

Article

Direct Comparison of HPV16 Viral Genomic Integration, Copy Loss, and Structural Variants in Oropharyngeal and Uterine Cervical Cancers Reveal Distinct Relationships to E2 Disruption and Somatic Alteration

Travis P. Schrank ^{1,*}, Sulgi Kim ^{1,2}, Hina Rehmani ^{1,2}, Aditi Kothari ^{1,2}, Di Wu ^{3,4}, Wendell G. Yarbrough ^{1,2,5} 
and Natalia Issaeva ^{1,2,5} 

¹ Department of Otolaryngology/Head and Neck Surgery, School of Medicine, The University of North Carolina, 170 Manning Drive, Chapel Hill, NC 27599, USA

² Lineberger Cancer Center, UNC, Chapel Hill, NC 27599, USA

³ Department of Biostatistics, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

⁴ Division of Oral and Craniofacial Health Sciences, Adams School of Dentistry, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

⁵ Department of Pathology and Lab Medicine, UNC, Chapel Hill, NC 27599, USA

* Correspondence: travis_schrank@med.unc.edu; Tel.: +1-919-966-8926



Citation: Schrank, T.P.; Kim, S.; Rehmani, H.; Kothari, A.; Wu, D.; Yarbrough, W.G.; Issaeva, N. Direct Comparison of HPV16 Viral Genomic Integration, Copy Loss, and Structural Variants in Oropharyngeal and Uterine Cervical Cancers Reveal Distinct Relationships to E2 Disruption and Somatic Alteration. *Cancers* **2022**, *14*, 4488. <https://doi.org/10.3390/cancers14184488>

Academic Editors: Paula A. Oliveira and Rui M. Gil da Costa

Received: 22 July 2022

Accepted: 13 September 2022

Published: 16 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Simple Summary: HPV16 causes approximately 60% of uterine cervical cancer and 95% of HPV-driven oropharynx cancers. Despite both being HPV-associated, these tumors are very different. We directly compare integration of the HPV16 genome into the human genome in these two diseases, finding that the viral gene E2 is frequently lost in cervical cancer, but usually maintained in oropharyngeal cancer. We also found that oropharyngeal cancers with integration have many more integration sites per tumor and these more frequently occur in genomic regions with a high density of genes.

Abstract: Squamous cell carcinoma of the oropharynx caused by HPV type 16 (HPV16+ OPSCC) is the most common HPV-associated malignancy in the USA and has many molecular differences from uterine cervical squamous cell carcinoma (UCSCC). Our understanding of HPV oncogenesis relied on studies of UCSCC revealing a consensus model reliant on HPV integration with a loss of E2. Here, we compare patterns of HPV integration in UCSCC and OPSCC by analysis of affinity capture sequencing of the HPV16 genome in 104 OPSCC and 44 UCSCC tumors. These cohorts were contemporaneously sequenced using an identical strategy. Integration was identified using discordant read pair clustering and assembly-based approaches. Viral integration sites, structural variants, and copy losses were examined. While large-scale deep losses of HPV16 genes were common in UCSCC and were associated with E2 loss, deep copy losses of the HPV16 genome were infrequent in HPV16+ OPSCC. Similarly, structural variants within HPV16 favored E2 loss in UCSCC but not OPSCC. HPV16 integration sites were non-random, with recurrent integration hot-spots identified. OPSCC tumors had many more integration sites per tumor when compared to UCSCC and had more integration sites in genomic regions with high gene density. These data show that viral integration and E2 disruption are distinct in UCSCC and OPSCC. Our findings also add to growing literature suggesting that HPV tumorigenesis in OPSCC does not follow the model developed based on UCSCC.

Keywords: HPV16; oropharynx; squamous cell carcinoma; uterine cervix; integration

1. Introduction

Papillomaviruses are species-specific, non-enveloped, double-stranded DNA viruses with an about 8kb circular genome protected by a 55 nm icosahedral capsid [1–4]. Human

papillomaviruses (HPV) are associated with benign and malignant epithelial lesions. Subsets of HPV are high-risk and are causative agents of cervical cancer [5,6] and the majority of oropharyngeal squamous cell carcinomas (OPSCC) [7–10]. Both cervical and oropharyngeal cancers share universal risk factors that involve sexual behavior and cigarette smoking.

Cervical cancer, the second most common cancer in woman and with a gradually decreasing incident rate in the United States, is usually associated with HPV. Development of cervical cancer is a multistep process that includes initial HPV infection followed by persistent infection resulting in a relatively well-established progression from pre-cancerous cervical intraepithelial neoplasia (CIN) to malignant lesions [11]. In early lesions, the HPV genome remains in an episomal state, but with advancement to higher grade lesions, HPV DNA frequently integrates into the host genome, facilitating cancer development [12,13]. Integration accelerates carcinogenesis through frequent disruption of the HPV E2 gene, causing increased expression of the major oncogenes, E6 and E7, that directly inhibit the key human tumor suppressors, p53 and Rb [14,15]. Maintenance of HPV oncogenes is critical, but other HPV genes are variably lost [16].

As the opposite of cervical cancer because it is constantly increasing in the USA, oropharyngeal cancer is detected mostly in men, with male-to-female ratio of about 4:1, and roughly 80% of this disease is driven by HPV. HPV tumorigenesis in the oropharynx has not been as intensively investigated, and it most likely differs from cervical lesions due to unique features of tonsillar epithelia. The epithelia of the lingual and palatine tonsils, the regions where HPV-associated cancer arises [17], are lymphoid-associated epithelia with specialized crypts that facilitate antigen presentation. Tonsils are well recognized sites for pathogenic viral infections, including Epstein–Barr virus [18], adenoviruses, influenza and parainfluenza viruses, herpes simplex virus, and respiratory syncytial virus [19]. Natural susceptibility to a viral infection of tonsils most likely explains why HPV-associated head and neck cancer nearly always arises in this anatomical site. In contrast to cervical cancer, defining particular steps of HPV-carcinogenesis in tonsillar epithelia has been hampered by the absence of clinically detectable pre-malignant lesions and established screening procedures [11].

In addition to epidemiological differences, although initiated by the infection of HPV, cervical and oropharyngeal malignancies are characterized by several biological and clinical dissimilarities that highlight distinct HPV-mediated carcinogenesis [20]. First, HPV16 and HPV18 account for approximately 75% of uterine cervical tumors, but HPV type 16 is responsible for roughly 92% of OPSCC, and HPV type 18 is extremely rare in OPSCC [21]. Second, HPV-associated OPSCC has distinct cellular gene and protein expression profiles compared to cervical cancer [22,23]. Third, patients with HPV-positive OPSCC generally respond better to treatment than patients with HPV-positive cervical cancer [11]. Fourth, HPV-positive OPSCCs harbor a different somatic mutation spectrum than cervical cancer [24,25]. Lastly, fifth, ~40% of oropharyngeal cancers display constitutively active NF- κ B, mostly due to genetic inactivation of *TRAF3* and *CYLD* genes that are rare in uterine cervical cancer [26]. The NF- κ B active subtype of OPSCC lacks HPV integration and has improved survival compared to OPSCC with low NF- κ B activity [27–30]. Moreover, a growing amount of data point out a remarkable diversity of tumor microenvironment with increased CD4+:CD8+ T cell ratio in oropharyngeal as compared to cervical cancer [31].

Both integrated and episomal HPV DNA occur in OPSCC and cervical malignancies, with some tumors containing only integrated HPV DNA, others containing only episomal DNA, and a third group containing both [25,32–35]. Since integration is implicated in HPV carcinogenesis and is found in more than 50% of both cervical and oropharyngeal cancers, recent studies have focused on mapping integration sites and exploring how integration contributes to carcinogenesis. Key findings from these studies include a wide number of mapped integration sites, recurrent integration hotspots, and discoveries of accompanying complex structural genomic rearrangements associated with modified expression of affected genes [25,27]. In addition, a recent study found non-chromosomal circular viral/human DNA with oncogenic properties in OPSCC [36].

Important differences between HPV driven oropharyngeal and uterine cervical carcinomas continue to emerge; however, no comprehensive study has compared HPV integration sites and HPV structural genomic aberrations present in these two cancers. To define similarities and differences, we used identical ultra-deep targeted sequencing of HPV16 from 104 oropharyngeal and 44 uterine cervical tumors. Here, we present an in-depth investigation that reveals drastic dissimilarities in HPV integrations and structural genomic aberrations in OPSCC and UCSCC. These findings emphasize their individual biology and suggest that HPV-driven carcinogenesis is different between these tumor types.

2. Materials and Methods

Human Subjects—DNA sequencing of data was performed as a part of the UNCseq tumor sequencing program. Tumor sample identifiers and genomic data were derived from the clinical trial LCCC1108: *Development of a Tumor Molecular Analyses Program and Its Use to Support Treatment Decisions*. This IRB-approved trial opened in 2011. All studies were done with the approval of our Institutional Review Board, patient participation required written informed consent, and all studies were conducted in accordance with recognized ethical guidelines as described in U.S. Common Rule. By means of a chart review of electronic medical records, demographic information was obtained for each study subject, including age, gender, race, and smoking history. The clinical stage at presentation according to the AJCC staging system (AJCC 8th edition) was recorded considering that many patients did not receive pathological staging.

