-positive cancers would confer better prognosis.

However, there exist published data that calls into question the validity of this hypothesis. As shown in Figure 3(A), within each age group, the HPV-positive 5-year survival rate is better than that of HPV-negative cases for cervical cancer. In Figure 3(B), the hazard ratios (HRs) show the risk of death for HPV-positive patients compared with HPV-negative patients of oropharyngeal cancer, which is a common type of head and neck cancer. Both unadjusted and adjusted HRs are shown, and adjusted
HPV infection partially mimics the somatic driver mutation accumulation process

We propose that the significant difference between the number of somatic driver mutations required for the induction of cancer in HPV-positive and -negative cases is accounted for by the HPV viral infection mechanisms. In the absence of some necessary driver mutations, HPV proteins target a set of genes that activate or inhibit certain cancer pathways to support cancer development, mimicking the effects of those somatic driver mutations. This mimicking mechanism could only be partial and thus could account for the prognosis difference between HPV-positive and -negative cancers. In other words, the HPV infection might not be as effective in inducing cancer as the driver mutations.

To investigate and evaluate the proposal above, we focused on the genes that were exclusively mutated in HPV-negative cases after the threshold. We analyzed KEGG pathways16 and performed public sources analysis to see how HPV infection mechanisms could substitute for these mutations and cause cancer in their absence. We also analyzed the differences between the effects of mutation and infection to look for specific factors that could contribute to the prognosis difference between HPV-positive and -negative groups. According to Figure 1 and Table 1, there are 17 of the exclusively mutated genes: TP53, CDKN2A, FAT1, CCND1, MYC (c-MYC), NFE2L2, CASP8, FGFR1, CDK6, BIRC2 (cIAP), AJUBA, CUL3, EPHA2, ERBB2, KEAP1, HRAS, and IGF1R.

**TP53**

TP53 is the gene coding for the P53 protein, the mutation of which is present in a wide variety of tumor cells. Most TP53 mutations are missense mutations that occur in the central sequence-specific DNA-binding domain and inactivate the tumor suppressor's transcriptional regulation functions31. Such inactivation and loss of function affect a significant number of P53 transcriptional targets including cell cycle inhibitors like P21, apoptosis related proteins such as BAX, and DNA-repair proteins like GADD45 (Figure 4A)32-34.

These effects of mutation-induced inactivation of P53 can be mimicked by HPV proteins through targeting of a set of genes ultimately linked to P53 or the P53 pathway. First of all, the E6 viral oncoprotein can bind with E6AP, and this E6/E6AP complex directly associates with P53. This association allows E6AP to catalyze the transfer of ubiquitin to P53, thereby activating the proteasome-mediated degradation of this tumor-suppressing protein. Thus, E6 inhibits P53’s ability to transcriptionally activate a number of proteins playing key roles in cell cycle regulation. For example, it inhibits the activation of P21, preventing P21 from causing G1 cell cycle arrest through its...
A. Pathway diagram of p53 mutation-induced carcinogenic effects. B. Pathway diagram of HPV-cancer-supporting effects that focus around p53. While HPV is able to mimic the p53 transcriptional inhibition induced by somatic mutations, it is unable to mimic the gain-of-function of oncogenes that mutations confer, as it doesn’t seem to be able to promote metastasis. Additionally, HPV’s inhibitory effects on p53 is counteracted by the E7 viral protein through E7’s inhibition of MDM2-p53 interaction. C. Key for interpreting the pathway diagram. This key will apply to all the pathway diagrams included in this paper.
binding and inhibition of CDK2. This interaction between HPV E6 and P53 is widely studied over the last several decades and is common to HPV infections in a variety of cancer types35,36. Secondly, the E6 oncoprotein can bind to and inhibit p300/CBP, which normally activates P53-dependent transcription. Thus, this inhibition can repress the transcriptional activity of P5337. However, this binding and inhibition has been noted by Zimmermann et al.37 to be limited to E6 proteins of high-risk HPVs associated with cervical cancer. Finally, in cervical infections, HPV E6 inactivates TADA3 function, abrogating TADA3’s role as a coactivator for P53-dependent transcriptional activation and disrupting the cell cycle arrest mechanism following DNA damage38.

Independent from its abnormal transcriptional functions, mutant P53 induces a series of oncogenic gain-of-function mechanisms, which are functions independent from and unrelated to the wild type P53 protein. Studies on Li-Fraumeni syndrome patients, who are subject to the effects of mutant P53, reveal that mutant P53 can increase genomic instability through its disruption of normal spindle checkpoint control, generating polyploidy cells39. It has also been shown by studies of myeloid leukemic cells that mutant P53 can confer antiapoptotic advantages to tumor cells, rendering cells significantly more resistant to cMyc and a number of other anticancer agents40,41. In addition, in studies performed on non-small cell lung carcinoma cells, with the cooperation of oncogenic Ras, mutant P53 activates the expression of a large set of chemokines and interleukins that have been shown to induce metastasis and angiogenesis42,43. While the gain-of-function effects of mutant P53 are observed in the above-stated non-HPV-associated cancers, mutant P53 could carry out similar effects in HPV-associated cancer types, and HPV seems to be unable to mimic all these gain-of-function effects of oncogenes induced by mutant P53. Thus, the mutant P53 in HPV-negative cases could have increased oncogenic effects when compared to P53 that is altered by HPV.

