bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Molecular mechanisms of human papillomavirus-related carcinogenesis in head and neck cancer

Farhoud Faraji*, Munfarid Zaidi*, Carole Fakhry, Daria A. Gaykalova

Department of Otolaryngology-Head and Neck Surgery, Johns Hopkins School of Medicine, Baltimore, MD.

* These authors contributed equally

Corresponding author:

Daria A. Gaykalova, PhD
Assistant Professor
Division of Head and Neck Research
Department of Otolaryngology-Head and Neck Surgery
Johns Hopkins University School of Medicine
Address: CRBII, Room 5M05C
1550 Orleans St., 5M North
Baltimore, MD 21287-0014
Phone: 410-502-5144
Fax: 410-614-1411
Email: dgaykal1@jhmi.edu

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

ABSTRACT

This review examines the general cellular and molecular underpinnings of human papillomavirus (HPV)-related carcinogenesis in the context of head and neck squamous cell carcinoma (HNSCC) and focuses on HPV-positive oropharyngeal squamous cell carcinoma in areas for which specific data is available. It covers the major pathways dysregulated in HPV-positive HNSCC and the genome-wide changes associated with this disease.

Keywords

HNSCC; squamous cell carcinoma; oropharyngeal; HPV; viral integration; chromatin

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

I. INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) encompasses a heterogeneous group of malignant neoplasms arising from the non-keratinizing epithelium of the upper aerodigestive tract. Anatomic subsites of HNSCC include the oral cavity, nasopharynx, oropharynx, hypopharynx, and larynx. Squamous cell carcinoma arising from these subsites collectively represents the sixth most common malignancy worldwide, accounting for 932,000 new cases and 379,000 deaths in 2015 [1].

Over the past four decades, striking epidemiological trends have been observed in HNSCC. Although the overall incidence of HNSCC has declined slightly, the relative contribution of each anatomic subsite to the overall incidence of HNSCC has shifted dramatically [2]. Incidence rates of tumors arising from non-oropharyngeal subsites (oral cavity, hypopharynx and larynx) have decreased while the incidence of oropharyngeal squamous cell carcinoma has steadily grown [3]. These subsite-specific epidemiological trends have been attributed to shifts in societal factors that have resulted in changes in exposure to two divergent, but complementary classes of HNSCC risk factors: (1) tobacco and alcohol consumption and (2) human papillomavirus (HPV) infection. Successful public health campaigns in high-income countries are largely credited with achieving population-level decreases in tobacco and alcohol consumption [4] with concomitant declines in tobacco-associated tumors such as non-oropharyngeal HNSCC and lung cancer [5]. Trends toward sexual practices that increase the risk of contracting sexually transmitted pathogens, like HPV, have been linked to the rise in HPV-associated cancers including oropharyngeal HNSCC (OPC) and anal cancers [5, 6]. Currently, HPV-positive OPC cases are surpassing the incidence of HPV-positive cervical cancer [3, 7].

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Human papillomavirus is the most common sexually transmitted infection in the United States and the primary infectious cause of HNSCC [8, 9]. Although the spread of high-risk HPV infection is pervasive, the virus is cleared by most people within 18 months [10]. It is believed that persistent infection is necessary for the development of HPV-associated malignancies. The oropharynx exhibits the strongest relationship to HPV. Indeed, HPV-positive oropharyngeal squamous cell carcinoma (OPC), is recognized as a distinct neoplastic entity with a unique molecular, histopathological, epidemiological, and clinical profile [11-13]. Patients with HPV-positive OPC diverge from the classical profile of HNSCC patients in that they are more often younger than 60 years in age and less likely to report a history of heavy tobacco and alcohol consumption [12]. HPV-positive OPC also exhibits marked sensitivity to treatment and a significantly better prognosis than HPV-negative HNSCC [11], observations which have led to the establishment of a staging system specific to this entity [14] and shifts in treatment paradigms [15]. These distinct epidemiological and clinical features are manifestations of the unique underlying biology of HPV-positive OPC. However, much of the research into the role of HPV in HNSCC was conducted before intimate links between HPV and OPC were widely recognized.

II. HUMAN PAPILLOMAVIRUS

Papillomaviridae is an ancient clade of non-enveloped viruses with a circular double-stranded DNA genome [16]. Within this family, approximately 200 human papillomavirus genotypes, or 'types', have been described based on viral genome sequence [17]. Twelve HPV types have been defined by the WHO as high-risk and exhibit high oncogenic potential [18]. At least 10 of these oncogenic HPV types (HPV16, 18, 31, 33, 45, 51, 52, 56, 58, and 59), as well as 6 low-risk HPV types (11, 32, 44, 53, and 81), have been isolated from HNSCC tumors [19-23] (Table 1). HPV16 represents the primary viral cause of HNSCC and is identified in at least 87%

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

of HPV-positive OPC [24]. HPV18 and HPV33, the next most prevalent types, account for most of the remainder of HPV-positive HNSCC [24].

All papillomaviruses possess a genome of approximately 8 kilobases encoding 8 open reading frames (ORFs) that enable viral genome replication and viral particle assembly (Figure 1). ORFs of the HPV genome are organized based on the timing of expression with respect to the viral life cycle: early (E1, E2, E4, E5, E6, and E7) and late (L1 and L2) genes [25]. The E1 viral helicase and E2 DNA-binding protein directly mediate viral genome replication, while E4, E5, E6, and E7 are accessory proteins that coordinate viral genome amplification and virulence. The late genes L1 and L2 encode viral capsid proteins necessary for the final stages of virion assembly and mediate viral entry into future host cells. The functional diversity of this limited set of ORFs is expanded through complex patterns of viral transcript splicing and stage-specific gene expression that is linked to host cell differentiation [26].

Infection

The host cells for HPV infection are keratinocyte progenitors located in the basal layer of stratified squamous epithelia and adhered to the epithelial basement membrane. Experimental models suggest that infection requires viral access to the basement membrane [27] (Figure 2). In the epidermis and anogenital tract, HPV gains access to basal cells through micro-abrasions that occur during sexual or other direct physical contact [28].

In the oropharynx, HPV infection may occur in the absence of epithelial abrasion. The palatine, lingual, tubal, and adenoid tonsils are lymphoid structures collectively known as Waldeyer's tonsillar ring. Constituents of Waldeyer's ring possess a specialized reticulated squamous epithelium infiltrated with lymphoid tissue. The reticulated epithelium contains a fenestrated, discontinuous basement membrane thought to allow immune cells access to oral

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

antigens [29]. These fenestrations also represent natural interruptions in epithelial barriers that likely provide HPV access to basal keratinocytes in the absence of traumatic epithelial disruption (Figure 2) [30]. Indeed, the unique epithelial properties of Waldeyer's ring may explain the disproportionate tendency of HPV to cause squamous cell carcinoma of the palatine and lingual tonsils [31].

Upon reaching the basal keratinocyte, HPV preferentially binds components of the extracellular matrix. The L1 capsid protein directly engages basement membrane heparin sulfate proteoglycans [32], triggering conformational changes in L1 and L2 that transfer virion particles to host cellular uptake receptor(s) necessary for viral internalization. Although cellular HPV receptors mediating viral entry have not been definitively identified, accumulating evidence suggests that tetraspanin-enriched microdomains on the plasma membrane serve as the primary platform for viral entry into the cell [33]. Once internalized, virion particles undergo endosomal transport, uncoating, and cellular sorting into the nucleus [34].

Upon entry and uncoating, the circular viral DNA is transported to the nucleus where it exists as an episome, a genetic element separate from the host cell genome, that employs host cell enzymes to replicate its genome along with host chromosomes and is maintained at low-copy number (~50-100 copies per cell) [35]. Tight regulation of low-copy replication of the viral genome in basal keratinocytes serves as one mechanism by which HPV evades the host immune system. Oncoproteins E6 and E7 are expressed prior to productive viral replication, driving cell cycle entry and cell proliferation in the basal and parabasal cells [26]. In oncogenic HPV genotypes, cell cycle dysregulation by E6 and E7 constitutes the initial steps driving HPV-related carcinogenesis.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

The expression of viral genes and virion production increases rapidly in daughter keratinocyte progenitors undergoing differentiation and progressing toward the mucosal surface [36]. In normal stratified squamous epithelia, only basal keratinocyte stem cells possess the potential to proliferate. Asymmetric cell division of these keratinocyte stem cells leads to two daughter cells: one renewed basal stem cell and one cell destined to become a terminally-differentiated keratinocyte. HPV-encoded genes perturb key keratinocyte differentiation and proliferation pathways to favor virus production and viral life cycle completion. In high-risk HPV types, these perturbations predispose cells to neoplastic transformation.

III. CANONICAL MECHANISMS OF HPV-DRIVEN CARCINOGENESIS

Knowledge derived from the study of HPV-related cervical cancer composes the general framework for understanding mechanisms of carcinogenesis in HPV-related HNSCC. High-risk HPV types encode virulent alleles of two viral proteins, E6 and E7, that endow viral oncogenic potential by targeting molecular pathways that underlie neoplastic transformation. Downstream consequences of high-risk HPV E6 and E7 action includes the promotion of cell cycle progression, disruption of keratinocyte differentiation, keratinocyte immortalization, genomic instability, and neoplastic transformation [37].

A central advance in the understanding of HPV-related carcinogenesis arose with the elucidation of mechanisms for the oncogenic potential of E6. High-risk HPV E6 proteins bind to and induce ubiquitin-dependent proteolysis of the p53 tumor suppressor protein [38]. P53 is the protein product of the *TP53* gene, a regulator of the DNA-damage response [39], G_2/M cell cycle transition [40], and the most commonly mutated tumor suppressor gene in cancer (Figure 3) [41]. Only E6 alleles from high-risk HPV types bind p53 and the strength of the binding affinity of E6 alleles for p53 correlates with that HPV type's oncogenic potential. For example, HPV16, which

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

is responsible for more HPV-related HNSCC than HPV18 [42], encodes for an E6 allele that binds p53 with twice the affinity of the HPV18 E6 allele [43].