DNA Isolation, Library Preparation, and Sequencing—A pathologist examined H&E-stained slides from each case to confirm the diagnosis of squamous cell carcinoma. Automated DNA extraction was from FFPE tissue sections using the Promega Maxwell MDx16™ instruments (Promega) and then fragmented by sonication. Subsequent quality assessments were performed by ultraviolet absorbance and quantity assessments. During DNA isolation and library preparation, DNA concentration was measured by fluorometry and DNA quality was evaluated using the Agilent 2100 Bioanalyzer high sensitivity assay. DNA libraries were pooled for deep sequencing using an Illumina HiSeq2500™ sequencer. The UNCseq targeted sequencing platform involves sequencing exons of a custom list of 650 human genes (covering 3.4 M bases) and 10 pathogen genome segments in fixed or frozen cancer tissue and matched germline DNA from consenting local patients. This custom sequencing platform provided deep coverage of all HPV16 open reading frames. For general metrics on human gene sequencing quality, read depth, and coverage statistics, please see our recent report which reviewed these features of the data set [37]. Excluding the hypervariable region 3150–3351, the median coverage of the HPV16 genome was 7904 with an IQR of 6867–7953. The hypervariable region 3150–3351 had a median coverage of 3307 with an IQR of 837–7377.

Inclusion Criteria—The UNCseq database was queried for p16+ tumors originating from the anatomic oropharynx (tonsil or tongue base) with available tumor sequencing data as well as data on stage, treatment strategy, clinical outcome, and histopathology available. HPV16 positivity was confirmed by DNA sequencing reads which mapped to the HPV16 genome (see above comments). Patients were excluded from clinical analyses if tumors were not p16+. The UNCseq database was also queried for squamous cell carcinomas of the uterine cervix; HPV16 positivity was confirmed by DNA sequencing reads which mapped to the HPV16 genome (see Viral Copy Number Analysis section below). For both tumors from the oropharynx and uterine cervix, tumors with non-HPV16 genotypes were excluded from the study.

Assigning HPV16 Positivity and Viral Copy Number Analysis—Raw reads were aligned to the human genome plus a comprehensive library of HPV virus sequences, using compiled reference sequences from the ViFi analysis pipeline [38]. An HPV16 viral copy number was estimated based on the ratio of reads mapping to the HPV16 (NC_001526.4) genome and HG19. $\text{Log}_{10}(\text{Reads}_{\text{HPV16}}/\text{Reads}_{\text{HG19}}) > -4$ was used as an empiric cut off for HPV positivity. This threshold was selected based on an obvious bimodal distribution in the

data; see our prior publication justifying this criterion [21]. The Samtools Idxstats() function was used to quantify reads mapping to HPV16 and human chromosomes [39]. For HPV16-positive tumors, $\text{Log}_{10}(\text{Reads}_{\text{HPV16}}/\text{Reads}_{\text{HG19}})$ calculated as above was considered as a relative metric of the HPV16 copy number.

Integration Site Calling—Raw reads were aligned to the human genome plus a comprehensive library of HPV virus sequences, using compiled reference sequences from the ViFi analysis pipeline [38]. The ViFi pipeline was used to identify discordant (human–viral) read pairs, which cluster to potential integration sites in the human and HPV genomes [38]. Breakpoints with more than 50 clustered discordant read pairs were classified as positive for integration. As a second method of detecting integration, an independent assembly structural variant caller, svABA, was utilized with default settings [40]. Alignment BAM files output from ViFi, which included hg19 as well as the HPV16 genome, were used as inputs for svABA. Contiguous sequences assembled by svABA were matched to discordant read pair clusters from ViFi. Discordant read pair clusters that were with fewer than 50 read pairs were also considered suggestive of integration if a continuous sequence supporting a HPV–human structural variant was also identified by svABA and passed the quality filter independently, or contained at least 5 reads spanning a HPV–human junction also identified by a ViFi cluster. Integration sites were identified from individual HPV–human breakpoints, looking for clusters of breakpoints within 1 Mb of each other using in-house scripts.

Assigning Gene Effects—The R SMITE [41] and GenomicRanges [42] packages were used to identify the nearest gene and minimum distance to it, for each high-confidence breakpoint. Based on prior reports demonstrating altered gene expression of genes as much as 500 Kb away from the gene of interest [7], we used ± 500 kb as a threshold for genes whose functions may be affected by integration.

Structural Variant Calling and Analysis—We also considered structural variants which joined disparate aspects of the HPV16 genome. To distinguish these from HPV16–human structural variants (i.e., integration breakpoints), we will refer to these events as HPV–HPV structural variants. HPV–HPV structural variants were identified from the svABA analysis as described above. HPV–HPV structural variants were only considered if they passed svABA quality filters. Additional filtering to assess for biologically relevant structural variants was performed based on variant allele frequencies (VAFs) estimated by svABA. The thresholds for VAF filters are described in the results section.

Viral Deep Copy Loss Analysis—Considering that HPV+ cancer cells often contain multiple copies of the viral genome, the loss of genomic material in some copies of the virus might result in a decreased copy number (shallow loss) of some viral genes. However, loss of genomic material in all or the vast majority of copies, or deep copy loss, would be expected to result in the most prominent functional differences. In this manuscript, we focused on copy number analysis on regions of deep viral copy losses, which represent the complete or nearly complete loss of the genomic regions in question. To identify these events, nucleotide-specific estimates of read depth were acquired using the Samtools depth function [39]. Relative tumor-specific read depths were calculated as the ratio of read depth to the tumor's median read depth in the HPV16 regions corresponding to E6 and E7. Nucleotides with read depths $<0.2\%$ of median E6/E7 coverage were considered to have deep loss of the involved nucleotides. Nucleotides with $<1\%$ of median E6/E7 coverage that bordered on both side areas of $<0.2\%$ deep loss were also considered to be a component of the same deep loss event. Deep losses involving only nucleotides in the hypervariable regions 3150–3351 were excluded from analysis as this area had low coverage across the entire data set.

3. Results

3.1. Integration Analyses

Viral integration sites were identified in approximately 75% of HPV16-associated oropharyngeal (OPSCC, $n = 104$) and uterine cervical (UCSCC, $n = 44$) cancers (see Figure 1A). Individual integration sites, defined as being within a mega-base region of a chromosome, had similar numbers of HPV–human breakpoint junctions per site in OPSCC and UCSCC, with some sites having 10 or more breakpoint junctions (Figure 1B). The viral copy number was not related to the viral integration status of tumors in either OPSCC or UCSCC; however, the viral copy number was higher in OPSCC as compared to UCSCC regardless of integration (Figure 1C). When analysis was restricted to OPSCC and UCSCC tumors with integration, the viral copy number correlated with the number of integration sites (Spearman's $Rho = 0.49$, $p < 5 \times 10^{-8}$). Amongst tumors with HPV16 integration, OPSCC tumors had significantly more integration sites per tumor (see Figure 1D). Although not statistically significant, integration preferences by chromosome were dissimilar between UCSCC and OPSCC, with chromosome 7 being preferred in UCSCC and chromosome 19 preferred in OPSCC (see Figure 1E). The distribution of sites differed significantly from a random model in OPSCC (Figure 1F), and based on chromosome size, integrations in chromosome 19 were independently enhanced above expectations (chi-squared test, $p < 0.001$). Integrations in areas of the genome with the highest gene density (top 5%) were also enhanced in OPSCC as compared to UCSCC (Figure 1G), and the relative number of genes with integration sites within 500 kB was higher in OPSCC (Figure 1H). Similarly, OPSCC tumors typically had more chromosomes affected by integration sites (see Figure 1I).

Individual sites were mapped to the closest gene, based on the median breakpoint position (see Figure 1J). In agreement with prior reports, recurrent integration sites were noted in both OPSCC and UCSCC. The most common genes associated with integration sites are displayed in Figure 1J or Figure 2. The genomic distribution of individual integration sites is displayed in Figure 2. A gross preference for peri-centromeric integrations were noted (see Figure 2) [43].

3.2. Deep Copy Loss Analyses

Deep copy-losses within the HPV genome were identified based on read depth as compared to E6 and E7 in the same tumor (see Methods). The regions of deep loss identified are displayed in Figure 3A. Deep losses were much more common and larger in UCSCC as compared to OPSCC (Figure 3B,C). Only 2 out of 104 (2%) OPSCC tumors had large-scale losses of $> 10\%$ of the viral genome, whereas 15 of 44 (34%) of UCSCC had such losses (Figure 3A). Tumors with deep losses of the HPV genome were usually integrated (Figure 3D), with only one UCSCC tumor with deep loss lacking integration. Deep losses involving E1, E2, and E5 were more common in UCSCC as compared to OPSCC (Figure 3E). E2 was the gene most commonly lost in UCSCC and the HPV gene most differentially lost as compared to OPSCC (Figure 3E). More specifically, 27% of UCSCC tumors and 5% of OPSCC tumors harbored deep losses involving E2. As expected, the upstream regulatory region (URR) containing the major early promoters, as well as E6 and E7 genes, were universally spared from deep losses in both diseases (Figure 3E).

Significance based on chi-squared test, *p*-value indicated in plot. (H) Box plot—number of genes affected by integration sites per tumor. Genes within 500 kb of the integration site were included. Significance based on Wilcoxon Rank-sum test, *p*-values are indicated in plot. (I) Box plot—number of chromosomes affected by integration sites per tumor. Significance based on Wilcoxon Rank-sum test, *p*-values are indicated in plot. (J) Bar plot—frequency of genes effected by integration events. Ordered by decreasing frequency across both OPSCC and UCSCC.