On the other hand, the E7 HPV protein, although commonly known for its carcinogenic effects when interacting with the retinoblastoma protein, produces antitumor effects with its interactions with wild type P53. In a study using mouse embryo fibroblasts, Seavey and co-authors showed that the E7 viral protein increased the level of P53 protein by inhibiting the interaction between P53 and MDM2, so that MDM2 cannot degrade P53 anymore. Such an inhibition minimizes the effects of E7/PP2A and E5/V-ATPase interactions, which were supposed to degrade P53 through the activation of MDM2. Specifically, without its inhibitory function, E7 would cause the degradation of P53 by directly binding to PP2A to displace its B subunit. This binding would leave PP2A unable to dephosphorylate and inhibit Akt, which is one of PP2A’s targets. As a result, high levels of Akt would enhance the activation of MDM2, which monoubiquitinites P53 and thus mediates P53’s degradation by proteasomes45,46. The overactivation of MDM2 would then result in degradation of P53 and thus cell-cycle dysregulation. However, this carcinogenic effect is inhibited by E7’s inhibition of MDM2–P53 interactions. Also, without this inhibitory function of E7, HPV E5 would use MDM2 to trigger P53 degradation. In transfected human keratinocytes, E5 has been shown to interact with the subunit C (a structural component) of V-ATPase to impair the enzyme’s functions47. Such an impairment triggers a series of events in the PI3K-Akt signaling pathway (Figure 4B) and ultimately leads to overactivation of MDM248. However, the effect is seriously mitigated by E7’s inhibition of MDM2-P53 interaction.

With the self-contradicting effects produced by this E7 function that inhibits MDM2–P53 interactions, we speculate that the effects of HPV E6 and E7 on P53 somewhat compensate for each other: As HPV E6 suppresses P53, E7 enhances P53 activity, so the effects of the two viral proteins contradict each other, possibly reducing the amount of repression P53 receives and thus conferring some survival advantage to HPV-positive cancer patients when compared to those HPV-negative, driver-mutation-based cancer patients. This speculation is supported by the fact observed in some studies that the expression of P53 as detected through immunohistochemistry is significantly higher in HPV-positive cancers than in HPV-negative ones for cancers of the upper respiratory tract and in laryngeal tumors49,50. The study results of Gillison and co-authors, and Maruyama and colleagues also support our speculations, as they showed an inverse relationship between HPV-presence and P53 mutation in oropharyngeal carcinomas, with HPV-positive tumors less likely to harbor P53 mutations than HPV-negative ones. However, there have also been experimental findings contradictory to this hypothesis, with Wang and colleagues stating that HPV-positive breast cancer tissues showed significantly lower expression of P53 and Gualberto and colleagues confirming that result in metastatic cancers of unknown primary in the head-and-neck region. Our speculation also contradicts previous theories that HPV-negative mutation-induced tumor cells accumulate excessive amounts of P53, as the frequently occurring missense mutations may render the P53 transcription factor resistant to proteolytic degradation by E3 ubiquitin ligases such as MDM2, maintaining high levels of the mutant P53 protein. If such theories were true, we would expect HPV-negative cancers to have higher levels of P53 expression than HPV-positive ones, and we would not see the HPV-positive anticancer effect of E7 on the expression of P53 proposed in our speculation.

We propose that these characteristics—the gain-
of function on oncogenes accompanying P53 somatic mutations and the antitumor effects of E7 and E5 induced by the inhibition of MDM2/P53 interaction—play roles in accounting for the prognosis difference between HPV-positive and HPV-negative tumors.

**CDKN2A and CDK6**

Somatic mutation of the CDKN2A gene results in functionally deficient p16ink4a and possibly p14arf proteins. Under normal circumstances, the p16ink4a binds to and inhibits CDK4/6, which promote cell cycle progression through the phosphorylation and activation of the retinoblastoma protein (pRb)54, and p14arf interacts with and inhibits P53-degrading MDM2 to lead to the stabilization of P53 and thus cell cycle arrest at the G1/S phase55. With somatic mutation, as demonstrated by Poli with colleagues in head and neck squamous cell carcinomas55, the mutant types of p16 showed significantly lower inhibitory activity than wild type p16, leading to dysregulated CDK4/6 levels and thus to uncontrolled cell cycle progression. On the other hand, functionally deficient mutant p14arf leads to increased levels of MDM2 and thus elevated levels of P53 destruction, further causing loss of cell cycle control55.