In an analogous mechanism to the action of E6 on p53, the HPV E7 oncoprotein exerts its oncogenic effect by facilitating the proteolytic destruction of Retinoblastoma (Rb)-family tumor suppressors [44]. Rb enacts its primary inhibitory effect on the cell cycle by regulating the nuclear accumulation of E2F and thereby controls the G1 to S-phase cell cycle transition (Figure 3). E7 induces enhanced degradation of Rb protein via a ubiquitin-proteasome pathway [44]. Degradation of Rb releases E2F family mitogenic transcription factors into the nucleus, promoting entry into S-phase and activating proliferative transcriptional programs [45, 46]. Degradation of Rb also releases *CDKN2A* gene expression, which codes for the tumor suppressor $p14^{ARF}$, as well as p16 (or $p16^{INK4A}$ - a surrogate marker of HPV infection [47, 48]). Low-risk E7 binds Rb at an affinity that is insufficient to induce levels of Rb degradation necessary to promote neoplastic transformation (Figure 3) [49, 50].

Beyond these well-characterized actions, E6 an E7 modulate the activity of numerous other cellular proteins. Recent data suggest that viral oncoproteins E6 and E7 bind and modify the activity of numerous proteins and molecular pathways [51]. E6 has been reported to interact with 83 cellular proteins and E7 with 254 [52].

IV. GENOME-WIDE SOMATIC MUTATIONS OF HPV-RELATED HNSCC

Notable studies have shed light on the unique mutational profile of HPV-positive HNSCC. They have served to distinguish HPV-positive HNSCC from their HPV-negative counterparts based on molecular, rather than simply clinical, markers. Insight gained from their contributions provide additional targets for which prevention and therapy can be developed.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

The Cancer Genome Atlas (TCGA) consortium presented the most comprehensive HNSCC genome-wide analysis to date [1]. Analysis of 279 HNSCC, of which 36 were HPV-positive, identified recurring activating mutations in *PIK3CA* in HPV-positive HNSCC [53] [1]. *PIK3CA* encodes the catalytic subunit of phosphoinositol 3-kinase (PI3K) of the AKT/PI3K/AKT/mTOR pathway, which is essential for protein synthesis, cell growth, proliferation, and survival [54]. Overexpression and constitutive activation of *PIK3CA* leads to inhibition of apoptosis and, thus further tipping the balance toward carcinogenesis [54]. TCGA found 56% of HPV-positive HNSCC showed activating mutations in this gene [53].

Amplification of the *E2F1* was another important HPV-specific finding [53]. *E2F1* encodes a mitogenic transcription factor. E2F1 target genes encode for proteins that are critical for progression of the cell cycle through the G1/S transition [55]. Interestingly, E2F1 is regulated by the Rb tumor suppressor [56], a target of HPV E7 [44]. Overexpression of *E2F1* leads to dysregulation of this cell cycle checkpoint, pushing the cell toward carcinogenesis. TCGA found activating mutations in *E2F1* in 19% of HPV-positive HNSCC, compared to 2% in HPV-negative HNSCC [53].

TCGA found the inactivation of *TRAF3*, found in 22% of HPV-positive tumors, to be another mutation typical of HPV-positive HNSCC [53]. The protein product of *TRAF3* is a negative regulator of the NF-κB signaling, a mediator of cellular proliferation and apoptosis blockade [57]. As an inhibitor of NF-κB, TRAF3 indirectly promotes apoptosis and suppresses proliferation. Inactivating mutations in TRAF3, thus, promote molecular pathways known to plays significant roles in carcinogenesis. Indeed, studies beyond TCGA have identified mutations in TRAF3 and CYLD, another regulator of NF-κB [58], as defining features of a subset of HPV-positive HNSCC [59].

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

In addition, both HPV-positive and -negative tumors contained recurrent focal amplifications for 3q26/28, a region that involves, among others, the squamous lineage transcription factor SOX2 [53]. This factor controls gene expression in pluripotent stem cells, contributing to the continuation of their pluripotency. Additionally, it plays a role in the reprogramming of somatic cells, reversing their epigenetic configuration back to the pluripotent embryonic state [60]. Evidence of amplification of this gene provides a mechanism for conversion of squamous cells to cancer stem cells, a concept that has been supported by other studies [61].

Other important genes mutated in HPV-positive HNSCC include *DDX3X*, *FGFR2, FGFR3*, *KRAS*, *MLL2/3*, and *NOTCH1* [62] – overlap with mutations in HPV-driven cervical tumors [63].

Mutational spectrum

Several genome-wide studies including TCGA have found that HNSCC exhibits mutational spectra that appear to be specific to HPV-status [53]. HPV-positive HNSCC display enrichment of cytosine-to-thymidine (C to T) mutations at TpC sites [62, 64], which are thought to result from dysregulation of APOBEC. APOBEC comprises a family cytosine deaminases that play key roles in viral immunity by mutating viral DNA and restricting viral replication [65]. However, high levels of APOBEC activity has been observed in several virally-induced cancers to also induce mutation of the host genome, and patterns of APOBEC-related mutagenesis with cytosine deaminase hyperactivity is a common event in virally transformed cancers [62, 64].

HPV-related dysfunction of the APOBEC family proteins in HPV-positive HNSCC [62, 64, 66] likely causes non-synonymous mutations within isolated hot-spots. Thus, the prototypical gene mutated by APOBEC dysregulation in HPV-positive HNSCC is *PIK3CA*, in which C to T

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

mutations at two hotspots cause recurring E542K and E545K amino acid substitutions that cause gain of function mutations and result in PIK3CA kinase activation [62, 67].

V. VIRAL INTEGRATION AND CARCINOGENESIS

The insertion of HPV genomic DNA into the host cellular genome, termed viral integration, is not considered a part of the natural life cycle of HPV; however, it may contribute to HPV-related carcinogenesis [68, 69]. Viral DNA integration into human genome more often occurs with oncogenic HPV types [70], and is frequently found in HPV-positive OPC [68, 69]. Indeed, half of HPV-positive HNSCC (39-71%) have the virus integrated into the host genome [13, 71, 72].

Viral transcripts expressed from integrated viral DNA have been found to be more stable and expressed at higher levels than those from episomal viral DNA [73]. Recent data suggest that integration of oncogenic viruses, including HPV, into the host genome, are non-random [73-76]. A common breakpoint of the viral genome occurs in E2, the virally-encoded transcriptional inhibitor of E6 and E7 (Figure 4) [71, 73]. In this manner, integration may result in the upregulation of HPV oncoproteins E6 and E7 [73, 77, 78]. Chromatin regulatory elements near the site of integration also may cause upregulation of integrated E6/E7 genes [72, 73]. Such elements include areas of transcriptionally-active "open" chromatin and non-coding transcriptional regulatory sites including enhancers or super-enhancers [79]. Thus, in the case of bovine papillomavirus 1 (BPV1), the correspondence between integration and open chromatin occurs because of BRD4, a reader of acetylated lysine residue 27 on histone 3 (H3K27Ac), mediates tethering of the viral genome to host chromatin [80-82]. Recent data correlate HPV16 integration sites in cervical cancer with super-enhancer domains of open chromatin, enriched by H3K27Ac, MED1 (RNA polymerase II [PolII] transcription subunit), and BRD4 proteins [79]. It is possible that a similar mechanism also mediates viral integration in HPV-positive HNSCC.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Further, HPV integration promotes genetic instability [68, 83], and confers an advantage in cellular proliferation [77]. While some evidence supports a role for HPV DNA integration in HPV-related carcinogenesis, the general principles HPV integration and its mechanistic role in HPV-related carcinogenesis remain controversial.

Early studies employing low-throughput molecular biological approaches sought to elucidate the potential relationship between HPV integration and HNSCC progression. Targeted analysis of viral copy number and RNA expression was conducted on HPV-positive OPC tumor samples by quantitative real time polymerase chain reaction (qRT-PCR) [72]. Twenty-nine of 75 tumor samples (39%) were found to amplify HPV16-human mRNA fusion products, evidence consistent with viral integration. Comparison of HPV-integrated and episomal OPC samples found no difference in viral copy number, expression of HPV16 interrupted genes, or expression of viral genes E2, E6, and E7 between HPV-integrated and episomal tumors [72]. This data suggested that HPV16 integration had no impact on HPV-dysregulated human transcripts and or the expression of key viral oncogenes [72]. In a subsequent qRT-PCR-based analysis of 7 HPV-positive HNSCC cell lines with integrated and/or episomal viral DNA, Olthof et al drew similar conclusions [71]. This latter study showed that viral E2, E6 and E7 gene expression was independent of viral load and the number of viral integration sites, suggesting that integration did not affect viral gene expression [71].

A more recent analysis of high-throughput data from TCGA, however, partially contradicted these findings [71, 72]. Parfenov et al defined HPV status by interrogating 279 HNSCC samples with associated RNA-sequencing data for samples expressing HPV E6 transcripts [73]. Viral integration status was then analyzed from associated DNA-sequencing data in the HPV-positive samples. Of 35 HPV-positive HNSCC samples, each of 25 tumors exhibited

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

non-random viral integration into 1 to 16 regions of the human genome, leaving *E6* or *E7* intact in all but one case. Interestingly, seventy-five of the 103 total integration events occurred in or within 20 kilobases of a known gene. Higher detail insertion-site analysis identified several putative mechanisms demonstrating that non-random HPV integration may confer a selective advantage to tumor cells and promote tumor progression.