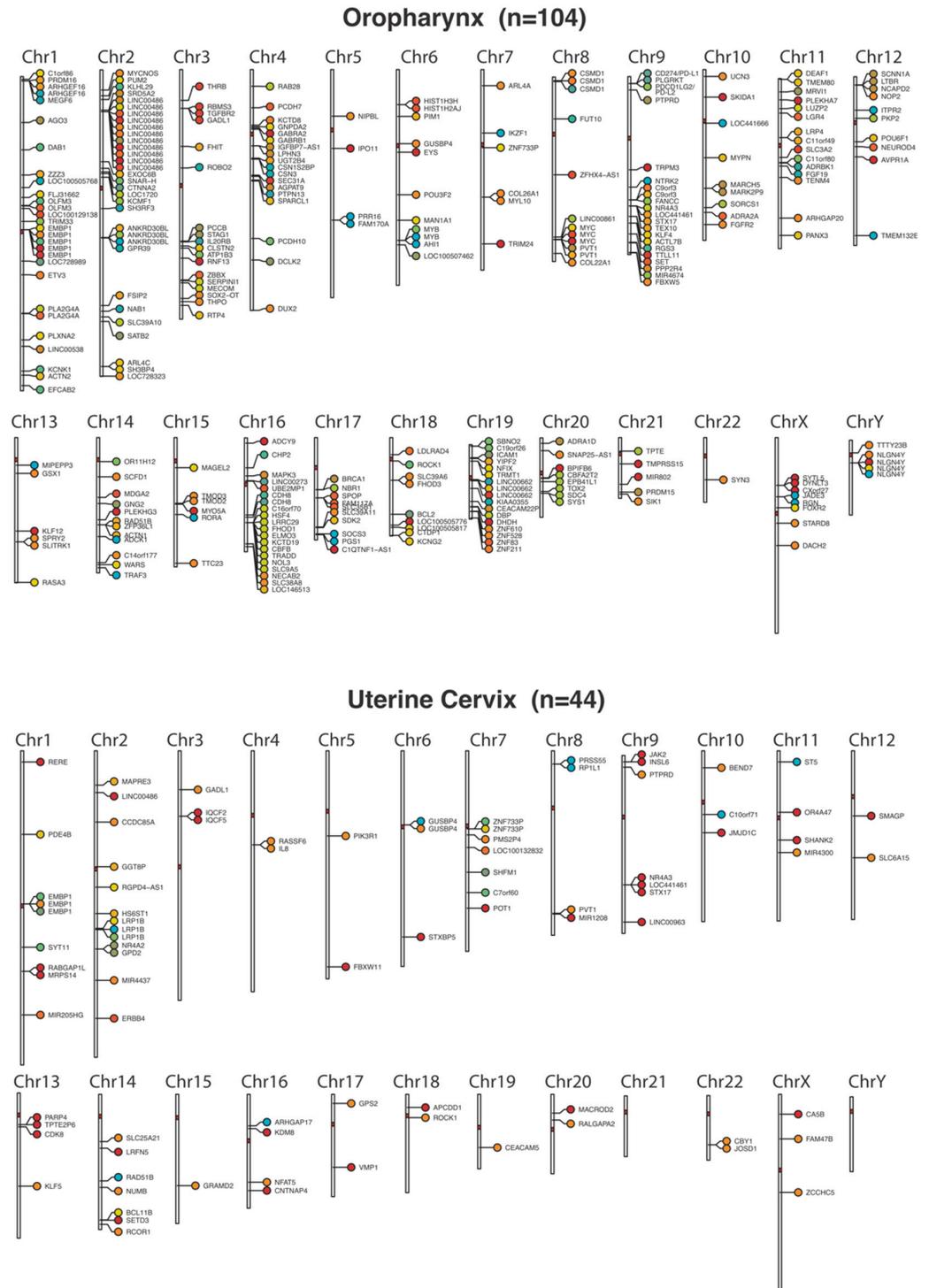


Figure 2. Locations of HPV16 integration sites across the human genome. Genes nearest to the integration sites are indicated. Colors identify tumor samples per chromosome, to allow identification of related sites in single tumors.

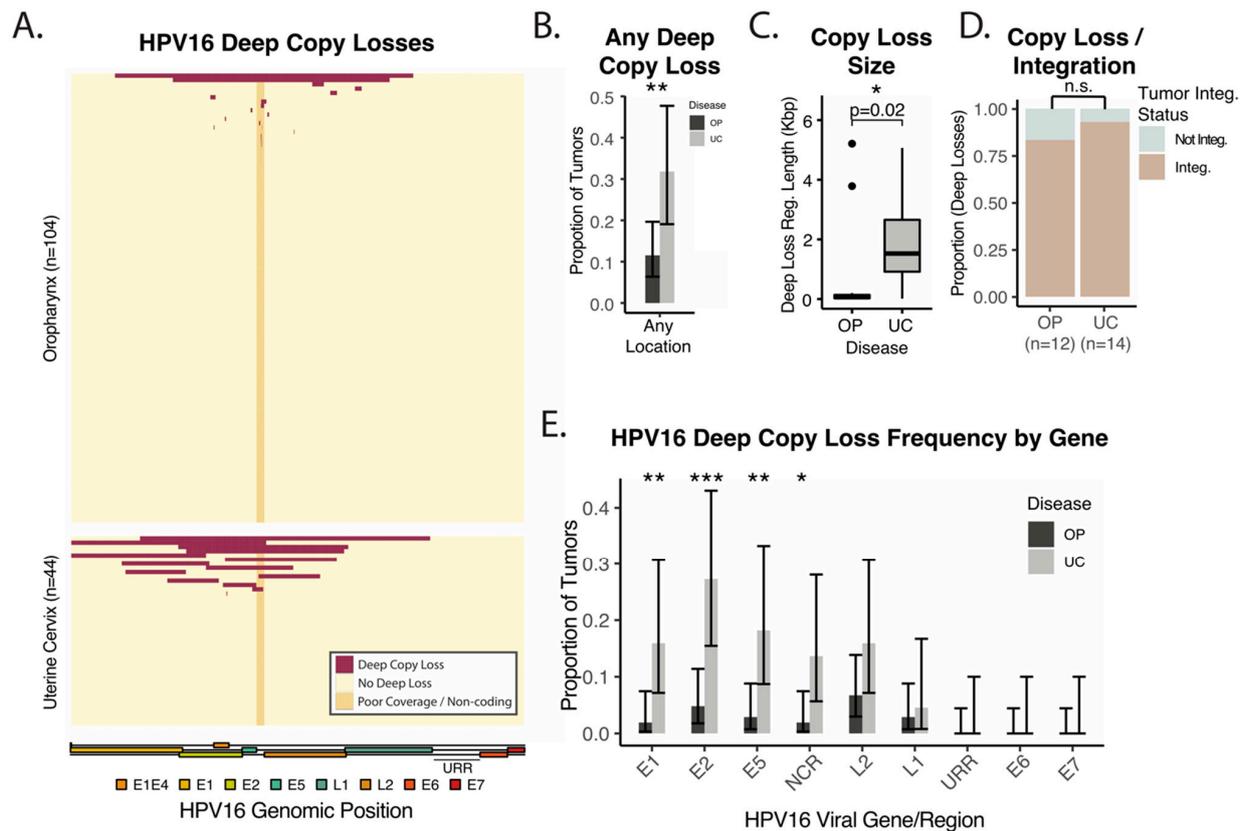


Figure 3. Deep losses of HPV16 genomic material in OPSCC and UCSCC. **(A)** Heatmap—regions of deep loss in the HPV16 genome by tumor. Columns—HPV16 genomic position. Rows—tumors. Dark Red—deep loss. Yellow—no deep loss. Orange—hypervariable non-coding region with poor coverage, excluded from deep loss analysis unless otherwise indicated. **(B)** Bar plot—proportion of tumors with any deep genomic loss in the HPV16 genome. Error bars represent 95% CI. **(C)** Boxplot—size distribution of regions of deep copy loss. p -value represents Wilcoxon rank-sum test. **(D)** Stacked bar plot—proportion of tumors with deep loss and viral integration. **(E)** Bar plot—proportion of tumors with deep genomic loss affecting the indicated regions of the HPV16 genome. Error bars represent 95% CI. NCR—hypervariable non-coding region with poor coverage (also orange in Panel A). URR—upstream regulatory region. * p -value < 0.05. ** p -value < 0.005. *** p -value < 0.0005.

3.3. HPV-HPV Structural Variant Analyses

We also investigated structural variants (SVs) within HPV16 itself (HPV–HPV SVs). These variants were identified by the svABA pipeline as passing the default quality filter, and were then further filtered by variant allele frequency to look for potentially biologically relevant variants (Figure 4A). There was a non-significant trend towards increasing numbers of SVs in OPSCC as compared to UCSCC (Figure 4B). Interestingly, HPV–HPV SVs of moderate VAF in UCSCC were mostly identified in tumors without integration, and this relationship was distinct from OPSCC (Figure 4C). The reliability of the SV calls and their potential biological relevance was bolstered by relative copy number alterations in HPV that correlated with the identified HPV–HPV SVs. More specifically, all high VAF HPV–HPV SVs had obvious discontinuities in coverage at the identified breakpoints, which also correlated with the orientation of the break-end pair. Two examples of this are shown in Figure 4D–E, where an HPV–HPV SV is the origin of a relative amplification involving L1 in an OPSCC tumor (Case A) and E2 is partially lost in a UCSCC tumor (Case B). Locations of HPV–HPV SV breakpoints appeared to be non-random and grossly favored regions of L2 and the URR, Figure 4F. Examining HPV genomic regions based on their size also demonstrated that breakpoint locations were not random for OPSCC or UCSCC, with UCSCC favoring breakpoints that

involved E2 (Figure 4G). Interesting HPV-HPV SVs that caused a break in E2 were more common in UCSCC as compared to OPSCC (Figure 4H). Considering that UCSCC had both an increased percentage of tumors with deep loss of E2 (Figure 3) and a greater number of HPV-HPV SVs causing loss of all or part of E2, UCSCC has a much larger percentage of tumors with E2 alteration as compared to OPSCC (Figure 4H).