In HPV-positive cases, the E7 viral protein has been demonstrated to trigger the overexpression of p16ink4a in both cervical and head and neck squamous cell carcinomas56,57,58. Although in HPV-negative cases, this overexpression would be linked to inhibition of CDK4/6 and thus to antitumor cell cycle arrest, the presence of HPV prevents cell cycle arrest and senescence in these tumors through E7-caused degradation of pRb56,59. This degradation eliminates the tumor-suppressing protein that is essential for the execution of the “arrest signal”59.

In addition to the indirect targeting of p16ink4a described above, HPV infection can bypass the p16ink4a protein and affect downstream targets of this CDKN2A-encoded tumor-suppressing protein. Here we assess the effects of HPV infection on the p16ink4a direct protein and affect downstream targets of this CDKN2A-encoded tumor-suppressing protein. Here we assess the effects of HPV infection on the p16ink4a direct and through mutations of cyclin D1 and CDK4/6. First of all, as previously mentioned, HPV E5 functionally impairs the V-ATPase enzyme, triggering a series of events in the PI3K-Akt signaling pathway and this time leading to decreased transcriptional activity in producing p27, which is an important factor for cell cycle progression from the G1 to the S phase through their interactions with CDK4/6 (Figure 5C)60. This decreased expression of p27 leaves the remaining p27 unable to inhibit the cell cycle progression factors cyclin D1 and CDK4/6, thus leading to dysregulated cell cycle progression16. As detailed by the KEGG pathway “Human Papillomavirus Infections” and as illustrated in Figure 5C as general HPV infection mechanisms, HPV E5’s activation of PDGFβ receptor (PDGFβR), E7’s inhibition of PP2A, E6’s inhibition of PTEN, the degradation of P53 by the E6/E6AP complex, and the inhibition of p27 by both E5 and E716 all lead to the same consequence of dysregulation of cyclin D1 and CDK4/6. In turn, these viral mechanisms contribute to the dysregulation of cell cycle progression, which is a hallmark of cancer, as the dysregulation can leave DNA damage and mutations uncorrected and thereby generate abnormal tumor cells.

Thus, HPV-infection mechanisms powerfully mimic the effects of functional alterations of p16arf induced through mutations, as the ultimate consequence of cell cycle dysregulation is supported in HPV-positive cases by triggering the PI3K-Akt pathway through seven different target genes. The accumulation of these seven separate triggering events could amplify the signal that leads to dysregulation. Although E7’s triggering to overexpression of p16ink4a functionally contradicts with its pRb-degrading mechanisms, the contradiction’s ultimate effect is pro-cancer, and we found no HPV effects that might decrease its carcinogenic efficiency in our searches.

Investigating how HPV mimics the effects of CDKN2A mutations that alter p14arf, we discovered that the E7 viral protein increases p14arf expression. It cleaves E2F-1/pRb binding, releases active E2F-1, and utilizes the ability of transcriptional factor E2F-1 to positively regulate p14arf and induce p14arf expression. Through this increased p14arf expression, E7 actually inhibits the P53-degrading MDM2, protecting P53 from ubiquitination and thus enhancing the cell cycle arrest mechanisms56,61. Experimental results from Kanao and co-authors56 support the above speculations, as they showed an overexpression of p14arf in all the HPV-positive cervical cancers they studied. Furthermore, as discussed in the TP53 section, the presence of HPV E7 inhibits the interaction between MDM2 and P53, as shown in mouse embryo fibroblasts60. So, even if the p14arf somatic driver mutation was already present before infection and had already induced elevated levels of MDM2, the introduction of HPV inhibits the mutant p14arf’s oncogenic effects, as HPV inhibits the interaction between MDM2 and P53 and protects P53 from degradation. This mechanism, on top of the increased p14arf expression, adds another layer of protection for P53.

So, E7 actually mimics the function of normal, nonmutated p14arf instead of mimicking the effects of p14arf with the somatic mutation. These mechanisms create interesting results, as while E7 protects P53 from degradation, E6 binds to E6AP to promote the degradation of P53, as described in the previous section about P53. This point of viral effects self-contradiction may contribute to the relative inefficiency of HPV infection compared to driver somatic mutations leading to better survivals.
CCND1

The CCND1 gene codes for the cyclin D1 protein, which plays an essential role in the G1 phase cell cycle progression. According to a 2018 study on the expression of CCND1 in endometrial adenocarcinomas, the somatic mutation of the CCND1 gene occurs most commonly in the C-terminus of CCND1, likely inhibiting GSK3β-mediated phosphorylation and inhibition of cyclin D1. With this inhibition, cyclin D1 levels become elevated, promoting cell cycle progression.