HPV was found to integrate into and alter the function of several known tumor-associated genes, including *RAD51B, ETS2, PDL1*. One tumor harbored three viral insertions into intron 8 of the *RAD51B* gene. Expression analysis of this tumor demonstrated significantly elevated expression of exons 9-11 and 13 and alternative transcripts unlikely to produce functional RAD51B protein. RAD51B is an essential component of DNA double-strand break repair machinery and, consistent with its role as a tumor suppressor gene, loss-of function *RAD51B* mutations have been reported in cancer [84]. Monoallelic HPV integration was detected in *ETS2*, another tumor suppressor, resulting in significant reductions in the expression of exons 7 and 8. This suggests alternative splicing pattern potentially may lead to biallelic loss of *ETS2* function by a dominant negative effect [85]. Similarly, intronic HPV integration into the *PDL1* gene resulted in augmented expression of an alternative *PDL1* transcripts. Alternative PDL1 transcripts have been associated with poor prognosis renal cell carcinoma [86].

Parfenov et al also identified HPV integration in intergenic loci known to regulate tumor-associated genes *TPRG1, TP63,* and *KLF5* [73]. The concomitant transcriptional upregulation of these target genes was directionally consistent with tumor-promoting roles of these genes [13, 87]. Similarly, HPV integration upstream of the *NR4A2* oncogene associated with *NR4A2* overexpression [88]. Consistent with these findings, analysis of 8 HPV-positive HNSCC cell lines by qRT-PCR identified HPV integration events and evidence consistent with dysregulation of

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

several cancer-related genes *TP63, DCC, JAK1, TERT, ATR, ETV6, PGR, PTPRN2,* and *TMEM237* [89]. Although preliminary, these findings set forth putative mechanisms by which HPV integration may promote tumor progression.

VI. TRANSCRIPTOME-LEVEL ALTERATIONS OF HPV-RELATED HNSCC

Gene Expression Signatures and HNSCC Subtypes

Not until recently have comprehensive studies of gene expression in HPV-positive and HPV-negative HNSCC tumor types been conducted to distinguish them based on underlying genetic aberrations. In addition to further distinguishing the molecular features of HPV-positive HNSCC, this research has led to the proposal of HNSCC subtypes that better capture its biological heterogeneity and may facilitate the development of subtype-specific targeted therapeutics.

This line of study was pioneered by Seiwert et al, who found that HPV-positive tumors showed a distinct genetic profile with unique mutations in *DDX3X, CYLD*, and *FGFR,* as well as enrichment for PI3K pathway alterations and rarer *KRAS* mutations common for HPV-negative HNSCC tissues. As an opposite, the mutational profile in HPV-negative HNSCC included enrichment for mutations in *TP53, CDKN2A, MLL2, CUL3, NSD1,* and *PIK3CA* genes, and copy-number increases in *EGFR*, *CCND1*, and *FGFR1*. The group also found that somatic aberrations in DNA-repair genes (*BRCA1/2*, Fanconi anemia genes, and *ATM*) may contribute to chemo- and/or radiosensitivity of HPV-positive tumors. These DNA-repair genes mutations occurred in HPV-positive tumors in non-/light smokers [66]. Considering a recent report that *RAD51B*, a BRCA2 cofactor [90], was a potential integration target for HPV16 leading to loss of its DNA-repair function, may indicate that the role of DNA repair gene disruption in HPV-positive HNSCC carcinogenesis warrants further investigation.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

One group, Keck et al, established subtype classification using gene expression-based consensus clustering, copy number profiling, and HPV status on a cohort of 134 locoregionally advanced HNSCCs. They identified five subtypes of HNSCC, including, most importantly, two biologically distinct HPV subtypes. The first HPV-positive HNSCC subtype is the classic HPV positive subtype (CL-HPV). This subtype is characterized by significant enrichment of the polyamine degradation pathway, which is relevant for detoxification, for example, related to tobacco use. More than 40% of CL-HPV tumors show keratinization, and this subtype has a higher proliferation rate compared to the other subtypes. The second HPV-positive subtype is the inflamed/mesenchymal HPV-positive subtype (IMS-HPV), which shows an immune and mesenchymal phenotype. Prominent tumor infiltration with CD8+ lymphocytes is characteristic of IMS-HPV, independent of HPV status. This tumor subtype shows no keratinization, and exhibits elevated mesenchymal markers. Establishment of these subtypes carries translational implications for the development of subtype-specific biomarkers, and it provides a biological justification for the use of different treatment approaches among HPV-positive HNSCC based on subtype [91].

The emerging role of ErbB family proteins in HPV-positive HNSCC therapeutics

ErbB family proteins are encoded by 4 genes that play a central role in carcinogenesis: *ERBB1* (also known as *HER1* or *EGFR*), *ERBB2* (also known as *HER2* or *CD340*), *ERBB3* (also known as *HER3*), and *ERBB4* (also known as *HER4*) [92, 93]. Pollock et al published a paper in 2015 which reported that expression of total HER2, total HER3, HER2:HER3 heterodimers, and the HER3:PI3K complex were significantly elevated in HPV-positive HNSCC [94]. They also reported that Afatinib, an anti-ErbB family small molecule inhibitor, significantly inhibited cell growth compared to cetuximab in HPV-positive cetuximab-resistant HNSCC cell lines. These

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

findings suggested that targeting ErbB family receptor tyrosine kinases may be a potentially effective treatment for HPV-positive HNSCC.

However, these results were not supported by the LUX Head and Neck 1 trial [95, 96], which showed that both patients with p16-positive tumors and patients whose tumors progressed after cetuximab therapy derived less benefit from afatinib. One possible explanation for this apparent discrepancy was a difference in definitions; for example, the LUX study considered previous exposure to cetuximab equivalent to cetuximab-resistance [95, 96].

The discrepancy was addressed with a recent study investigating the relationship between HPV oncoproteins and HER3-mediated signaling, and the role of HER3 as a potential molecular target in HPV-positive HNSCC [97]. This study proposed a novel mechanism by which HPV can promote tumor growth through the mTOR/AKT/PI3K pathway. HER3 expression was found to be significantly increased in HPV-positive cell lines compared to HPV-negative cell lines. Brand et al then investigated the role of E6 and E7 oncoproteins in *HER3* expression. They found that E6/E7 gene silencing led to diminished *HER3* production, while increased E6/E7 oncoprotein production led to increased *HER3* production. The results indicate a relationship between HPV infection and *HER3* expression that is mediated by the E6 and E7 oncoproteins [97].

Having found evidence to support this relationship, the group then studied how *HER3* expression might affect cellular proliferation and tumor growth in HPV-positive HNSCC cell lines. Their suspicions were confirmed; genetic ablation of *HER3* resulted in statistically significant inhibition of cellular proliferation (21-55%) [97]. Furthermore, *HER3* knockdown led to reduced AKT and RPS6 phosphorylation [97]. The reduction of these proteins, which are components of the same mTOR/AKT/PI3K pathway discussed in the genomics section of this review, suggests

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

that HER3 exerts its proliferative effect by signaling through the mTOR/AKT/PI3K signaling pathway [97].

Based on their findings suggesting that *HER3* is overexpressed in HPV-positive cell lines, patient-derived xenografts, and human tumors, Brand et al investigated the sensitivity of HPV-positive preclinical models to anti-HER3 monoclonal antibody therapy, such as KTN3379. KTN3379 was shown to be effective in growth inhibition in 4 out of 4 HPV-positive HNSCC cell lines, versus 2 out of 5 HPV- lines [97]. KTN3379 inhibited phosphorylation of HER3 in all HNSCC cell lines regardless of HPV status, but analysis of downstream PI3K signaling revealed that phospho-AKT and phospho-RPS6 were more significantly inhibited in three HPV-positive cell lines versus three HPV-negative cell lines [97]. HPV-positive human HNSCC tumors xenografted into mice also showed statistically-significant reduction in tumor volume, as well as reduced levels of total *HER3*, phospho-HER3, phospho-AKT, and phospho-rpS6 [97]. These findings indicate that anti-HER3 therapy may have potential as an tumor type-specific HPV-positive HNSCC therapeutic target. Further supporting a potential role of ErbB family proteins, retrospective study of a patient with HPV-positive OPC with complete response to HER2-targeted therapy found that a dermal metastasis from this patient exhibited HER amplification[98].

VII. EPIGENETIC ALTERATIONS IN HPV-RELATED HNSCC

Additional amplification of the host genome at the site of HPV integration [53], and dysregulation of p53 and Rb tumor-suppressor genes by HPV oncoproteins E6 and E7 do not fully account for the genome-wide spectrum of gene expression changes in HPV-positive HNSCC [53]. Chromatin structure and epigenetic modifications, such as DNA methylation and histone-modifications, regulate gene expression, and their changes may explain genome-wide dysregulation found in HPV-positive HNSCC.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Methylation

The drive to understand the molecular underpinnings of HPV-positive HNSCC and, based on that knowledge, to develop biomarkers predictive of prognosis and treatment has now progressed to include epigenetic changes. The disruption of epigenetic signatures can cause alterations in gene function that lead to malignant changes in cells [99]. DNA methylation, the addition of a methyl group to a DNA molecule, especially the C5 position of cytosine, is the most common form of epigenetic modification. Given how often promoter methylation is found to be the mechanism of action in transcriptional silencing in HNSCC [100], an exploration of the unique methylation patterns inherent in HPV-positive HNSCC may be helpful in creating clearer molecular profiles and tailored treatment plans.

One such exploration was conducted by Kostareli at al, found a negative correlation between gene promoter hypermethylation and gene transcript levels in $ALDH1A2^{lo}$, $OSR2^{lo}$, $GATA4^{hi}$, $GRIA4^{hi}$, and $IRX4^{hi}$ in HPV-related tumors. These hypermethylation signatures were positively correlated with better patient prognostic outcomes. These findings may serve as a basis for the accurate identification of patients who are at greater risk of treatment failure, and distinguish them from patients with better prognostic indicators. Patients at a higher risk of treatment failure would benefit from multimodal treatment options, whereas patients with more favorable prognosis may benefit from lower toxicities associated with treatment de-escalation [101].