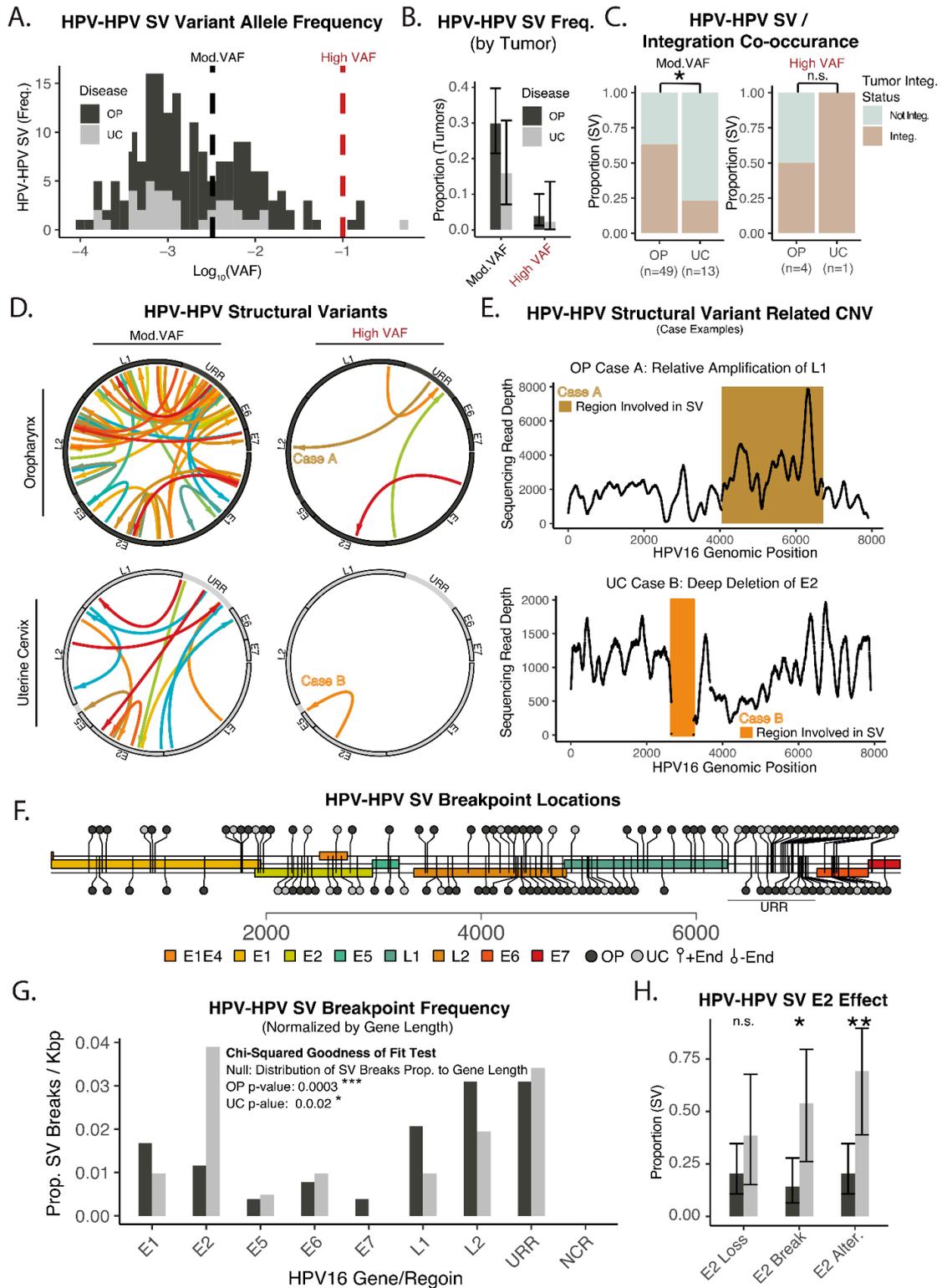


Figure 4. Structural variants in the HPV16 genome in OPSCC and UCSCC. (A) Histogram—estimated

variant allele frequency (VAF) of HPV–HPV structural variants (SVs). High-quality variant calls were investigated based on empiric VAF thresholds indicated. Black dashed line—moderate VAF threshold; variants with VAF greater than this cutoff were analyzed further. Red dashed line—high VAF threshold, variants with VAF greater than this cutoff were independently analyzed and displayed where indicated. **(B)** Bar plot—proportion of tumors with HPV–HPV SVs. Error bars represent 95% CI. **(C)** Stacked bar plot—proportion of tumors with HPV–HPV SVs and viral integration. Significance testing represents chi-squared test. * p -value < 0.05. **(D)** Circos Plots—location and orientation of HPV–HPV SVs. Arrow origin—upstream aspect of structural variant breakpoint junction. Arrow head—downstream aspect of structural variant breakpoint junction. VAF groups are as defined in Panel **(A)**. Colors in each individual plot represent a single tumor sample. **(E)** Genomic coverage plots—HPV copy number alterations related to HPV–HPV structural variants in tumors with high VAF SVs. The color and case number link coverage plots to the (left) adjacent circos plots in panel **(D)**. **(F)** Lollipop plot—breakpoint locations of HPV–HPV SVs with moderate VAF. Light grey—HPV–HPV SV from UCSCC tumor. Dark grey—HPV–HPV SV from OPSCC tumor. **(G)** Bar plot—proportion of HPV–HPV SV breakpoint sites by HPV16 gene or region. Significance based on chi-squared goodness of fit test, with sites distribution based on gene/region size as null hypothesis. * p -value < 0.05. *** p -value < 0.0005. **(H)** Bar plot—proportion of HPV–HPV SVs and effects on E2. Error bars represent 95% CI. Significance based on chi-squared test. * p -value < 0.05. ** p -value < 0.005.

Taken together, these results indicate that both integration and HPV–HPV SVs are selected for E2 loss in UCSCC, but not OPSCC. Conversely, the higher frequency of integration sites per tumor and preferential integration in a gene-dense area of the genome, suggests that integration sites in OPSCC are selected for the oncogenic properties resulting from effects on the host genome.

4. Discussion

Our direct comparison of HPV integration in oropharyngeal and cervical tumors revealed several important dissimilarities supporting the idea that HPV-driven carcinogenesis is different between these cancers. Below, we summarize major points highlighting their discrete biology.

The Different Role of E2 in HPV Carcinogenesis—Integration of the HPV genome into the human genome is an established hallmark of HPV-mediated carcinogenesis. Given that the HPV E2 protein negatively regulates expression of viral oncoproteins E6 and E7, another hallmark of this model is E2 loss. This HPV carcinogenesis model was developed from analysis of UCSCC, and our analyses of UCSCC are supportive since integration was detected in the vast majority and E2 was the HPV gene most commonly altered by deletion (30%) and SVs (40%). On the other hand, integration was less frequent in OPSCC, and even in integrated tumors, E2 loss by deep deletion or HPV–HPV SVs was significantly less frequent in OPSCC, where L2 was found to be the HPV gene most frequently altered by deletion and SVs (Figure 3E or Figure 4G). Although more limited in number of tumors analyzed and without direct comparison to UCSCC, Huang et al. noted a similar lack of E2 loss related to viral integration in squamous cell carcinomas of the penis [44]. The indifference toward the loss of E2 and a higher percentage of non-integrated tumors in OPSCC contradict tenets of most accepted models of HPV carcinogenesis, but is not without precedent in the literature. Our group and others have reported that some HPV+ OPSCC tumors lack integrations, maintain all HPV genes, and express E6 and E7, albeit at lower levels [27]. As a refinement of these findings, here, we report that integration can be detected in 75% of HPV+ OPSCC tumors. This estimate is higher than many prior reports, perhaps reflecting the sensitivity of our assay to detect low frequency variants that could be sub-clonally present in cancer cells. However, even with a very sensitive assay, 25% of tumors lacked HPV integration. Groves et al. demonstrated constitutive expression of HPV oncogenes E6 and E7 without E2 loss [45]. Other groups have suggested that HPV integration, or lack thereof, was not correlated with differential expression of E6 and E7 [46,47]. Furthermore, maintained expression of E2, E4, and E5 have been purported to

be critical for an alternate carcinogenic pathway in HPV+ OPSCC, which was associated with a lack of genomic integration [48]. The present work further highlights a distinct relationship of E2 to HPV+ OPSCC as compared to UCSCC, to the extent that a significant role for E2 disruption in HPV+ OPSCC should be questioned.

A Higher Number of HPV16 Integration Sites Per Tumor in OPSCC—Considering that E2 loss may be less important in providing selective advantages for cancer and pre-malignant cells in the oropharynx, other roles for viral integration should be considered, especially since we found HPV integration in 75% of OPSCC cases. Furthermore, we found a strikingly higher number of integration sites per tumor in HPV+ OPSCC as compared to UCSCC, suggesting that maintenance of these integrated HPV copies contributes to cancer development or maintenance. Accompanying higher integration sites per tumor, we also found that viral copy number was higher in OPSCC. Although integration itself was not associated with a higher or lower copy number in either disease, the number of integration sites in an individual tumor was correlated to viral copy number. We note that other groups have associated higher viral copy number in tumors harboring episomal HPV; however, our data do not support this association [49].

The ability of HPV integration to alter expression of nearby human genes (within 500 Kb) has been demonstrated in both OPSCC and UCSCC [7,12] with recurrent integration sites identified near genes implicated in cancer such as *MYC* [7,50], *MYB* [7], *PVT1* [51–53], *RAD51B* [50,54], *EMBP1* [55], *CD274/PD-L1* [7,56], and *SOX2* [7]. Our results agree with these prior reports confirming integration in or near these genes. Our study identified additional non-recurrent, cancer-associated genes in or near integration sites in HPV+ OPSCC including *TGFBR2*, *FGFR2* [20], *PD-L2* [57], *TRAF3* [26], *BRCA1*, *SPOP* [58], and *BCL-2* [59]. Our results bolster prior reports that modulation of the function of genes near HPV integration sites is key in the oncogenesis of OPSCC [7]. As other studies investigating HPV integration in UCSCC have found, we noted HPV16 integration near *NR4A2*, *MYC* [60], *LRP1B* [61–63], and *MACROD2* [54]. Unfortunately, additional tissue for expression analysis was not available for the vast majority of the analyzed cases; however, there is a strong precedent in the literature that HPV integration can modulate the expression of nearby cancer genes in both OPSCC [7,50,64] and UCSCC [50,60,65]. Our finding that areas of high gene density are more likely to harbor integration sites in OPSCC as compared to UCSCC, as well as the increased number of integration sites per tumor, are both consistent with the hypothesis that the modulation of human cancer genes may carry increased significance in OPSCC as compared to UCSCC.