In HPV infection, the exact same set of genes leading to changes in CDK4/6 expression, as discussed above, are affected by the virus to support changes in CCND1. The only difference exists in that in addition to inhibition of P27, inhibition of CCND1 can be repressed through Akt’s inhibition of GSK3β (see the two separate branches stemming from Akt in Figure 5C). Similar to the CDK4/6 scenario, the inhibition of P27 leads to insufficient repression of cyclin D1 and thereby to dysregulated progression of the cell cycle, which is a significant...
contributing factor to unchecked cell growth and the development of genetically abnormal tumor cells.

The nearly complete similarity between the way HPV affects the P16INK4a downstream targets, CDK4/6, and the way it affects cyclin D1 is not surprising, as the cyclin D1 protein actually forms a complex with CDK4/6, known as the cyclin D/CDK4/6 complex. 

**FGFR1**

The fibroblast growth factor receptor 1 (FGFR1), a cytokine receptor, initiates a cascade of intracellular signaling in the PI3K-Akt and the MAPK signaling pathways upon its binding to fibroblast growth factors, eventually leading to gene transcriptions that are related to cell cycle regulation, migration and cell survival. FGFR’s “family members” FGFR2 and FGFR3 have been shown to interact with the bovine papillomavirus (BPV) E2 protein to impair viral replication, but the study by DeSmet and colleagues failed to identify any associations between FGFR1 and HPV E2. The FGFR1 mutations or more frequently, amplifications, can cause aberrant activations of its downstream pathways through abnormal phosphorylation and can lead to antiapoptotic and mitogenic effects. The mutation in this gene has also been shown to induce transformation in the NIH3T3 mouse embryonic fibroblast cell line.

Interestingly, heparan sulfate proteoglycans, which modulate binding of FGF to its receptor, have been shown to be essential for successful HPV infection. In fact, one of the heparan sulfate proteoglycans, syndecan1, is demonstrated to be involved in HPV penetration to the cell. With this association, we speculate that the heparan sulfate proteoglycans facilitate easier activation of pathways that lead to cell growth and thus facilitate carcinogenesis, as membrane heparan sulfate proteoglycans can act as coreceptors for a series of tyrosine kinase-type growth factor receptors and lower their activation thresholds. We didn’t find any other links between FGFR1 and HPV.

**IGF1R**

IGF1R has long been associated with tumorigenesis and growth. Activated by its ligands IGF1 and IGF2, it protects cells from apoptosis. The C-terminal domain of IGF1R mediates protein-to-protein interactions, which in turn play roles in IGF1R downstream signaling to multiple pathways including the PI3K-Akt and the MAPK signaling pathways, leading to IGF1R’s implication in antiapoptotic responses. Regarding IGF1R mutations, however, studies have shown conflicting results. Some demonstrated that mutations increase IGF1R’s ability to help tumor cells evade apoptosis through either increased binding of a positive mediator of IGF1R signaling, blocked binding of an inhibitory protein, or deletions of the C-terminus, while others demonstrated a promotion of apoptotic functions through IGF1R C-terminus point mutations.

While the HPV Infection KEGG pathway didn’t indicate any associations between HPV and IGF1R, Steller and co-authors through their study using embryonic mouse fibroblasts with disrupted IGF1R to investigate the role of IGF1R in HPV E6 and E7-induced transformation, suggested that HPV E6 may functionally substitute for IGF1R to protect cells from apoptosis, and Pickard with colleagues suggested that HPV E6 and E7 can deplete IGFBP2 expression, thus generating signaling through the IGF1R and FGFR2b to IGF1R downstream pathways.

**ERBB2 (HER2)**

The ERBB2 gene, also known as HER2, is a member of the EGFR family of transmembrane receptor tyrosine kinases and activates signaling pathways regulating cellular proliferation and survival. Herter-Sprie and co-authors summarized three types of somatic mutations of ERBB2: missense mutations in the kinase domain, missense mutations in the extracellular domain and large deletions of the extracellular domain. All three types of somatic mutations activate the ERBB2 protein, leading to downstream signaling in a multitude of pathways including the MAPK and the PI3K-Akt signaling pathway. Such mutational activation and induced downstream signaling lead to proliferation and apoptosis evasion.

Studies have proposed that HPV viral proteins cooperate with ERBB2 to stimulate transformation of normal oral and cervical epithelium. In experiments with oral epithelial cells, the co-expression of E6/E7 and ERBB2 lead to the downregulation of the E-cadherin/catenin complex, which plays important roles in epithelial cell-to-cell adhesion. This downregulation causes cell-adhesion molecule loss, which has been causally associated with loss of intercellular adhesion, gain of invasive properties and the development of epithelial cancers rather than being designated as a consequence of tumor progression. Therefore, the presence of HPV E6/E7 can lead to tumor cell invasion and contribute to metastasis.