Sartor et al HPV-positive HNSCC overall were found to have greater DNA methylation, both in terms of targeted genic methylation and global methylation, compared to HPV-negative HNSCC. The group found higher promoter methylation of polycomb repressive complex 2 target genes and increased expression of *DNMT3A,* in HPV-positive cells, implicating this DNA

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

methyltransferase in the methylation patterns observed in HPV-positive HNSCC. *RUNX2, IRS-1* and *CCNA1* had higher methylation and lower expression in HPV-positive cells, whereas *CDKN2A* (the surrogate marker of HPV infection [48]) and *KRT8* had higher methylation and lower expression in HPV-negative cells [102].

Similar to Sartor et al, the independent group of Lleras et al reported a greater level of differential DNA methylation in HPV-positive OPC cases compared to HPV-negative cases, and identified that this increase was due in large part to hypermethylation of CpG islands in HPV-positive cases. There were 28 genes found in association with these methylation patterns in HPV-positive OPC. Some novel findings included the association of differential methylation patterns with *ALX4*, *CUTL2*, and *HOXA7*, as well as previously identified cancer-related genes such as *FBX039*, *IGSF4*, *PLOD2* and *SLITRK3*. Two genes identified in this study which were also reported in the Sartor group discussed above included: *CDKN2A*, and *HOXA7*. Notably, Lleras et al found increased promoter hypermethylation in HPV-positive cases, whereas Sartor found increased promoter hypermethylation in HPV-negative cases. A possible explanation for this discrepancy may be found in study design: Sartor et al excluded samples containing less than 40% tumor from subsequent analysis for nodal metastasis signature and HPV status, while Lleras et al included such samples in the initial identification of hyper- and hypo-methylated CpG loci [103].

Chromatin

Out of different functional chromatin domains, promoters and enhancers play a major role in gene expression regulation via recruitment of specific transcription factors (TFs) to DNA [104]. The enhancers may drive transcriptional dysregulation responsible for the development and progression of HPV-positive HNSCC at a genome-wide scale. [78, 104-110]. HPV-E6/E7 proteins were also found to interact with p300 and MYC, main constituents of enhancer regions [111-114].

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

These data suggest that targets of viral oncoproteins play the central role in activation of super-enhancers, the epigenetic regulators of gene expression.

Current literature suggests that super-enhancers at the site of HPV integration upregulate the expression of integrated HPV-E6/E7 oncogenes [79]. These overexpressed viral oncoproteins most likely activate their direct or indirect protein targets, which activate super-enhancer complexes genome-wide in disease-specific manner [115]. These disease-specific super-enhancers with unique components can be therapeutically targeted [115]. Indeed, the leading HNSCC oncoproteins NF-κB and STAT [116], were also found to participate in enhancer formation during Epstein Barr Virus-caused carcinogenesis [115] In addition, since E6/E7 lack enzymatic activities and don't directly bind to DNA they exert their functions by binding with a vast number of cellular proteins [117]. Published data suggest that direct or indirect targets of HPV E6/E7 affect the chromatin landscape of cancer cells [118-120]. Indeed, E7 induces the expression of KDM6B (demethylase of H3K27me3, the mark of repressive chromatin), which increases chromatin accessibility and enhances gene transcription [118]. HPV16 E7 activates HATs (histone acetyltransferases) and inactivates HDACs (histone deacetylases) [119, 120], leading to acetylation of histone tails and chromatin opening. Additionally, binding and inactivation of E2F6, Rb, and p130 by HPV E7 induces acetylation of H3K9 and methylation of H3K4, both of which are marks of open chromatin and active transcription [119, 121, 122]. These and other data suggest that reorganization and activation of the chromatin are primary mechanisms of HPV-related carcinogenesis.

VIII. CONCLUDING REMARKS

In this review, we briefly described the prevailing molecular alterations observed in HPV-positive HNSCC, a disease currently lacking effective targeted therapeutics. The current literature

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

suggests that viral oncoproteins lead to significant epigenetic cascades and that aberrant chromatin structure plays an important role in the integration of HPV into the host genome. A "chicken and the egg" question remains unanswered for the potentially interdependent phenomena of HPV integration and HNSCC aberrations in chromatin structure, and requires further mechanistic investigation. Furthermore, unlike cervical cancer in which most tumors exhibit viral integration, HPV-positive OPC often display episomal HPV DNA. Whether these observations represent a divergence in the mechanisms of HPV-related carcinogenesis in HNSCC also remains an unanswered question. Epigenetic dysregulation and APOBEC-signature mutations most likely drive genome wide dysregulation of the gene expression. Many of these genes, including ErbB-family proteins, represent promising targets for the development of tumor type-specific therapeutics. The significant progress in the understanding of HPV-related carcinogenesis in HNSCC reviewed here will serve as the foundations toward the development of novel diagnostic and therapeutic modalities to quell the burden imposed by this emerging viral malignancy.

CONFLICTS OF INTERESTS

The authors have no conflicts of interest to disclose.

ACKNOWLEDGEMENTS

D.A.G. receives funding from NIDCR, grant number R21-DE025398.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