HPV+ OPSCC is strikingly more prevalent in men as compared to women [66]. In HPV-negative HNSCC, men typically have a worse prognosis, yet loss of the Y chromosome has been associated with poorer outcomes [67]. Our data identify an interesting integration hotspot in chromosome Y near/at *NLGN4Y* (Figure 1F). Previous work has associated this gene with complex neurocognitive traits such as male homosexuality and disorders including autism [68,69]. A homologous gene on chromosome X, *NLGN4X*, has been shown to have non-identical signaling properties despite 97% sequence homology to *NLGN4Y*, which have been used to explain X-linked phenomenon related to this XY-gene-pair [68]. However, we found only scant reports linking *NLGN4Y* to cancer [70,71]. The most frequent integration hot-spot we identified in OPSCC was near *LINC00486*, on chromosome 2; however, it was notably uncommon for these sites to be called by assembly-based reconstruction, so they should be interpreted with some caution. However, HPV16 integration sites in this area have been confirmed and extensively characterized in UCSCC cell lines [72]. In this study, the integration sites were near an enhancer element that is thought to be related to epithelial differentiation [72], and HPV integration has been demonstrated to be more likely to be found in or near enhancer elements [50].

Mechanistic Considerations Regarding HPV16 Integration—The strikingly higher number of integration sites in OPSCC as compared to UCSCC per integrated tumor (Figure 1D, median shift of ~5x), merits consideration from a mechanistic perspective. Integration, like all structural variants, necessitate DNA double-strand breaks as a com-

ponent of the process [73,74]. HPV+ OPSCC has been noted to be particularly deficient in double-strand break repair [75–78] and this has been related to the relative radiosensitivity of these tumors [79]. Deficiency of homologous recombination [75] and replication stress [80] have been linked to HPV integration and pathogenicity. Work from our group and others have demonstrated that APOBEC plays a key role in HPV-mediated carcinogenesis in the oropharynx [81], and increased expression and activity of APOBEC3B and APOBEC3A, respectively, have been linked to direct actions of HPV oncoproteins E6 and E7 [82]. APOBEC3s have been associated with genomic instability across multiple cancer types [83]. Maybe not surprisingly, APOBEC3A (A3A) protein expression has been strongly correlated with increased double-strand breaks and HPV viral integration [84]. Others have shown that integration also directly correlates with genomic instability [85]. Whether differences in DNA damage repair and global genomic instability between UCSCC and OPSCC may explain the differential rates of integration is unknown but merits additional focused analysis.

Similarly, one could consider if the molecular process of integration is variable between the two diseases, and indeed, at least three structural classes of integration sites have been identified: single viral genomic insertions (Type I), insertions of tandem viral repeats (Type II), and tandem repeats of hybrid viral–human DNA (Type III) [72]. Unfortunately, without long reads or optical techniques [72,85], it is impossible to distinguish these types, as in our data. However, integrations were commonly found to be quite complex in both OPSCC and UCSCC, with four or more human–viral genomic junctions; these do not fit well into any of the described models in the literature. We note that experimental studies on LINC00486-related integration sites were Type III, and the complex/repetitive nature of the site might explain the failure of assembly-based methods in these cases [72].

Our study was limited in the ability to resolve some complexities of integration events, primarily due to the targeted affinity capture technique, which does not allow unbiased analysis of the copy number of small regions of non-target human genomic DNA, as is possible with whole genome sequencing. More specifically, because of this, we were not able to determine if the integration sites detected were related to chromosomal or extrachromosomal DNA, although other groups using whole genome sequencing have reported detecting human–HPV hybrid extrachromosomal DNA in OPSCC tumors [36].

Clinical Implications of HPV16 Integration in OPSCC—We did not find any prognostic relationship to HPV16 viral integration in HPV+ OPSCC in this study. However, this is a frequently discussed association in the literature, with most studies suggesting that integration is a poor prognostic marker [21,86], but conflicting reports exist [27,87]. We notice that other DNA-sequencing-based studies have failed to demonstrate prognostic relationship to integration [7,36]. Interestingly, the strongest evidence for HPV integration being a poor prognostic factor has been based on RNA sequencing [21,86], while conflicting studies as well as studies with no association between HPV integrations and survival were based on DNA technology [7,27,36,87]; therefore, we hypothesize that only transcriptionally active integration sites are prognostic.

5. Conclusions

We directly compared 44 cases of HPV16+ UCSCC and 104 cases of HPV16+ OPSCC. E2 loss was common in HPV+ UCSCC, mediated by frequent copy number losses or structural variant breakpoints. Conversely, E2 loss was not selected for in HPV+ OPSCC. Tumors with no evidence of integration represented ~25% of both UCSCC and OPSCC cases. Amongst tumors with genomic integration, OPSCC tumors had many more integration sites per tumor, and were more likely to have integration sites in high gene density areas of the genome. These data highlight clear differences between HPV16-associated UCSCC and OPSCC related to the physical state of the HPV16 genome, and there may be implications for distinct HPV-associated carcinogenesis in the two diseases.

Author Contributions: Conceptualization, T.P.S., W.G.Y. and N.I.; methodology, T.P.S.; software, T.P.S.; validation, T.P.S.; formal analysis, T.P.S.; investigation, T.P.S., W.G.Y. and N.I.; resources, T.P.S., W.G.Y. and N.I.; data curation, T.P.S.; writing—all authors contributed to writing the original draft and its review and editing; data visualization, T.P.S.; supervision, T.P.S., W.G.Y. and N.I.; project administration, T.P.S., W.G.Y. and N.I.; funding acquisition, T.P.S., W.G.Y. and N.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NIH NIDCR R01DE027942, Awarded to N.I. and W.G.Y. and NIH NIDCR KO8-DE029241, Awarded to T.S.

Institutional Review Board Statement: IRB Number: 17-2947 (Date Finalized After Modifications: 6 June 2022). This IRB-approved trial opened in 2011. All studies were done with the approval of our Institutional Review Board, patient participation required written informed consent, and all studies were conducted in accordance with recognized ethical guidelines as described in U.S. Common Rule.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Raw sequencing data have been deposited to dbGaP; accession no. phs001713.v1.p1.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ibeanu, O.A. Molecular Pathogenesis of Cervical Cancer. *Cancer Biol. Ther.* **2011**, *11*, 295–306. [[CrossRef](#)] [[PubMed](#)]
2. Burd, E.M. Human Papillomavirus and Cervical Cancer. *Clin. Microbiol. Rev.* **2003**, *16*, 1–17. [[CrossRef](#)] [[PubMed](#)]
3. Longworth, M.S.; Laimins, L.A. Pathogenesis of Human Papillomaviruses in Differentiating Epithelia. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 362–372. [[CrossRef](#)] [[PubMed](#)]
4. McMurray, H.R.; Nguyen, D.; Westbrook, T.F.; McAnce, D.J. Biology of Human Papillomaviruses. *Int. J. Exp. Pathol.* **2001**, *82*, 15–33. [[CrossRef](#)] [[PubMed](#)]
5. Schiffman, M.H.; Bauer, H.M.; Hoover, R.N.; Glass, A.G.; Cadell, D.M.; Rush, B.B.; Scott, D.R.; Sherman, M.E.; Kurman, R.J.; Wacholder, S. Epidemiologic Evidence Showing That Human Papillomavirus Infection Causes Most Cervical Intraepithelial Neoplasia. *J. Natl. Cancer Inst.* **1993**, *85*, 958–964. [[CrossRef](#)] [[PubMed](#)]
6. Walboomers, J.M.; Jacobs, M.V.; Manos, M.M.; Bosch, F.X.; Kummer, J.A.; Shah, K.V.; Snijders, P.J.; Peto, J.; Meijer, C.J.; Muñoz, N. Human Papillomavirus Is a Necessary Cause of Invasive Cervical Cancer Worldwide. *J. Pathol.* **1999**, *189*, 12–19. [[CrossRef](#)]
7. Symer, D.E.; Akagi, K.; Geiger, H.M.; Song, Y.; Li, G.; Emde, A.-K.; Xiao, W.; Jiang, B.; Corvelo, A.; Toussaint, N.C.; et al. Diverse Tumorigenic Consequences of Human Papillomavirus Integration in Primary Oropharyngeal Cancers. *Genome Res.* **2022**, *32*, 55–70. [[CrossRef](#)]
8. D’Souza, G.; Kreimer, A.R.; Viscidi, R.; Pawlita, M.; Fakhry, C.; Koch, W.M.; Westra, W.H.; Gillison, M.L. Case-Control Study of Human Papillomavirus and Oropharyngeal Cancer. *N. Engl. J. Med.* **2007**, *356*, 1944–1956. [[CrossRef](#)]
9. Gillison, M.L.; Shah, K.V. Human Papillomavirus-Associated Head and Neck Squamous Cell Carcinoma: Mounting Evidence for an Etiologic Role for Human Papillomavirus in a Subset of Head and Neck Cancers. *Curr. Opin. Oncol.* **2001**, *13*, 183–188. [[CrossRef](#)]
10. Sisk, E.A.; Bradford, C.R.; Jacob, A.; Yian, C.H.; Staton, K.M.; Tang, G.; Harris, M.O.; Carey, T.E.; Lancaster, W.D.; Gregoire, L. Human Papillomavirus Infection in “Young” versus “Old” Patients with Squamous Cell Carcinoma of the Head and Neck. *Head Neck* **2000**, *22*, 649–657. [[CrossRef](#)]
11. Berman, T.A.; Schiller, J.T. Human Papillomavirus in Cervical Cancer and Oropharyngeal Cancer: One Cause, Two Diseases. *Cancer* **2017**, *123*, 2219–2229. [[CrossRef](#)] [[PubMed](#)]
12. Cullen, A.P.; Reid, R.; Campion, M.; Lörintz, A.T. Analysis of the Physical State of Different Human Papillomavirus DNAs in Intraepithelial and Invasive Cervical Neoplasm. *J. Virol.* **1991**, *65*, 606–612. [[CrossRef](#)] [[PubMed](#)]
13. Hudelist, G.; Manavi, M.; Pischinger, K.I.D.; Watkins-Riedel, T.; Singer, C.F.; Kubista, E.; Czerwenka, K.F. Physical State and Expression of HPV DNA in Benign and Dysplastic Cervical Tissue: Different Levels of Viral Integration Are Correlated with Lesion Grade. *Gynecol. Oncol.* **2004**, *92*, 873–880. [[CrossRef](#)]
14. Scheffner, M.; Werness, B.A.; Huibregtse, J.M.; Levine, A.J.; Howley, P.M. The E6 Oncoprotein Encoded by Human Papillomavirus Types 16 and 18 Promotes the Degradation of P53. *Cell* **1990**, *63*, 1129–1136. [[CrossRef](#)]
15. Dyson, N.; Howley, P.M.; Münger, K.; Harlow, E. The Human Papilloma Virus-16 E7 Oncoprotein Is Able to Bind to the Retinoblastoma Gene Product. *Science* **1989**, *243*, 934–937. [[CrossRef](#)]
16. Münger, K.; Baldwin, A.; Edwards, K.M.; Hayakawa, H.; Nguyen, C.L.; Owens, M.; Grace, M.; Huh, K. Mechanisms of Human Papillomavirus-Induced Oncogenesis. *J. Virol.* **2004**, *78*, 11451–11460. [[CrossRef](#)]