Additionally, the HPV-positive status has been linked to higher levels of ERBB2 expression in HNSCC. This ERBB2 overexpression has been associated with the gain of function of several oncogenic pathways caused by p53 mutants, which induce genomic instability and metastasis (as mentioned in the p53 section) that HPV-affected p53 doesn’t seem to be able to mimic. Also, we speculate that this overexpression can suggest that HPV affects ERBB2, as both the increased ERBB2 expressions and ERBB2 somatic mutations can lead to downstream signaling in the MAPK and PI3K pathways. However, the specific mechanisms behind such mimics are not well understood.
HRAS

HRAS is a member of the Ras gene family that plays important roles in the MAPK signaling pathway, which receives external signals that arise from the presence of mitogens and promotes cell growth and proliferation. The activation of the MAPK pathway is actually triggered by activations of previously mentioned ERBB2, IGF1R, and FGFR1 genes (Figure 6A). Mutations in HRAS maintain the HRAS GTPase in its GTP-bound active state, thus producing permanently activated HRAS proteins and therefore constantly activated MAPK pathway signaling process, which eventually leads to dysregulated cell cycle progression and division that are significant hallmarks of cancer. Somatic missense mutations in the Ras family of genes have also been shown to confer gain-of-function mechanisms to oncogenes. As previously mentioned, Ras’ cooperation with mutant types of p53 induce metastasis and angiogenesis, as observed in non-small cell lung carcinoma cells.

Figure 6: Pathway diagrams for mutation and HPV-induced effects of FGFR1, IGF1R, ERBB2, HRAS & MYC.

Processes (either mutations or HPV infections) for which the effects are still under debate or completely unknown are not included in this diagram. This statement applies to FGFR1 in the HPV infection diagram and to IGF1R in both the mutation and HPV infection diagrams. As the diagrams illustrate, HPV is largely able to mimic the loss of cell cycle regulation effects caused by mutations through the targeting of genes that function in the MAPK and PI3K-Akt pathways. It is also able to support loss of intracellular adhesion and thus promote metastasis, effects not achievable by the mutations. However, it is not unable to mimic the gain-of-function-of-oncogene effects of mutations.

A. Pathway diagram for the effects of somatic mutations of FGFR1, IGF1R, ERBB2, HRAS and MYC.

B. Pathway diagram for the effects of HPV infection focusing around FGFR1, IGF1R, ERBB2, HRAS and MYC.
The HPV E5 viral protein, which inhibits the functions of V-ATPase when studied in transfected human keratinocytes, seems to be the main mutation-mimicking agent for HRAS. Through the inhibition, E5 leaves V-ATPase unable to inhibit EGFR's activation of PI3KCA and thus the MAPK signaling pathway. Thus, E5 triggers a cascade of effects that will eventually activate the Ras family of genes, thus indirectly activating the MAPK signaling pathway (Figure 6B). In addition, as shown in the HPV Infection KEGG Pathway, HPV E5 directly activates PDGF/R, which in turn activates the MAPK signaling pathway through Grb2 (Figure 6B). However, some studies have indicated that interaction with PDGF/R is only a characteristic of BPV94,95, raising doubts about the previously described HPV interaction with PDGF/R.

**MYC**

MYC is an important regulator of both cell growth and cell metabolism, functioning in multiple different pathways. MYC binds to MAX, and the complex's normal functions involve the transcriptional repression of cell cycle inhibitor-genes such as p15 and p21,97, transcriptional activation of pro-proliferative cell-cycle regulating genes such as Cyclin D1 and control of DNA replication. With these important responsibilities, MYC is normally tightly regulated by mitogenic signals.

In tumor cells, however, the expression of MYC is almost always increased, sometimes by mutations in the gene itself but usually by the induction of its expression by upstream oncogenic alterations, such as HRAS and IGF1R mutations leading to changes in the MAPK signaling pathway. Such an overexpression causes rapid cell proliferations enabled through the overactivation of pro-proliferative genes, under-inhibition of cell-cycle inhibitor-genes and increased DNA replication activities.

Previously, it has been demonstrated that the integration of HPV viral DNA in genital tumors activated MYC genes. Other studies showed that HPV E6 and MYC associate in vitro and in vivo in a complex. Such an interaction enhances MYC activity and thus is a mutation-mimicking mechanism by the virus.

In addition, as mentioned above, changes in MYC are often caused by changes in MYC's upstream regulators. Therefore, HPV-targeting of MYC's upstream regulators in the MAPK signaling pathway indirectly affects MYC as well, with E5-induced activation of PDGF/R and E5's inhibition of V-ATPase eventually leading to the activation of MYC (Figure 6B). These mechanisms all contribute to HPV's mimicry of mutation leading to MYC overexpression.