REFERENCES

[1] Global Burden of Disease Cancer C, Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. JAMA Oncol 2017; 3:524-548.
[2] Carvalho AL, Nishimoto IN, Califano JA, Kowalski LP. Trends in incidence and prognosis for head and neck cancer in the United States: a site-specific analysis of the SEER database. Int J Cancer 2005;114:806-16.
[3] Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J Clin Oncol 2011;29:4294-301.
[4] Tiwana MS, Wu J, Hay J, Wong F, Cheung W, Olson RA. 25 year survival outcomes for squamous cell carcinomas of the head and neck: population-based outcomes from a Canadian province. Oral Oncol 2014;50:651-6.
[5] Chaturvedi AK, Anderson WF, Lortet-Tieulent J, Curado MP, Ferlay J, Franceschi S, et al. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. J Clin Oncol 2013;31:4550-9.
[6] D'Souza G, Agrawal Y, Halpern J, Bodison S, Gillison ML. Oral sexual behaviors associated with prevalent oral human papillomavirus infection. J Infect Dis 2009;199:1263-9.
[7] Panwar A, Batra R, Lydiatt WM, Ganti AK. Human papilloma virus positive oropharyngeal squamous cell carcinoma: a growing epidemic. Cancer Treat Rev 2014;40:215-9.
[8] Satterwhite CL, Torrone E, Meites E, Dunne EF, Mahajan R, Ocfemia MC, et al. Sexually transmitted infections among US women and men: prevalence and incidence estimates, 2008. Sex Transm Dis 2013;40:187-93.
[9] Plummer M, de Martel C, Vignat J, Ferlay J, Bray F, Franceschi S. Global burden of cancers attributable to infections in 2012: a synthetic analysis. Lancet Glob Health 2016;4:e609-16.
[10] Giuliano AR, Lu B, Nielson CM, Flores R, Papenfuss MR, Lee JH, et al. Age-specific prevalence, incidence, and duration of human papillomavirus infections in a cohort of 290 US men. J Infect Dis 2008;198:827-35.
[11] Fakhry C, Westra WH, Li S, Cmelak A, Ridge JA, Pinto H, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. J Natl Cancer Inst 2008;100:261-9.
[12] Gillison ML, D'Souza G, Westra W, Sugar E, Xiao W, Begum S, et al. Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. J Natl Cancer Inst 2008;100:407-20.
[13] Cancer Genome Atlas Research N. Comprehensive genomic characterization of squamous cell lung cancers. Nature 2012;489:519-25.
[14] Lydiatt WM, Patel SG, O'Sullivan B, Brandwein MS, Ridge JA, Migliacci JC, et al. Head and Neck cancers-major changes in the American Joint Committee on cancer eighth edition cancer staging manual. CA Cancer J Clin 2017;67:122-137.
[15] Kelly JR, Husain ZA, Burtness B. Treatment de-intensification strategies for head and neck cancer. Eur J Cancer 2016;68:125-33.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[16] Garcia-Vallve S, Alonso A, Bravo IG. Papillomaviruses: different genes have different histories. Trends Microbiol 2005;13:514-21.
[17] de Villiers EM. Cross-roads in the classification of papillomaviruses. Virology 2013;445:2-10.
[18] Bouvard V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, et al. A review of human carcinogens--Part B: biological agents. Lancet Oncol 2009;10:321-2.
[19] Anantharaman D, Gheit T, Waterboer T, Abedi-Ardekani B, Carreira C, McKay-Chopin S, et al. Human papillomavirus infections and upper aero-digestive tract cancers: the ARCAGE study. J Natl Cancer Inst 2013;105:536-45.
[20] Michaud DS, Langevin SM, Eliot M, Nelson HH, Pawlita M, McClean MD, et al. High-risk HPV types and head and neck cancer. Int J Cancer 2014;135:1653-61.
[21] Viens LJ, Henley SJ, Watson M, Markowitz LE, Thomas CC, Thompson TD, et al. Human Papillomavirus-Associated Cancers - United States, 2008-2012. MMWR Morb Mortal Wkly Rep 2016;65:661-6.
[22] Steinau M, Saraiya M, Goodman MT, Peters ES, Watson M, Cleveland JL, et al. Human papillomavirus prevalence in oropharyngeal cancer before vaccine introduction, United States. Emerg Infect Dis 2014;20:822-8.
[23] Kreimer AR, Clifford GM, Boyle P, Franceschi S. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. Cancer Epidemiol Biomarkers Prev 2005;14:467-75.
[24] Ndiaye C, Mena M, Alemany L, Arbyn M, Castellsague X, Laporte L, et al. HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis. Lancet Oncol 2014;15:1319-31.
[25] Doorbar J. The papillomavirus life cycle. J Clin Virol 2005;32 Suppl 1:S7-15.
[26] Doorbar J. Molecular biology of human papillomavirus infection and cervical cancer. Clin Sci (Lond) 2006;110:525-41.
[27] Kines RC, Thompson CD, Lowy DR, Schiller JT, Day PM. The initial steps leading to papillomavirus infection occur on the basement membrane prior to cell surface binding. Proc Natl Acad Sci U S A 2009;106:20458-63.
[28] Roberts JN, Buck CB, Thompson CD, Kines R, Bernardo M, Choyke PL, et al. Genital transmission of HPV in a mouse model is potentiated by nonoxynol-9 and inhibited by carrageenan. Nat Med 2007;13:857-61.
[29] Perry ME. The specialised structure of crypt epithelium in the human palatine tonsil and its functional significance. J Anat 1994;185 (Pt 1):111-27.
[30] Westra WH. The morphologic profile of HPV-related head and neck squamous carcinoma: implications for diagnosis, prognosis, and clinical management. Head Neck Pathol 2012;6 Suppl 1:S48-54.
[31] Paz IB, Cook N, Odom-Maryon T, Xie Y, Wilczynski SP. Human papillomavirus (HPV) in head and neck cancer. An association of HPV 16 with squamous cell carcinoma of Waldeyer's tonsillar ring. Cancer 1997;79:595-604.
[32] Joyce JG, Tung JS, Przysiecki CT, Cook JC, Lehman ED, Sands JA, et al. The L1 major capsid protein of human papillomavirus type 11 recombinant virus-like particles interacts with heparin and cell-surface glycosaminoglycans on human keratinocytes. J Biol Chem 1999;274:5810-22.
[33] DiGiuseppe S, Bienkowska-Haba M, Guion LG, Sapp M. Cruising the cellular highways: How human papillomavirus travels from the surface to the nucleus. Virus Res 2017;231:1-9.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[34] Sapp M, Bienkowska-Haba M. Viral entry mechanisms: human papillomavirus and a long journey from extracellular matrix to the nucleus. FEBS J 2009;276:7206-16.
[35] Pyeon D, Pearce SM, Lank SM, Ahlquist P, Lambert PF. Establishment of human papillomavirus infection requires cell cycle progression. PLoS Pathog 2009;5:e1000318.
[36] Stanley M. Immune responses to human papillomavirus. Vaccine 2006;24 Suppl 1:S16-22.
[37] Sherman L, Jackman A, Itzhaki H, Stoppler MC, Koval D, Schlegel R. Inhibition of serum- and calcium-induced differentiation of human keratinocytes by HPV16 E6 oncoprotein: role of p53 inactivation. Virology 1997;237:296-306.
[38] Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell 1990;63:1129-36.
[39] Bunz F, Dutriaux A, Lengauer C, Waldman T, Zhou S, Brown JP, et al. Requirement for p53 and p21 to sustain G2 arrest after DNA damage. Science 1998;282:1497-501.
[40] Taylor WR, Stark GR. Regulation of the G2/M transition by p53. Oncogene 2001;20:1803-15.
[41] Hollstein M, Sidransky D, Vogelstein B, Harris CC. p53 mutations in human cancers. Science 1991;253:49-53.
[42] Castellsague X, Alemany L, Quer M, Halec G, Quiros B, Tous S, et al. HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. J Natl Cancer Inst 2016;108:djv403.
[43] Werness BA, Levine AJ, Howley PM. Association of human papillomavirus types 16 and 18 E6 proteins with p53. Science 1990;248:76-9.
[44] Boyer SN, Wazer DE, Band V. E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. Cancer Res 1996;56:4620-4.
[45] Ikeda MA, Jakoi L, Nevins JR. A unique role for the Rb protein in controlling E2F accumulation during cell growth and differentiation. Proc Natl Acad Sci U S A 1996;93:3215-20.
[46] Nevins JR. E2F: a link between the Rb tumor suppressor protein and viral oncoproteins. Science 1992;258:424-9.
[47] Khleif SN, DeGregori J, Yee CL, Otterson GA, Kaye FJ, Nevins JR, et al. Inhibition of cyclin D-CDK4/CDK6 activity is associated with an E2F-mediated induction of cyclin kinase inhibitor activity. Proc Natl Acad Sci U S A 1996;93:4350-4.
[48] Riethdorf S, Neffen EF, Cviko A, Loning T, Crum CP, Riethdorf L. p16INK4A expression as biomarker for HPV 16-related vulvar neoplasias. Hum Pathol 2004;35:1477-83.
[49] Munger K, Werness BA, Dyson N, Phelps WC, Harlow E, Howley PM. Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. EMBO J 1989;8:4099-105.
[50] Barbosa MS, Edmonds C, Fisher C, Schiller JT, Lowy DR, Vousden KH. The region of the HPV E7 oncoprotein homologous to adenovirus E1a and Sv40 large T antigen contains separate domains for Rb binding and casein kinase II phosphorylation. EMBO J 1990;9:153-60.
[51] Guo T, Gaykalova DA, Considine M, Wheelan S, Pallavajjala A, Bishop JA, et al. Characterization of functionally active gene fusions in human papillomavirus related oropharyngeal squamous cell carcinoma. Int J Cancer 2016;139:373-82.
[52] Guirimand T, Delmotte S, Navratil V. VirHostNet 2.0: surfing on the web of virus/host molecular interactions data. Nucleic Acids Res 2015;43:D583-7.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[53] Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 2015;517:576-82.
[54] Laplante M, Sabatini DM. mTOR signaling in growth control and disease. Cell 2012;149:274-93.
[55] Johnson DG, Schwarz JK, Cress WD, Nevins JR. Expression of transcription factor E2F1 induces quiescent cells to enter S phase. Nature 1993;365:349-52.
[56] Sellers WR, Rodgers JW, Kaelin WG, Jr. A potent transrepression domain in the retinoblastoma protein induces a cell cycle arrest when bound to E2F sites. Proc Natl Acad Sci U S A 1995;92:11544-8.
[57] Zhang Q, Lenardo MJ, Baltimore D. 30 Years of NF-kappaB: A Blossoming of Relevance to Human Pathobiology. Cell 2017;168:37-57.
[58] Sun SC. CYLD: a tumor suppressor deubiquitinase regulating NF-kappaB activation and diverse biological processes. Cell Death Differ 2010;17:25-34.
[59] Hajek M, Sewell A, Kaech S, Burtness B, Yarbrough WG, Issaeva N. TRAF3/CYLD mutations identify a distinct subset of human papilloma virus-associated head and neck squamous cell carcinoma. Cancer 2017; doi: 10.1002/cncr.30570. Ahead of print.
[60] Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006;126:663-76.
[61] Keysar SB, Le PN, Miller B, Jackson BC, Eagles JR, Nieto C, et al. Regulation of Head and Neck Squamous Cancer Stem Cells by PI3K and SOX2. J Natl Cancer Inst 2017;109:1-12.
[62] Hayes DN, Van Waes C, Seiwert TY. Genetic Landscape of Human Papillomavirus-Associated Head and Neck Cancer and Comparison to Tobacco-Related Tumors. J Clin Oncol 2015;33:3227-34.
[63] The Cancer Genome Atlas Research N. Integrated genomic and molecular characterization of cervical cancer. Nature 2017; 16;543(7645):378-384.
[64] Henderson S, Chakravarthy A, Su X, Boshoff C, Fenton TR. APOBEC-mediated cytosine deamination links PIK3CA helical domain mutations to human papillomavirus-driven tumor development. Cell Rep 2014;7:1833-41.