17. Begum, S.; Cao, D.; Gillison, M.; Zahurak, M.; Westra, W.H. Tissue Distribution of Human Papillomavirus 16 DNA Integration in Patients with Tonsillar Carcinoma. *Clin. Cancer Res.* **2005**, *11*, 5694–5699. [[CrossRef](#)]
18. Mirghani, H.; Mortuaire, G.; Armas, G.L.; Hartl, D.; Aupérin, A.; El Bedoui, S.; Chevalier, D.; Lefebvre, J.L. Sinonasal Cancer: Analysis of Oncological Failures in 156 Consecutive Cases. *Head Neck* **2014**, *36*, 667–674. [[CrossRef](#)]
19. Syrjänen, S. HPV Infections and Tonsillar Carcinoma. *J. Clin. Pathol.* **2004**, *57*, 449–455. [[CrossRef](#)]
20. Pan, C.; Issaeva, N.; Yarbrough, W.G. HPV-Driven Oropharyngeal Cancer: Current Knowledge of Molecular Biology and Mechanisms of Carcinogenesis. *Cancers Head Neck* **2018**, *3*, 12. [[CrossRef](#)]
21. Schrank, T.P.; Landess, L.; Stepp, W.H.; Rehmani, H.; Weir, W.H.; Lenze, N.; Lal, A.; Wu, D.; Kothari, A.; Hackman, T.G.; et al. Comprehensive Viral Genotyping Reveals Prognostic Viral Phylogenetic Groups in HPV16-Associated Squamous Cell Carcinoma of the Oropharynx. *Mol. Cancer Res.* **2022**. [[CrossRef](#)] [[PubMed](#)]
22. Pyeon, D.; Newton, M.A.; Lambert, P.F.; den Boon, J.A.; Sengupta, S.; Marsit, C.J.; Woodworth, C.D.; Connor, J.P.; Haugen, T.H.; Smith, E.M.; et al. Fundamental Differences in Cell Cycle Deregulation in Human Papillomavirus-Positive and Human Papillomavirus-Negative Head/Neck and Cervical Cancers. *Cancer Res.* **2007**, *67*, 4605–4619. [[CrossRef](#)] [[PubMed](#)]
23. Schlecht, N.F.; Burk, R.D.; Adrien, L.; Dunne, A.; Kawachi, N.; Sarta, C.; Chen, Q.; Brandwein-Gensler, M.; Prystowsky, M.B.; Childs, G.; et al. Gene Expression Profiles in HPV-Infected Head and Neck Cancer. *J. Pathol.* **2007**, *213*, 283–293. [[CrossRef](#)] [[PubMed](#)]
24. Cancer Genome Atlas Network. Comprehensive Genomic Characterization of Head and Neck Squamous Cell Carcinomas. *Nature* **2015**, *517*, 576–582. [[CrossRef](#)] [[PubMed](#)]
25. Cancer Genome Atlas Research Network; Albert Einstein College of Medicine; Analytical Biological Services; Barretos Cancer Hospital; Baylor College of Medicine; Beckman Research Institute of City of Hope; Buck Institute for Research on Aging; Canada’s Michael Smith Genome Sciences Centre; Harvard Medical School; Helen, F. Graham Cancer Center & Research Institute at Christiana Care Health Services; et al. Integrated Genomic and Molecular Characterization of Cervical Cancer. *Nature* **2017**, *543*, 378–384. [[CrossRef](#)]
26. Hajek, M.; Sewell, A.; Kaech, S.; Burtneess, B.; Yarbrough, W.G.; Issaeva, N. TRAF3/CYLD Mutations Identify a Distinct Subset of Human Papillomavirus-Associated Head and Neck Squamous Cell Carcinoma. *Cancer* **2017**, *123*, 1778–1790. [[CrossRef](#)]
27. Parfenov, M.; Pedamallu, C.S.; Gehlenborg, N.; Freeman, S.S.; Danilova, L.; Bristow, C.A.; Lee, S.; Hadjipanayis, A.G.; Ivanova, E.V.; Wilkerson, M.D.; et al. Characterization of HPV and Host Genome Interactions in Primary Head and Neck Cancers. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15544–15549. [[CrossRef](#)]
28. Schrank, T.P.; Prince, A.C.; Sathe, T.; Wang, X.; Liu, X.; Alzhanov, D.T.; Burtneess, B.; Baldwin, A.S.; Yarbrough, W.G.; Issaeva, N. NF- κ B over-Activation Portends Improved Outcomes in HPV-Associated Head and Neck Cancer. *Oncotarget* **2022**, *13*, 707–722. [[CrossRef](#)]
29. Tilborghs, S.; Corthouts, J.; Verhoeven, Y.; Arias, D.; Rolfo, C.; Trinh, X.B.; van Dam, P.A. The Role of Nuclear Factor-Kappa B Signaling in Human Cervical Cancer. *Crit. Rev. Oncol. Hematol.* **2017**, *120*, 141–150. [[CrossRef](#)]
30. DA Costa, R.M.G.; Bastos, M.M.S.M.; Medeiros, R.; Oliveira, P.A. The NF κ B Signaling Pathway in Papillomavirus-Induced Lesions: Friend or Foe? *Anticancer Res.* **2016**, *36*, 2073–2083.
31. Santegoets, S.J.; van Ham, V.J.; Ehsan, I.; Charoentong, P.; Duurland, C.L.; van Unen, V.; Höllt, T.; van der Velden, L.-A.; van Egmond, S.L.; Kortekaas, K.E.; et al. The Anatomical Location Shapes the Immune Infiltrate in Tumors of Same Etiology and Affects Survival. *Clin. Cancer Res.* **2019**, *25*, 240–252. [[CrossRef](#)] [[PubMed](#)]
32. Mellin, H.; Dahlgren, L.; Munck-Wikland, E.; Lindholm, J.; Rabbani, H.; Kalantari, M.; Dalianis, T. Human Papillomavirus Type 16 Is Episomal and a High Viral Load May Be Correlated to Better Prognosis in Tonsillar Cancer. *Int. J. Cancer* **2002**, *102*, 152–158. [[CrossRef](#)]
33. Koskinen, W.J.; Chen, R.W.; Leivo, I.; Mäkitie, A.; Bäck, L.; Kontio, R.; Suuronen, R.; Lindqvist, C.; Auvinen, E.; Molijn, A.; et al. Prevalence and Physical Status of Human Papillomavirus in Squamous Cell Carcinomas of the Head and Neck. *Int. J. Cancer* **2003**, *107*, 401–406. [[CrossRef](#)] [[PubMed](#)]
34. Kim, S.-H.; Koo, B.-S.; Kang, S.; Park, K.; Kim, H.; Lee, K.R.; Lee, M.J.; Kim, J.M.; Choi, E.C.; Cho, N.H. HPV Integration Begins in the Tonsillar Crypt and Leads to the Alteration of P16, EGFR and c-Myc during Tumor Formation. *Int. J. Cancer* **2007**, *120*, 1418–1425. [[CrossRef](#)]
35. Lace, M.J.; Anson, J.R.; Klussmann, J.P.; Wang, D.H.; Smith, E.M.; Haugen, T.H.; Turek, L.P. Human Papillomavirus Type 16 (HPV-16) Genomes Integrated in Head and Neck Cancers and in HPV-16-Immortalized Human Keratinocyte Clones Express Chimeric Virus-Cell MRNAs Similar to Those Found in Cervical Cancers. *J. Virol* **2011**, *85*, 1645–1654. [[CrossRef](#)]
36. Pang, J.; Nguyen, N.; Luebeck, J.; Ball, L.; Finegersh, A.; Ren, S.; Nakagawa, T.; Flagg, M.; Sadat, S.; Mischel, P.S.; et al. Extrachromosomal DNA in HPV-Mediated Oropharyngeal Cancer Drives Diverse Oncogene Transcription. *Clin. Cancer Res.* **2021**, *27*, 6772–6786. [[CrossRef](#)] [[PubMed](#)]
37. Schrank, T.P.; Lenze, N.; Landess, L.P.; Hoyle, A.; Parker, J.; Lal, A.; Sheth, S.; Chera, B.S.; Patel, S.N.; Hackman, T.G.; et al. Genomic Heterogeneity and Copy Number Variant Burden Are Associated with Poor Recurrence-Free Survival and 11q Loss in Human Papillomavirus-Positive Squamous Cell Carcinoma of the Oropharynx. *Cancer* **2021**, *127*, 2788–2800. [[CrossRef](#)] [[PubMed](#)]
38. Nguyen, N.-P.D.; Deshpande, V.; Luebeck, J.; Mischel, P.S.; Bafna, V. ViFi: Accurate Detection of Viral Integration and MRNA Fusion Reveals Indiscriminate and Unregulated Transcription in Proximal Genomic Regions in Cervical Cancer. *Nucleic Acids Res.* **2018**, *46*, 3309–3325. [[CrossRef](#)]

39. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. 1000 Genome Project Data Processing Subgroup The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
40. Wala, J.A.; Bandopadhyay, P.; Greenwald, N.F.; O'Rourke, R.; Sharpe, T.; Stewart, C.; Schumacher, S.; Li, Y.; Weischenfeldt, J.; Yao, X.; et al. SvABA: Genome-Wide Detection of Structural Variants and Indels by Local Assembly. *Genome Res.* **2018**, *28*, 581–591. [[CrossRef](#)]
41. Wijetunga, N.A.; Johnston, A.D.; Maekawa, R.; Delahaye, F.; Ulahannan, N.; Kim, K.; Grealley, J.M. SMITE: An R/Bioconductor Package That Identifies Network Modules by Integrating Genomic and Epigenomic Information. *BMC Bioinform.* **2017**, *18*, 41. [[CrossRef](#)] [[PubMed](#)]
42. Lawrence, M.; Huber, W.; Pagès, H.; Aboyoun, P.; Carlson, M.; Gentleman, R.; Morgan, M.T.; Carey, V.J. Software for Computing and Annotating Genomic Ranges. *PLoS Comput. Biol.* **2013**, *9*, e1003118. [[CrossRef](#)] [[PubMed](#)]
43. Garza-Rodríguez, M.L.; Oyervides-Muñoz, M.A.; Pérez-Maya, A.A.; Sánchez-Domínguez, C.N.; Berlanga-Garza, A.; Antonio-Macedo, M.; Valdés-Chapa, L.D.; Vidal-Torres, D.; Vidal-Gutiérrez, O.; Pérez-Ibave, D.C.; et al. Analysis of HPV Integrations in Mexican Pre-Tumoral Cervical Lesions Reveal Centromere-Enriched Breakpoints and Abundant Unspecific HPV Regions. *Int. J. Mol. Sci.* **2021**, *22*, 3242. [[CrossRef](#)]
44. Huang, K.-B.; Guo, S.-J.; Li, Y.-H.; Zhang, X.-K.; Chen, D.; Spiess, P.E.; Li, Z.-S.; Deng, C.-Z.; Chen, J.-P.; Zhou, Q.-H.; et al. Genome-Wide Profiling Reveals HPV Integration Pattern and Activated Carcinogenic Pathways in Penile Squamous Cell Carcinoma. *Cancers* **2021**, *13*, 6104. [[CrossRef](#)] [[PubMed](#)]
45. Groves, I.J.; Coleman, N. Pathogenesis of Human Papillomavirus-Associated Mucosal Disease. *J. Pathol.* **2015**, *235*, 527–538. [[CrossRef](#)] [[PubMed](#)]
46. Olthof, N.C.; Speel, E.-J.M.; Kolligs, J.; Haesevoets, A.; Henfling, M.; Ramaekers, F.C.S.; Preuss, S.F.; Drebber, U.; Wieland, U.; Silling, S.; 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. [[CrossRef](#)]
47. Olthof, N.C.; Huebbers, C.U.; Kolligs, J.; Henfling, M.; Ramaekers, F.C.S.; Cornet, I.; van Lent-Albrechts, J.A.; Stegmann, A.P.A.; Silling, S.; Wieland, U.; et al. Viral Load, Gene Expression and Mapping of Viral Integration Sites in HPV16-Associated HNSCC Cell Lines. *Int. J. Cancer* **2015**, *136*, E207–E218. [[CrossRef](#)]
48. Ren, S.; Gaykalova, D.A.; Guo, T.; Favorov, A.V.; Fertig, E.J.; Tamayo, P.; Callejas-Valera, J.L.; Allevato, M.; Gilardi, M.; Santos, J.; et al. HPV E2, E4, E5 Drive Alternative Carcinogenic Pathways in HPV Positive Cancers. *Oncogene* **2020**, *39*, 6327–6339. [[CrossRef](#)]
49. Chera, B.S.; Kumar, S.; Beaty, B.T.; Marron, D.; Jefferys, S.; Green, R.; Goldman, E.C.; Amdur, R.; Sheets, N.; Dagan, R.; et al. Rapid Clearance Profile of Plasma Circulating Tumor HPV Type 16 DNA during Chemoradiotherapy Correlates with Disease Control in HPV-Associated Oropharyngeal Cancer. *Clin. Cancer Res.* **2019**, *25*, 4682–4690. [[CrossRef](#)]
50. Warburton, A.; Markowitz, T.E.; Katz, J.P.; Pipas, J.M.; McBride, A.A. Recurrent Integration of Human Papillomavirus Genomes at Transcriptional Regulatory Hubs. *NPJ Genom. Med.* **2021**, *6*, 1–15. [[CrossRef](#)]
51. Liu, Y.; Lu, Z.; Xu, R.; Ke, Y. Comprehensive Mapping of the Human Papillomavirus (HPV) DNA Integration Sites in Cervical Carcinomas by HPV Capture Technology. *Oncotarget* **2015**, *7*, 5852–5864. [[CrossRef](#)] [[PubMed](#)]
52. Iden, M.; Fye, S.; Li, K.; Chowdhury, T.; Ramchandran, R.; Rader, J.S. The LncRNA PVT1 Contributes to the Cervical Cancer Phenotype and Associates with Poor Patient Prognosis. *PLoS ONE* **2016**, *11*, e0156274. [[CrossRef](#)] [[PubMed](#)]
53. Mainguéné, J.; Vacher, S.; Kamal, M.; Hamza, A.; Masliah-Planchon, J.; Baulande, S.; Ibadioune, S.; Borcoman, E.; Cacheux, W.; Calugaru, V.; et al. Human Papilloma Virus Integration Sites and Genomic Signatures in Head and Neck Squamous Cell Carcinoma. *Mol. Oncol.* **2022**. [[CrossRef](#)]
54. Kamal, M.; Lameiras, S.; Deloger, M.; Morel, A.; Vacher, S.; Lecerf, C.; Dupain, C.; Jeannot, E.; Girard, E.; Baulande, S.; et al. Human Papilloma Virus (HPV) Integration Signature in Cervical Cancer: Identification of MACROD2 Gene as HPV Hot Spot Integration Site. *Br. J. Cancer* **2021**, *124*, 777–785. [[CrossRef](#)] [[PubMed](#)]
55. Yang-chun, F.; Sen-yu, W.; Yuan, Z.; Yan-chun, H. Genome-Wide Profiling of Human Papillomavirus DNA Integration into Human Genome and Its Influence on PD-L1 Expression in Chinese Uygur Cervical Cancer Women. *J. Immunol. Res.* **2020**, *2020*, 6284960. [[CrossRef](#)] [[PubMed](#)]
56. Kataoka, K.; Ogawa, S. Novel Mechanism of Immune Evasion Involving PD-L1 in Various Cancers. *Transl. Cancer Res.* **2016**, *5*, S428–S432. [[CrossRef](#)]
57. Fan, Z.; Wu, C.; Chen, M.; Jiang, Y.; Wu, Y.; Mao, R.; Fan, Y. The Generation of PD-L1 and PD-L2 in Cancer Cells: From Nuclear Chromatin Reorganization to Extracellular Presentation. *Acta Pharm. Sin. B* **2022**, *12*, 1041–1053. [[CrossRef](#)]
58. Clark, A.; Burleson, M. SPOP and Cancer: A Systematic Review. *Am. J. Cancer Res.* **2020**, *10*, 704–726.
59. Pena, J.C.; Thompson, C.B.; Recant, W.; Vokes, E.E.; Rudin, C.M. Bcl-XL and Bcl-2 Expression in Squamous Cell Carcinoma of the Head and Neck. *Cancer* **1999**, *85*, 164–170. [[CrossRef](#)]
60. Rusan, M.; Li, Y.Y.; Hammerman, P.S. Genomic Landscape of Human Papillomavirus-Associated Cancers. *Clin. Cancer Res.* **2015**, *21*, 2009–2019. [[CrossRef](#)]
61. Hu, Z.; Zhu, D.; Wang, W.; Li, W.; Jia, W.; Zeng, X.; Ding, W.; Yu, L.; Wang, X.; Wang, L.; 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*, 158–163. [[CrossRef](#)] [[PubMed](#)]