**KEAP1, NFE2L2 and CUL3**

KEAP1-NRF2 (protein encoded by the NFE2L2 gene) interaction is important in protecting cells from endogenous and exogenous stresses. In this interaction, KEAP1 functions as a substrate adaptor protein for the CUL3-containing E3 ubiquitin ligase complex, which mediates NRF2 ubiquitylation. This ubiquitylation marks NRF2 for degradation and thus maintains a low expression of genes regulated by NRF2 (cytoprotective proteins HO-1, NQO1, TXNRD1 and the glutathione S-transferase (GSTs) (Figure 6A) that help eliminate oxidative stress when needed to do so. At the same time, KEAP1 functions as a sensor for chemical signals induced by oxidative and electrophilic stresses. Upon the detection of these chemical signals, KEAP1 loses the ability to mediate NRF2 ubiquitylation, allowing NRF2 to accumulate and activate the expression of its cytoprotective target genes, which mainly balance the free radicals and antioxidants in the body to ameliorate the oxidative stress problem.

A study of KEAP1 and NFE2L2 mutations in non-small cell lung squamous cell carcinomas identified that somatic mutations of KEAP1 usually implicate a KEAP1 loss of function, while those of NFE2L2 interrupt binding of NRF2 to KEAP1. A separate study of CUL3 mutations in papillary renal cell carcinomas identified that the mutations of CUL3 cause a complete CUL3 loss of function through deletion. Thus, all of these somatic mutations disable the NRF2-degradation mechanisms and thereby increase the level of intracellular NRF2 and the synthesis of antioxidant and detoxification enzymes, which are the cytoprotective targets of NRF2. While the increased synthesis of antioxidant and detoxification enzymes aren't necessarily carcinogenic effects, the effects of these mutations are also linked to resistance to chemotherapy. A past study by Zhang et al. demonstrated that downregulation of NRF2 triggered chemotherapy sensitivity, implying that upregulation of NRF2 would induce the opposite effect of chemotherapy resistance (Figure 7B). Despite this effect, Frank and co-authors argue that neither NFE2L2 nor Keap1 mutations are drivers, as the majority of non-small-cell lung carcinomas they observed harbored these mutations in co-occurrence with other cancer-related mutations. In fact, it has been shown that oncogenic K-Ras and MYC contribute to the transcriptional induction of Nrfr2.

By mapping a global network of HPV virus-host interactions, Eckhardt with colleagues recently showed that in HPV-positive cases, E1 viral protein physically interacts with and binds to KEAP1 to inactivate KEAP1's functions, thereby leading to freed NRF2 and thus to expression of NRF2's cytoprotective target genes (Figure 7C). As Eckhardt and co-authors pointed out, this E1-KEAP1 interaction mimics the inactivating mutations in the interaction between KEAP1, NRF2 and CUL3, but the effect mimicked isn't necessarily carcinogenic but can be linked to the resistance against chemotherapy.
In addition, HPV E6 has been shown to support oxidative stress\textsuperscript{110} (Figure 7C), and the reactive oxygen species released during oxidative stress can chemically damage the DNA\textsuperscript{111}, contributing to virus-induced mutagenesis.

**FADD and CASP8**

FADD plays an important role in receptor-mediated extrinsic apoptosis, which is triggered by extracellular signals and is to be distinguished from intrinsic apoptosis that is triggered by intracellular signals such as oxidative stress and DNA damage\textsuperscript{112}. In response to extracellular triggers such as killer lymphocytes, Fas is activated (by the Fas ligands carried by those triggers) and binds to FADD via their homologous death domains. This binding triggers the construction of a death-inducing signaling complex (DISC) to eventually lead to the caspase cascade, activating multiple caspases, which then cleave intracellular substrates such as lamin A and PARP to cause apoptosis\textsuperscript{112}. CASP8 is one of these caspases activated (Figure 8A).

A point mutation in FADD inhibits its ability to bind to Fas\textsuperscript{113}, thereby disrupting the apoptosis response to extracellular signals, which now can’t reach the proteins that actually carry out apoptosis. Somatic mutations in CASP8 have been associated with loss of apoptotic functions in a variety of cancers, as CASP8 is rendered nonfunctional\textsuperscript{114-116}. In addition, CASP8 mutations have been shown to confer gain of function mechanisms to other oncoproteins in HNSCC, specifically by activating the NF-κB signaling pathway\textsuperscript{117}. Such an activation leads to anti-apoptotic and pro-proliferation effects by NF-κB\textsuperscript{16,118} (Figure 8B).

During HPV infection, HPV E6 binds to and degrades FADD (Figure 8C), thus protecting cells from apoptosis by preventing transmissions of apoptotic signals through the Fas-FADD pathway and accomplishing similar effects with FADD somatic mutations\textsuperscript{16,119}. However, these similar effects are achieved in slightly different ways, as mutations disrupt FADD’s ability to bind with Fas, while HPV degrades FADD altogether.