[65] Henderson S, Fenton T. APOBEC3 genes: retroviral restriction factors to cancer drivers. Trends Mol Med 2015;21:274-84.
[66] Seiwert TY, Zuo Z, Keck MK, Khattri A, Pedamallu CS, Stricker T, et al. Integrative and comparative genomic analysis of HPV-positive and HPV-negative head and neck squamous cell carcinomas. Clin Cancer Res 2015;21:632-41.
[67] Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat Genet 2013;45:970-6.
[68] Wentzensen N, Vinokurova S, von Knebel Doeberitz M. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. Cancer Res 2004;64:3878-84.
[69] Tommasino M. The human papillomavirus family and its role in carcinogenesis. Semin Cancer Biol 2014;26:13-21.
[70] Vinokurova S, Wentzensen N, Kraus I, Klaes R, Driesch C, Melsheimer P, et al. Type-dependent integration frequency of human papillomavirus genomes in cervical lesions. Cancer Res 2008;68:307-13.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[71] Olthof NC, Huebbers CU, Kolligs J, Henfling M, Ramaekers FC, Cornet I, et al. Viral load, gene expression and mapping of viral integration sites in HPV16-associated HNSCC cell lines. Int J Cancer 2015;136:E207-18.
[72] Olthof NC, Speel EJ, Kolligs J, Haesevoets A, Henfling M, Ramaekers FC, et al. Comprehensive analysis of HPV16 integration in OSCC reveals no significant impact of physical status on viral oncogene and virally disrupted human gene expression. PloS one 2014;9:e88718.
[73] Parfenov M, Pedamallu CS, Gehlenborg N, Freeman SS, Danilova L, Bristow CA, et al. Characterization of HPV and host genome interactions in primary head and neck cancers. Proc Natl Acad Sci U S A 2014;111:15544-9.
[74] Schmitz M, Driesch C, Jansen L, Runnebaum IB, Durst M. Non-random integration of the HPV genome in cervical cancer. PloS one 2012;7:e39632.
[75] Adams AK, Wise-Draper TM, Wells SI. Human papillomavirus induced transformation in cervical and head and neck cancers. Cancers (Basel) 2014;6:1793-820.
[76] Richards RI. Fragile and unstable chromosomes in cancer: causes and consequences. Trends Genet 2001;17:339-45.
[77] Jeon S, Allen-Hoffmann BL, Lambert PF. Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. J Virol 1995;69:2989-97.
[78] Wilson GA, Lechner M, Koferle A, Caren H, Butcher LM, Feber A, et al. Integrated virus-host methylome analysis in head and neck squamous cell carcinoma. Epigenetics 2013;8:953-61.
[79] Dooley KE, Warburton A, McBride AA. Tandemly Integrated HPV16 Can Form a Brd4-Dependent Super-Enhancer-Like Element That Drives Transcription of Viral Oncogenes. MBio 2016;7: e01446-16.
[80] You J, Croyle JL, Nishimura A, Ozato K, Howley PM. Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes. Cell 2004;117:349-60.
[81] McPhillips MG, Oliveira JG, Spindler JE, Mitra R, McBride AA. Brd4 is required for e2-mediated transcriptional activation but not genome partitioning of all papillomaviruses. J Virol 2006;80:9530-43.
[82] Baxter MK, McPhillips MG, Ozato K, McBride AA. The mitotic chromosome binding activity of the papillomavirus E2 protein correlates with interaction with the cellular chromosomal protein, Brd4. J Virol 2005;79:4806-18.
[83] Hu Z, Zhu D, Wang W, Li W, Jia W, Zeng X, et al. Genome-wide profiling of HPV integration in cervical cancer identifies clustered genomic hot spots and a potential microhomology-mediated integration mechanism. Nat Genet 2015;47:158-63.
[84] Song H, Dicks E, Ramus SJ, Tyrer JP, Intermaggio MP, Hayward J, et al. Contribution of Germline Mutations in the RAD51B, RAD51C, and RAD51D Genes to Ovarian Cancer in the Population. J Clin Oncol 2015;33:2901-7.
[85] Kumar-Sinha C, Tomlins SA, Chinnaiyan AM. Recurrent gene fusions in prostate cancer. Nat Rev Cancer 2008;8:497-511.
[86] Frigola X, Inman BA, Lohse CM, Krco CJ, Cheville JC, Thompson RH, et al. Identification of a soluble form of B7-H1 that retains immunosuppressive activity and is associated with aggressive renal cell carcinoma. Clin Cancer Res 2011;17:1915-23.
[87] Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, et al. The mutational landscape of head and neck squamous cell carcinoma. Science 2011;333:1157-60.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[88] Ke N, Claassen G, Yu DH, Albers A, Fan W, Tan P, et al. Nuclear hormone receptor NR4A2 is involved in cell transformation and apoptosis. Cancer Res 2004;64:8208-12.
[89] Walline HM, Goudsmit CM, McHugh JB, Tang AL, Owen JH, Teh BT, et al. Integration of high-risk human papillomavirus into cellular cancer-related genes in head and neck cancer cell lines. Head Neck 2017;39:840-852.
[90] Cousineau I, Abaji C, Belmaaza A. BRCA1 regulates RAD51 function in response to DNA damage and suppresses spontaneous sister chromatid replication slippage: implications for sister chromatid cohesion, genome stability, and carcinogenesis. Cancer Res 2005;65:11384-91.
[91] Keck MK, Zuo Z, Khattri A, Stricker TP, Brown CD, Imanguli M, et al. Integrative analysis of head and neck cancer identifies two biologically distinct HPV and three non-HPV subtypes. Clin Cancer Res 2015;21:870-81.
[92] Subramaniam D, He AR, Hwang J, Deeken J, Pishvaian M, Hartley ML, et al. Irreversible multitargeted ErbB family inhibitors for therapy of lung and breast cancer. Curr Cancer Drug Targets 2015;14:775-93.
[93] Androutsopoulos G, Adonakis G, Liava A, Ravazoula P, Decavalas G. Expression and potential role of ErbB receptors in type II endometrial cancer. Eur J Obstet Gynecol Reprod Biol 2013;168:204-8.
[94] Pollock NI, Wang L, Wallweber G, Gooding WE, Huang W, Chenna A, et al. Increased Expression of HER2, HER3, and HER2:HER3 Heterodimers in HPV-Positive HNSCC Using a Novel Proximity-Based Assay: Implications for Targeted Therapies. Clin Cancer Res 2015;21:4597-606.
[95] Machiels JP, Galot R, Schmitz S. HER2 and HER3 in HPV+ and HPV- HNSCC--Letter. Clin Cancer Res 2016;22:1825.
[96] Machiels JP, Haddad RI, Fayette J, Licitra LF, Tahara M, Vermorken JB, et al. Afatinib versus methotrexate as second-line treatment in patients with recurrent or metastatic squamous-cell carcinoma of the head and neck progressing on or after platinum-based therapy (LUX-Head & Neck 1): an open-label, randomised phase 3 trial. Lancet Oncol 2015;16:583-94.
[97] Brand TM, Hartmann S, Bhola NE, Peyser ND, Li H, Zeng Y, et al. Human Papillomavirus Regulates HER3 Expression in Head and Neck Cancer: Implications for Targeted HER3 Therapy in HPV+ Patients. Clin Cancer Res 2016; doi: 10.1158/1078-0432.CCR-16-2203. Ahead of print.
[98] Chung CH, Germain A, Subramaniam RM, Heilmann AM, Fedorchak K, Ali SM, et al. Genomic alterations in human epidermal growth factor receptor 2 (HER2/ERBB2) in head and neck squamous cell carcinoma. Head Neck 2017;39:E15-E9.
[99] Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. Carcinogenesis 2010;31:27-36.
[100] Ha PK, Califano JA. Promoter methylation and inactivation of tumour-suppressor genes in oral squamous-cell carcinoma. Lancet Oncol 2006;7:77-82.
[101] Kostareli E, Holzinger D, Bogatyrova O, Hielscher T, Wichmann G, Keck M, et al. HPV-related methylation signature predicts survival in oropharyngeal squamous cell carcinomas. J Clin Invest 2013;123:2488-501.
[102] Sartor MA, Dolinoy DC, Jones TR, Colacino JA, Prince ME, Carey TE, et al. Genome-wide methylation and expression differences in HPV(+) and HPV(-) squamous cell carcinoma cell lines are consistent with divergent mechanisms of carcinogenesis. Epigenetics 2011;6:777-87.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[103] Lleras RA, Smith RV, Adrien LR, Schlecht NF, Burk RD, Harris TM, et al. Unique DNA methylation loci distinguish anatomic site and HPV status in head and neck squamous cell carcinoma. Clin Cancer Res 2013;19:5444-55.
[104] Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 2009;459:108-12.
[105] Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 2013;153:307-19.
[106] Hnisz D, Weintraub AS, Day DS, Valton AL, Bak RO, Li CH, et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. Science 2016;351:1454-8.
[107] Kulaeva OI, Nizovtseva EV, Polikanov YS, Ulianov SV, Studitsky VM. Distant activation of transcription: mechanisms of enhancer action. Mol Cell Biol 2012;32:4892-7.
[108] Heyn H, Vidal E, Ferreira HJ, Vizoso M, Sayols S, Gomez A, et al. Epigenomic analysis detects aberrant super-enhancer DNA methylation in human cancer. Genome Biol 2016;17:11.
[109] Hon GC, Rajagopal N, Shen Y, McCleary DF, Yue F, Dang MD, et al. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. Nat Genet 2013;45:1198-206.
[110] Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 2012;485:376-80.
[111] Veldman T, Liu X, Yuan H, Schlegel R. Human papillomavirus E6 and Myc proteins associate in vivo and bind to and cooperatively activate the telomerase reverse transcriptase promoter. Proc Natl Acad Sci U S A 2003;100:8211-6.
[112] Xie X, Piao L, Bullock BN, Smith A, Su T, Zhang M, et al. Targeting HPV16 E6-p300 interaction reactivates p53 and inhibits the tumorigenicity of HPV-positive head and neck squamous cell carcinoma. Oncogene 2014;33:1037-46.
[113] Patel D, Huang SM, Baglia LA, McCance DJ. The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. EMBO J 1999;18:5061-72.
[114] Miller DL, Puricelli MD, Stack MS. Virology and molecular pathogenesis of HPV (human papillomavirus)-associated oropharyngeal squamous cell carcinoma. Biochem J 2012;443:339-53.
[115] Zhou H, Schmidt SC, Jiang S, Willox B, Bernhardt K, Liang J, et al. Epstein-Barr virus oncoprotein super-enhancers control B cell growth. Cell Host Microbe 2015;17:205-16.
[116] Gaykalova DA, Manola JB, Ozawa H, Zizkova V, Morton K, Bishop JA, et al. NF-kappaB and stat3 transcription factor signatures differentiate HPV-positive and HPV-negative head and neck squamous cell carcinoma. Int J Cancer 2015;137:1879-89.
[117] Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. Nat Rev Cancer 2010;10:550-60.
[118] McLaughlin-Drubin ME, Crum CP, Munger K. Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming. Proc Natl Acad Sci U S A 2011;108:2130-5.
[119] Zhang B, Laribee RN, Klemsz MJ, Roman A. Human papillomavirus type 16 E7 protein increases acetylation of histone H3 in human foreskin keratinocytes. Virology 2004;329:189-98.
[120] Bodily JM, Mehta KP, Laimins LA. Human papillomavirus E7 enhances hypoxia-inducible factor 1-mediated transcription by inhibiting binding of histone deacetylases. Cancer Res 2011;71:1187-95.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[121] Zhang B, Chen W, Roman A. The E7 proteins of low- and high-risk human papillomaviruses share the ability to target the pRB family member p130 for degradation. Proc Natl Acad Sci U S A 2006;103:437-42.
[122] Attwooll C, Oddi S, Cartwright P, Prosperini E, Agger K, Steensgaard P, et al. A novel repressive E2F6 complex containing the polycomb group protein, EPC1, that interacts with EZH2 in a proliferation-specific manner. J Biol Chem 2005;280:1199-208.