62. Jiang, Y.; Zhu, C.; He, D.; Gao, Q.; Tian, X.; Ma, X.; Wu, J.; Das, B.C.; Severinov, K.; Hitzeroth, I.I.; et al. Cytological Immunostaining of HMGA2, LRP1B, and TP63 as Potential Biomarkers for Triaging Human Papillomavirus-Positive Women. *Transl. Oncol.* **2019**, *12*, 959–967. [[CrossRef](#)] [[PubMed](#)]
63. Yang, W.; Liu, Y.; Dong, R.; Liu, J.; Lang, J.; Yang, J.; Wang, W.; Li, J.; Meng, B.; Tian, G. Accurate Detection of HPV Integration Sites in Cervical Cancer Samples Using the Nanopore MinION Sequencer Without Error Correction. *Front. Genet.* **2020**, *11*, 660. [[CrossRef](#)] [[PubMed](#)]
64. Walline, H.M.; Komarck, C.M.; McHugh, J.B.; Bellile, E.L.; Brenner, J.C.; Prince, M.E.; McKean, E.L.; Chepeha, D.B.; Wolf, G.T.; Worden, F.P.; et al. Genomic Integration of High-Risk HPV Alters Gene Expression in Oropharyngeal Squamous Cell Carcinoma. *Mol. Cancer Res.* **2016**, *14*, 941–952. [[CrossRef](#)]
65. McBride, A.A.; Warburton, A. The Role of Integration in Oncogenic Progression of HPV-Associated Cancers. *PLoS Pathog.* **2017**, *13*, e1006211. [[CrossRef](#)]
66. Ernster, J.A.; Sciotto, C.G.; O'Brien, M.M.; Finch, J.L.; Robinson, L.J.; Willson, T.; Mathews, M. Rising Incidence of Oropharyngeal Cancer and the Role of Oncogenic Human Papilloma Virus. *Laryngoscope* **2007**, *117*, 2115–2128. [[CrossRef](#)]
67. Hollows, R.; Wei, W.; Cazier, J.-B.; Mehanna, H.; Parry, G.; Halford, G.; Murray, P. Association between Loss of Y Chromosome and Poor Prognosis in Male Head and Neck Squamous Cell Carcinoma. *Head Neck* **2019**, *41*, 993–1006. [[CrossRef](#)]
68. Nguyen, T.A.; Wu, K.; Pandey, S.; Lehr, A.W.; Li, Y.; Bembem, M.A.; Badger, J.D.; Lauzon, J.L.; Wang, T.; Zaghoul, K.A.; et al. A Cluster of Autism-Associated Variants on X-Linked NLGN4X Functionally Resemble NLGN4Y. *Neuron* **2020**, *106*, 759–768.e7. [[CrossRef](#)]
69. Bogaert, A.F.; Skorska, M.N.; Wang, C.; Gabrie, J.; MacNeil, A.J.; Hoffarth, M.R.; VanderLaan, D.P.; Zucker, K.J.; Blanchard, R. Male Homosexuality and Maternal Immune Responsivity to the Y-Linked Protein NLGN4Y. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 302–306. [[CrossRef](#)]
70. Gong, Y.; Wang, L.; Chippada-Venkata, U.; Dai, X.; Oh, W.K.; Zhu, J. Constructing Bayesian Networks by Integrating Gene Expression and Copy Number Data Identifies NLGN4Y as a Novel Regulator of Prostate Cancer Progression. *Oncotarget* **2016**, *7*, 68688–68707. [[CrossRef](#)]
71. Zhao, M.; Li, M.; Chen, Z.; Bian, Y.; Zheng, Y.; Hu, Z.; Liang, J.; Huang, Y.; Yin, J.; Zhan, C.; et al. Identification of Immune-Related Gene Signature Predicting Survival in the Tumor Microenvironment of Lung Adenocarcinoma. *Immunogenetics* **2020**, *72*, 455–465. [[CrossRef](#)] [[PubMed](#)]
72. Warburton, A.; Redmond, C.J.; Dooley, K.E.; Fu, H.; Gillison, M.L.; Akagi, K.; Symer, D.E.; Aladjem, M.I.; McBride, A.A. HPV Integration Hijacks and Multimerizes a Cellular Enhancer to Generate a Viral-Cellular Super-Enhancer That Drives High Viral Oncogene Expression. *PLoS Genet.* **2018**, *14*, e1007179. [[CrossRef](#)] [[PubMed](#)]
73. Muramoto, N.; Oda, A.; Tanaka, H.; Nakamura, T.; Kugou, K.; Suda, K.; Kobayashi, A.; Yoneda, S.; Ikeuchi, A.; Sugimoto, H.; et al. Phenotypic Diversification by Enhanced Genome Restructuring after Induction of Multiple DNA Double-Strand Breaks. *Nat. Commun.* **2018**, *9*, 1995. [[CrossRef](#)]
74. Burrell, R.A.; McClelland, S.E.; Endesfelder, D.; Groth, P.; Weller, M.-C.; Shaikh, N.; Domingo, E.; Kanu, N.; Dewhurst, S.M.; Gronroos, E.; et al. Replication Stress Links Structural and Numerical Cancer Chromosomal Instability. *Nature* **2013**, *494*, 492–496. [[CrossRef](#)] [[PubMed](#)]
75. Wallace, N.A. Catching HPV in the Homologous Recombination Cookie Jar. *Trends Microbiol.* **2020**, *28*, 191–201. [[CrossRef](#)]
76. Weaver, A.N.; Cooper, T.S.; Rodriguez, M.; Trummell, H.Q.; Bonner, J.A.; Rosenthal, E.L.; Yang, E.S. DNA Double Strand Break Repair Defect and Sensitivity to Poly ADP-Ribose Polymerase (PARP) Inhibition in Human Papillomavirus 16-Positive Head and Neck Squamous Cell Carcinoma. *Oncotarget* **2015**, *6*, 26995–27007. [[CrossRef](#)]
77. Nickson, C.M.; Moori, P.; Carter, R.J.; Rubbi, C.P.; Parsons, J.L. Misregulation of DNA Damage Repair Pathways in HPV-Positive Head and Neck Squamous Cell Carcinoma Contributes to Cellular Radiosensitivity. *Oncotarget* **2017**, *8*, 29963–29975. [[CrossRef](#)] [[PubMed](#)]
78. Hajek, M.; Biktasova, A.; Sewell, A.; Gary, C.; Cantalupo, P.; Anderson, K.S.; Yarbrough, W.G.; Issaeva, N. Global Genome Demethylation Causes Transcription-Associated DNA Double Strand Breaks in HPV-Associated Head and Neck Cancer Cells. *Cancers* **2020**, *13*, 21. [[CrossRef](#)] [[PubMed](#)]
79. Bamps, M.; Dok, R.; Nuyts, S. The DNA Damage Response Is Differentially Involved in HPV-Positive and HPV-Negative Radioresistant Head and Neck Squamous Cell Carcinoma. *Cancers* **2021**, *13*, 3717. [[CrossRef](#)]
80. Moody, C.A. Impact of Replication Stress in Human Papillomavirus Pathogenesis. *J. Virol.* **2019**, *93*, e01012-17. [[CrossRef](#)]
81. Cannataro, V.L.; Gaffney, S.G.; Sasaki, T.; Issaeva, N.; Grewal, N.K.S.; Grandis, J.R.; Yarbrough, W.G.; Burtneess, B.; Anderson, K.S.; Townsend, J.P. APOBEC-Induced Mutations and Their Cancer Effect Size in Head and Neck Squamous Cell Carcinoma. *Oncogene* **2019**, *38*, 3475–3487. [[CrossRef](#)]
82. Riva, G.; Albano, C.; Gugliesi, F.; Pasquero, S.; Pacheco, S.F.C.; Pecorari, G.; Landolfo, S.; Biolatti, M.; Dell'Oste, V. HPV Meets APOBEC: New Players in Head and Neck Cancer. *Int. J. Mol. Sci.* **2021**, *22*, 1402. [[CrossRef](#)] [[PubMed](#)]
83. Jakobsdottir, G.M.; Brewer, D.S.; Cooper, C.; Green, C.; Wedge, D.C. APOBEC3 Mutational Signatures Are Associated with Extensive and Diverse Genomic Instability across Multiple Tumour Types. *BMC Biol.* **2022**, *20*, 117. [[CrossRef](#)] [[PubMed](#)]
84. Kondo, S.; Wakae, K.; Wakisaka, N.; Nakanishi, Y.; Ishikawa, K.; Komori, T.; Moriyama-Kita, M.; Endo, K.; Muro, S.; Wang, Z.; et al. APOBEC3A Associates with Human Papillomavirus Genome Integration in Oropharyngeal Cancers. *Oncogene* **2017**, *36*, 1687–1697. [[CrossRef](#)] [[PubMed](#)]

85. Labarge, B.; Hennessy, M.; Zhang, L.; Goldrich, D.; Chartrand, S.; Purnell, C.; Wright, S.; Goldenberg, D.; Broach, J.R. Human Papillomavirus Integration Strictly Correlates with Global Genome Instability in Head and Neck Cancer. *Mol. Cancer Res.* **2022**, *20*, 1420–1428. [[CrossRef](#)] [[PubMed](#)]
86. Koneva, L.A.; Zhang, Y.; Virani, S.; Hall, P.B.; McHugh, J.B.; Chepeha, D.B.; Wolf, G.T.; Carey, T.E.; Rozek, L.S.; Sartor, M.A. HPV Integration in HNSCC Correlates with Survival Outcomes, Immune Response Signatures, and Candidate Drivers. *Mol. Cancer Res.* **2018**, *16*, 90–102. [[CrossRef](#)]
87. Pinatti, L.M.; Sinha, H.N.; Brummel, C.V.; Goudsmit, C.M.; Geddes, T.J.; Wilson, G.D.; Akervall, J.A.; Brenner, C.J.; Walline, H.M.; Carey, T.E. Association of Human Papillomavirus Integration with Better Patient Outcomes in Oropharyngeal Squamous Cell Carcinoma. *Head Neck* **2021**, *43*, 544–557. [[CrossRef](#)]