Interestingly, Filippova and colleagues\textsuperscript{120} identified different effects on CASP8 for different isoforms of HPV E6 that arise from alternative splicing\textsuperscript{120,121}. The large E6 isoforms were shown to degrade CASP8, while the small isoforms stabilized CASP8. Thus, the large isoforms inhibited the interaction between FADD and CASP8 and in effect prevented the apoptosis signal from being carried
out, while the small isoforms were not able to do so. HPV also seems to mimic the gain-of-function effect that CASP8 mutations confer on the NF-κB pathway, although through a different mechanism. Specifically, HPV E7 inhibits PP2A to increase Akt levels, thus triggering an indirect activation of NF-κB (Figure 8C).

**BIRC2 (cIAP)**

BIRC2 is in the inhibitors of apoptosis protein (IAP) family and inhibits apoptosis through an inactivation of multiple caspases including CASP7 and CASP9, as shown in the Apoptosis KEGG Pathway⁴⁴, and through its regulatory roles within the NF-κB pathway. The effects of somatic mutations of BIRC2 are still not clear, but these mutations might activate the NF-κB pathway to achieve anti-apoptosis effects in HNSCC and other types of cancers, as suggested by Leemans et al.¹²².

We didn't find any interactions between HPV and BIRC2 through our literature searches and pathway analyses. However, we'd like to note that if BIRC2 somatic mutations do activate the NF-B pathway, a similar carcinogenic effect is already achieved by HPV's inhibition of PP2A, perhaps eliminating the need for the mimicry of BIRC2 mutation effects.

**Discussion**

Several perplexing questions surround the topic of HPV-supported carcinogenesis: why is HPV alone incapable of causing cancer, and why do HPV-positive cancers exhibit better prognosis than HPV-negative cancers? To address these questions, in this study, we have presented a concept that HPV infection substitutes for certain driver mutations by mimicking their mutation effects on the pathway to cancer. This mimicry may be accomplished through viral proteins targeting a set of genes that may differ from the
genes that are mutated in HPV-negative cancers, but the ultimate effects are similar. Such a model requires HPV infection to be coupled with an appropriate set of mutations in order for progression to cancer to occur, since HPV does not target the full set of cancer driver genes.

We have found ample evidence that support our concept. First of all, we noticed that in HPV-associated cancers such as cervical cancer and head-and-neck cancer, HPV-negative cases generally have more somatic mutations than HPV-positive cases do\textsuperscript{17,19}, and that the majority of mutated genes have a higher mutation frequency in HPV-negative cases\textsuperscript{19}. These data confirm our proposal, as they correspond to the reasoning that since HPV mimics certain somatic mutations, HPV-positive cancers would require the presence of a smaller number of somatic mutations because the presence of HPV supports the carcinogenic effects that originally needed to be induced by somatic mutations. However, we'd like to point out that this mutation frequency difference is subject to intratumor heterogeneity, as the mutation profiles differ to a certain degree across different regions within the same tumor and across metastatic tumor cells. Therefore, the mutation frequencies data we referenced\textsuperscript{17,19} might not capture the entire mutational landscape and furthermore could be subject to the limitations of the whole-exome sequencing techniques used. In a 2013 study by Zhang et al. investigating the heterogeneity of HNSCC\textsuperscript{123}, the use of whole-genome instead of whole-exome sequencing showed greater HPV-positive HNSCC mutation frequencies than previously observed, revealing a large number of mutations in the non-coding introns. The impacts of these mutations in non-coding sections of the genome are becoming increasingly accepted\textsuperscript{123} and are very pertinent to our study, as they can confer HPV-positive tumors additional oncogenic potential and could possibly undermine the difference in mutation frequencies we noted based on past studies. The effects of these mutations are worth investigating in the future.

Secondly, we raise an objection to the previous hypothesis that the prognosis difference between HPV-positive and HPV-negative cancers are due to age instead of the presence of HPV, as data from studies conducted by Lei and co-authors\textsuperscript{14} and Retting with colleagues\textsuperscript{15} clearly demonstrate that regardless of age, HPV-positive prognosis rates are better than HPV-negative ones. Thus, establishing HPV as an important prognostic factor, we are able to relate our proposal to the prognosis differences between HPV-positive and HPV-negative cases, as we reason that the differences between mutation and HPV viral mechanisms can possibly account for the prognosis difference.