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

FIGURE LEGENDS

Figure 1: The HPV16 genome

Schematic of HPV16 genome designed from Ref Seq entry: NC_001526. Viral genes E6 and E7 are marked in light red; E1, E2, E4, E5 in purple; and L1 and L2 in blue. Green regions marked E1, E2 are binding sites for the corresponding proteins. LCR: long control region; TSS_E: transcription start site for viral early genes; TSS_L: transcription start site for viral late genes; TATA in red represents the TATA box transcription factor binding site. SP1 represents the binding site for the host SP1 transcription factor. Although numerical base markers represent viral genomic positions, the coding elements are not to scale.

Figure 2: Tumor progression model for HPV-related carcinogenesis.

(A) In the head and neck, HPV demonstrates tropism for lymphoid-associated structure of the oropharynx, including the palatine and lingual tonsils. (B) In the oropharynx, HPV gains access to basal keratinocyte progenitors through fenestrations in the reticulated epithelium of the tonsillar crypts. (C) Infection of the tonsillar epithelium result aberrant basal cell differentiation, dysplasia, carcinoma in situ, and finally invasive carcinoma. (D) Hypothetical model for somatic mutations in a multistage tumor progression model. Genes and loci in red are upregulated or activated or show increase in copy number. Genes in green undergo loss of function mutation or deletion.

Figure 3: Cell cycle perturbation by human papillomavirus.

(A) Schematic representing physiologic cell cycle regulation. Tumor suppressor proteins are marked blue, proto-oncoproteins are marked red. The *CDKN2A* gene encodes $p14^{ARF}$ and $p16^{INK4A}$ proteins. $p14^{ARF}$ inhibits MDM2, thereby disinhibiting p53 to activate p21 and stop progression through the G2/M checkpoint into mitosis. $p16^{INK4A}$ inhibits the CyclinD1/CDK4 and CyclinD1/CDK6 complexes. These complexes catalyze phosphorylation of the retinoblastoma

bioRxiv preprint doi: https://doi.org/10.1101/137653; this version posted May 16, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

protein (Rb), inducing it to release E2F family transcription factors to enter the nucleus and activate transcription of S-phase promoting transcripts. Phosphorylation of Rb also results in feedback inhibition of $p16^{INK4A}$ expression. (B) HPV E6 oncoprotein binds and targets p53 for degradation, resulting in loss of G2/M checkpoint regulation. HPV E7 oncoprotein binds and targets Rb for degradation, resulting in the nuclear translocation of E2F and promotion of S-phase transition. In addition, downregulation of Rb results loss of feedback inhibition and overexpression of $p16^{INK4A}$.

Figure 4: Model for HPV integration promoting tumorigenesis.

(A) Breakpoints in the viral genome often occur in the E2 gene. (B) Through interruption of the E2 gene, viral linearization leads to unregulated expression of E6 and E7 viral oncogenes, represented by the bold arrow at TSS_E. (C) The linearized viral genome inserts into a segment of open chromatin. Although this schematic shows viral DNA integrating at an enhancer, it may also integrate at genic or other transcriptionally active loci. (D) Viral DNA integration may result in the disruption in the expression of nearby genes. In a simplified example depicted here, integration of the viral genome upstream of a host tumor suppressor gene results in its transcriptional inhibition. However, viral integration into gene introns has also been demonstrated gene alter gene expression and alternative splicing [73].

Table 1: HPV genotypes identified in HNSCC

High-risk	Low-risk
Genotype	Genotype
16	11
18	32
31	44
33	53
45	57
51	81
52	
56	
58	
59	

Figure 1

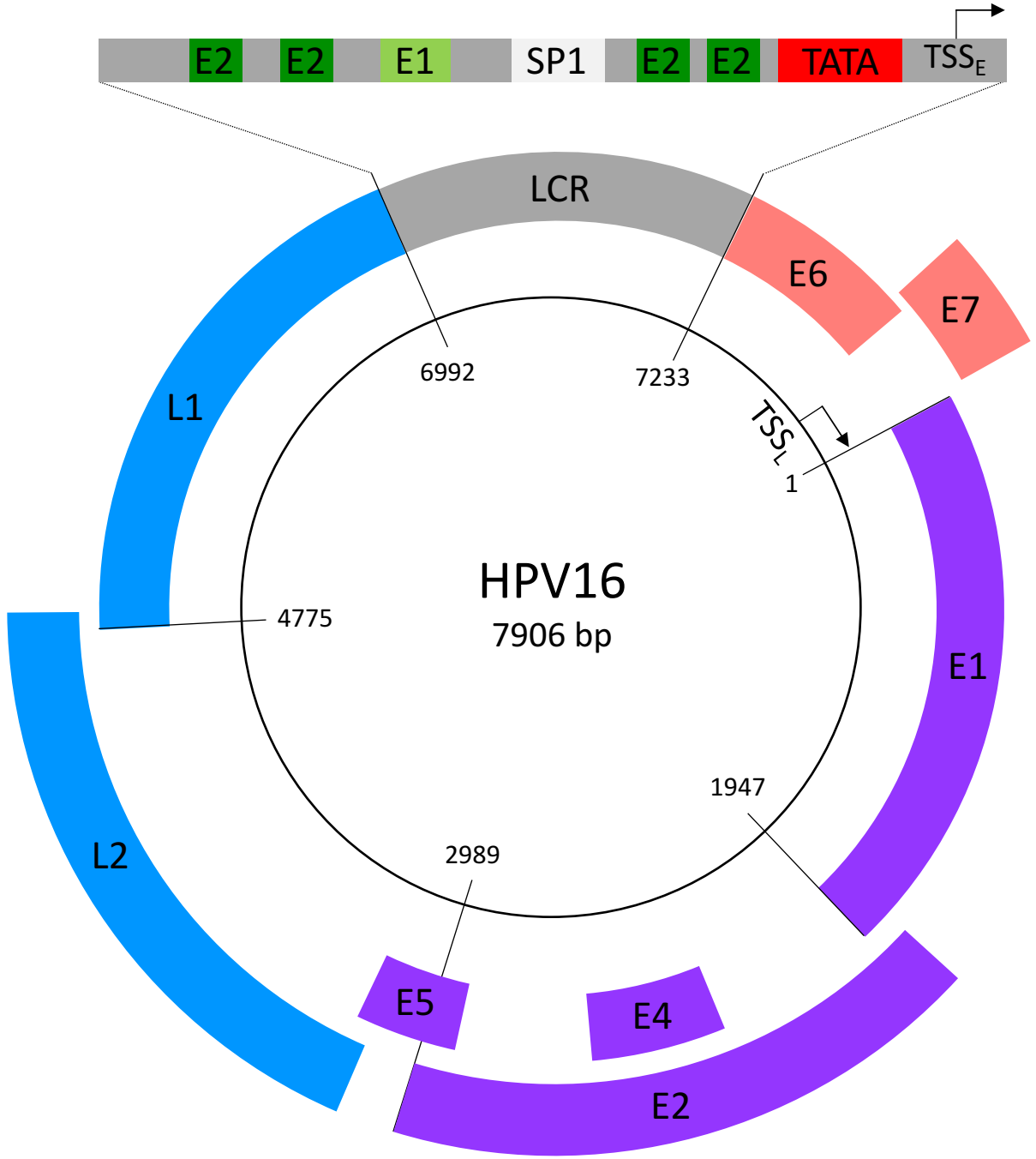
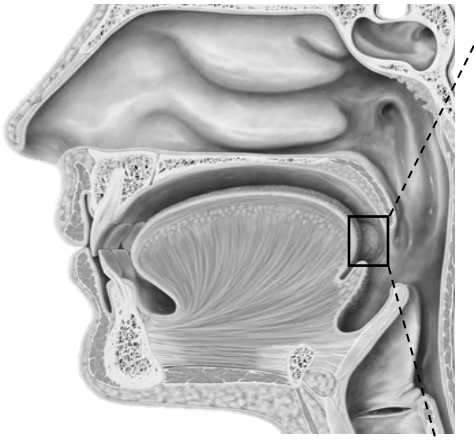
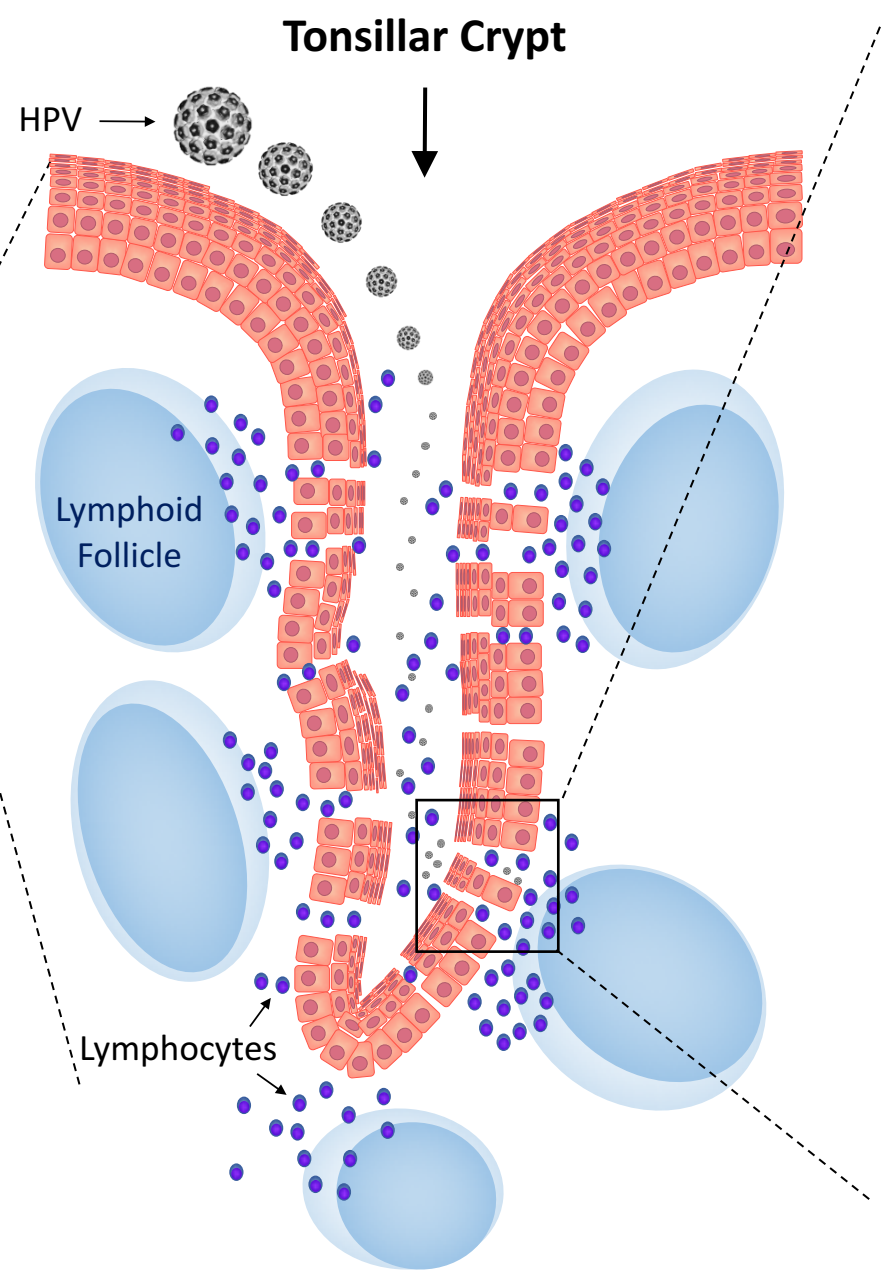
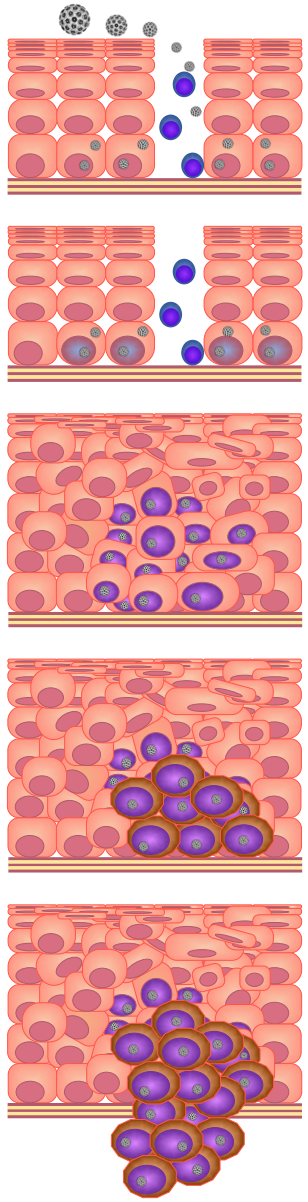

E2
E2
E1
SP1
E2
E2
TATA
TSS$_E$
LCR
E6
E7
L1
6992
7233
TSS$_L$
1
HPV16
7906 bp
4775
E1
1947
L2
2989
E5
E4
E2

Figure 2







CNV

Mutation

Infection of Tonsillar Crypt Epithelium

E6
E7

Aberrant Basal Stem Cell Differentiation

E2F1

Dysplasia

TRAF3
FHIT
PTEN

Carcinoma in situ

3q26/28

PIK3CA
SOX2
TP63

Invasive Carcinoma

Figure 3

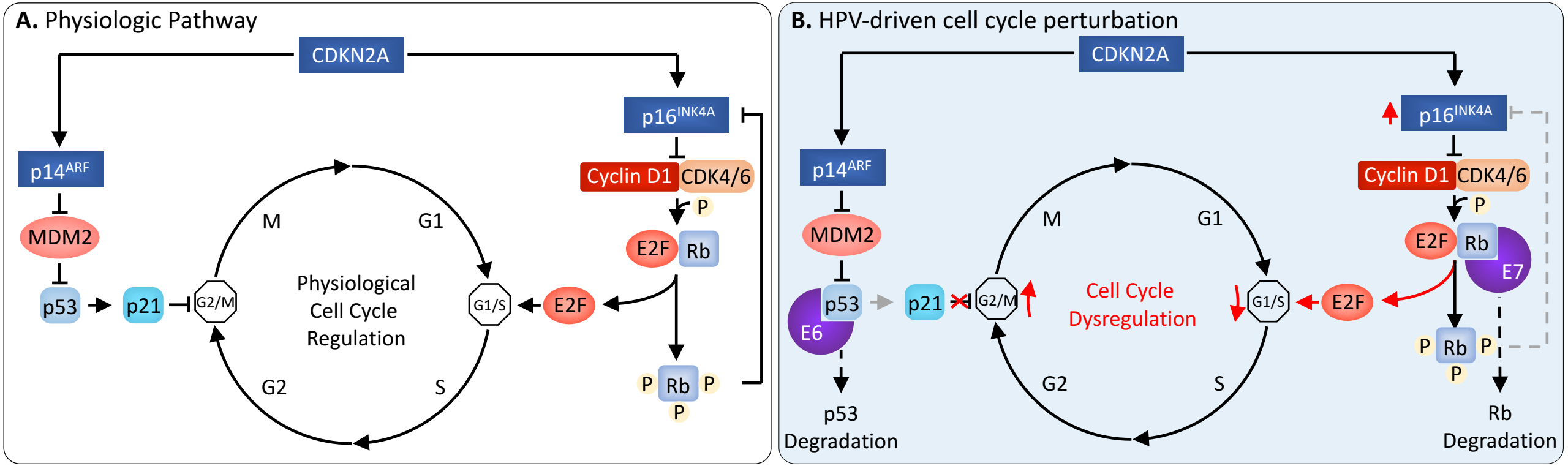
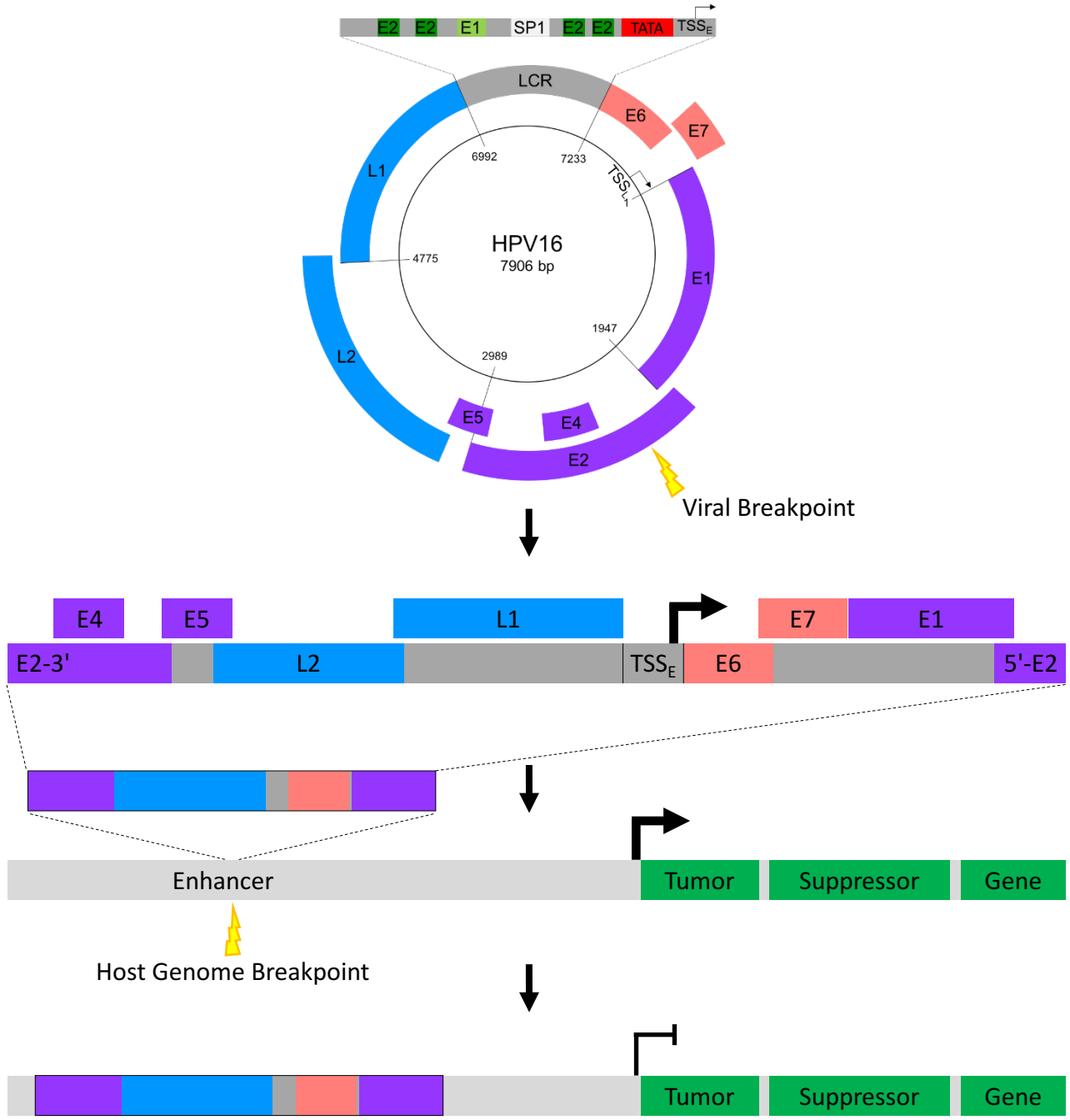

A. Physiologic Pathway
CDKN2A
p14ARF
MDM2
p53
p21
G2/M
M
G1
Physiological
Cell Cycle
Regulation
G1/S
E2F
G2
S
p16INK4A
Cyclin D1
CDK4/6
P
E2F
Rb
P
Rb
P
P
B. HPV-driven cell cycle perturbation
CDKN2A
p14ARF
MDM2
p53
E6
p53
Degradation
p21
G2/M
M
G1
Cell Cycle
Dysregulation
G1/S
G2
S
E2F
p16INK4A
Cyclin D1
CDK4/6
P
E2F
Rb
E7
P
Rb
P
P
Rb
Degradation

Figure 4


E2
E2
E1
SP1
E2
E2
TATA
TSS$_E$
LCR
E6
E7
L1
6992
7233
TSS$_{L1}$
HPV16
7906 bp
4775
E1
1947
L2
2989
E5
E4
E2
Viral Breakpoint
E4
E5
L1
E7
E1
E2-3'
L2
TSS$_E$
E6
5'-E2
Enhancer
Tumor
Suppressor
Gene
Host Genome Breakpoint
Tumor
Suppressor
Gene