Finally, in an effort to elucidate how HPV could account for the prognosis difference, we analyzed the differences between the effects of HPV viral infections and those of somatic mutations on 17 specific genes: TP53, CDKN2A, FAT1, CCND1, MYC (c-MYC), NFE2L2, CASP8, FGFR1, CDK6, BIRC2 (cIAP), AJUBA, CUL3, EPFA2, ERBB2, KEAP1, HRAS, and IGF1R. These genes are extracted from The Cancer Genome Atlas mutation frequency database and are the ones that are exclusively mutated in HPV-negative cases. In this analysis, by reviewing past experimental evidence of virus-host interactions and of mutation effects, we have pieced together many ways in which HPV powerfully mimics the effects of somatic mutations. With the exception of EPFA2, AJUBA, BIRC2 and FGFR1, all other extracted genes’ functions were altered in one way or another by HPV to achieve mutation-mimicking effects. In other words, the carcinogenic effects of 13 out of the 17 selected genes are mimicked by HPV. These oncogenic effects of mutations in EPFA2, AJUBA, BIRC2 and FGFR1 that are not mimicked by HPV (to the best of our knowledge at the time of this study) potentially contribute to the prognosis difference conferred by the different HPV statuses, as mutations in HPV-negative cases disrupt cancer pathways that are not as affected in HPV-positive cases. Additionally, it is worth noting that some of the effects are achieved by mutations and by HPV in drastically different ways. For example, while an HRAS mutation directly causes the permanent activation of the MAPK pathway and leads to dysregulated cell cycle progression, HPV E5 accomplishes a similar effect by inhibiting V-ATPase, a far upstream regulator of HRAS. This inhibition would then lead to a cascade of effects that eventually activate the MAPK pathway. Although achieving the similar effect of MAPK-pathway activation, these different methods may confer differences in the degrees of severity of the effect, as direct activations of the pathway by mutations may be more efficient and effective than indirect activations by viruses. As an additional example, mutant p16 leads to uncontrolled cell cycle progression as it demonstrates a reduced inhibition of cell cycle progression promoter CDK4/6 when compared to its wild type counterpart, while HPV triggers the overexpression of p16, which would naturally induce cell cycle arrest, but then prevents cell cycle arrest as it degrades pRb, thereby eliminating the protein that is essential for carrying out the cell cycle arrest. These two methods achieve the similar effect of inducing cell cycle dysregulation, but the degree of severity of the effects might be different. These degree-of-severity differences are not addressed in this current study but should be investigated in the future, as they could be major contributors to the prognosis difference between HPV-positive and HPV-negative cases.

Further addressing the prognosis difference, we have also identified and described multiple HPV infection mechanisms that have been previously uncovered by experimental procedures and that possibly reduce the virus’ efficiency in supporting cancer. First of all, a viral protein sometimes supports effects that undermine the carcinogenic effects of another viral protein. In other
words, the virus sometimes has self-contradicting effects. Such a scenario can be seen in HPV E7’s inhibition of interaction between MDM2 and p53, as this inhibition undermines a combined effort of multiple viral proteins to activate MDM2, which degrades p53. Secondly, the virus is unable to support some of the gain-of-function of other oncogene effects that somatic mutations induce. This is seen in the case of TP53, the mutation of which, on top of inducing effects through the inhibition of p53’s transcription mechanisms, increases genomic instability and confers more resistance to a number of antitumor agents. Finally, the differences between various isoforms of viral proteins might introduce another source of “inferiority” for the virus in carcinogenesis. As pointed out in the *FADD and CASP8* section, different effects on *CASP8* have been identified for different isoforms of HPV E6. Some isoforms confer anti-apoptosis advantages to the tumor cells, while others do not. Such differences between isoforms perhaps render a portion of HPV-supported cancers more dangerous than others. Also, the existence of variants of the same viral proteins due to natural genetic differences raises uncertainty to the consistency of the functions of HPV infection. The same point of argument can be extended to the multiple types of HPVs, including both the high-risk and the low-risk types. More research into the specific differences between isoforms and HPV types would contribute to our overall understanding of the carcinogenic effects of HPV infections.

Overall, our analysis yielded ample evidence supporting our hypothesis for head-and-neck cancers. We expect this hypothesized model to apply to other HPV-associated cancers, such as cervical cancer, because of the model’s ability to account for the differences in the number of somatic mutations in tumors of different HPV statuses, and because of the commonality of this difference across almost all types of HPV-associated cancers.

**Conclusion**

We proposed a novel explanation for the role of HPV in inducing cancer: the mimicry of the effects of certain somatic mutations. We found ample evidence for this proposal by compiling and analyzing the previously identified viral infection mechanisms that target certain genes to induce carcinogenic effects. This proposal explains the insufficiency of HPV alone to induce cancer, as HPV infections substitute for some, not all, somatic driver mutations. It also implies the acceleration of the carcinogenic process in HPV-positive cases, and thus we see the trend that HPV-positive cancer patients are younger than HPV-negative ones. Finally, our proposal offers a possible explanation for the better prognosis conferred by HPV-positive cancers, as we showed through compilation of past experimental evidence that the viral mechanisms achieve only partial mimicry of oncogenic mutation effects, that the effectiveness of viral mechanisms differ by the types of HPV, and that certain infection mechanisms contradict other viral carcinogenic effects.

**References**


underlie a phenotype within the RASopathy spectrum and contribute to leukaoeogenesis. Human molecular genetics. 2014; 23(16): 4315-4327.


