

G OPEN ACCESS

Citation: McBride AA, Warburton A (2017) The role of integration in oncogenic progression of HPV-associated cancers. PLoS Pathog 13(4): e1006211. https://doi.org/10.1371/journal.ppat.1006211

Editor: Katherine R. Spindler, University of Michigan Medical School, UNITED STATES

Published: April 6, 2017

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the <u>Creative</u> <u>Commons CC0</u> public domain dedication.

Funding: The research of the McBride laboratory is supported by the Intramural Research Program of the NIAID, NIH (NIAID intramural 1ZIAAI001073-08, https://projectreporter.nih.gov/project_info_ description.cfm?aid=9161637&icde= 32244971&ddparam=&ddvalue=&ddsub=&cr= 4&csb=default&cs=ASC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

PEARLS

The role of integration in oncogenic progression of HPV-associated cancers

Alison A. McBride*, Alix Warburton

Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, United States of America

* amcbride@nih.gov

Introduction

Persistent infection with a subset of "high oncogenic risk" human papillomaviruses (HPVs) can promote the development of cancer. In these cancers, the extrachromosomal viral genome has often become integrated into the host genome. The integration event is thought to drive oncogenesis by dysregulating expression of the E6 and E7 viral oncogenes, leading to inactivation of critical cell cycle checkpoints and increased genetic instability in the host. This Pearl reviews the evidence that gave rise to the current textbook paradigm of HPV integration events and their consequences and incorporates new findings that demonstrate that stochastic integration events can promote oncogenesis in many ways.

Is HPV integration part of the papillomavirus life cycle?

Papillomaviruses have a resourceful life cycle that takes advantage of the tissue renewal process of stratified epithelia. Only the lower, basal cells in the epithelium proliferate, but they can divide either symmetrically (to produce more basal cells) or asymmetrically (one of the daughter cells leaves the basal layer and begins the differentiation process). Differentiating daughter cells move up through the epithelium, acquiring specialized properties until they are released from the surface as part of the process of tissue renewal. Papillomaviruses exploit this process; they access and infect the basal cells through a micro-abrasion and establish a long-term infection in these dividing cells. The viral E1 and E2 proteins support replication of the viral genome at a low copy number in the basal cells as a small, dsDNA nuclear plasmid of about 7–8 kbp. It is only when these infected cells differentiate and move towards the surface of the epithelium that high levels of viral DNA are synthesized, packaged in virions, and sloughed from the surface of the epithelium in viral-laden squames [1]. HPVs are often found integrated in premalignant lesions and a range of anogenital and oropharyngeal cancers [2–4], but this is not part of the viral life cycle. In fact, integration is a dead end for the virus, as it is no longer able to form a small, circular genome that can be packaged and transmitted to a new host.

How does integration of HPV promote oncogenesis?

Almost all HPV integration events that have been studied in detail to date are related to HPV oncogenesis. HPV integration events can be detected in premalignant lesions, but the percentage of cells containing integrated HPV increases as cells progress to invasive cancer [5]. Integration usually results in dysregulation of expression of the viral E6 and E7 oncogenes, which promotes cellular proliferation, abrogates cell cycle checkpoints, and causes progressive genetic instability. This gives cells a selective growth advantage and promotes oncogenic progression [6]. Clonal outgrowth of cells with integrated HPV highlights the importance of dysregulated oncogene expression. In fact, HPV-associated cancers are dependent on the expression of the viral E6 and E7 oncogenes for continued proliferation and survival [7].

HPV integration can be classified into two types: in Type 1, a single genome is integrated into cellular DNA; and in Type 2, multiple tandem head-to-tail repeats of the genome, in some cases with intervening cellular flanking sequences, are found at a single genomic locus [6] (see Fig 1). In both types, RNA encoding the E6 and E7 oncogenes is initiated from the major promoter in the viral upstream regulatory region (URR) and is spliced from a splice donor in the viral genome to a splice acceptor in the host DNA [8]. In Type 2 integration, there is evidence that usually only the 3' junctional copy of the viral genome is transcriptionally active [9]. Continued expression of the E1 and E2 replication proteins from integrated genomes causes focal genomic instability at the integration locus [10, 11], and in most cases, when complete genomes are tandemly integrated, the internal copies are silenced by DNA methylation [12].

Viral genome integration events usually result in dysregulation of E6 and E7 gene expression compared to that expressed from extrachromosomal viral genomes, and this can be achieved in a number of ways (see Fig 2 and Table 1). The earliest model proposed that the integration event disrupts the E2 gene, alleviating E2 transcriptional repression of the E6 and

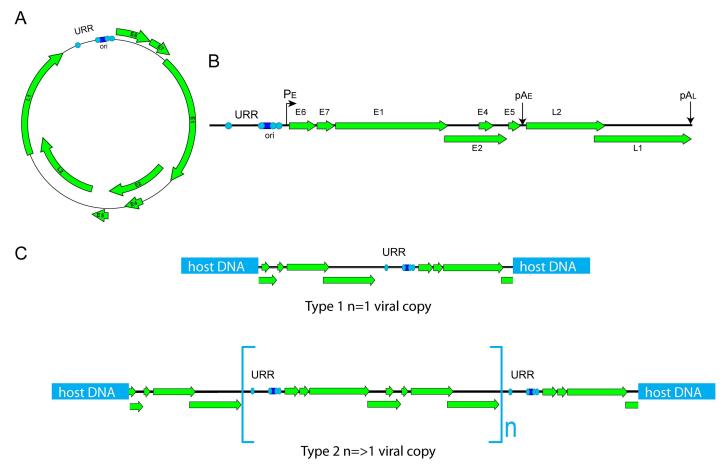


Fig 1. Types of HPV integration. A. Circular HPV genome. B. Linear HPV genome. URR (upstream regulatory region), P_E (early promoter), and pA_E and pA_L (early and late polyadenylation sites) are indicated. The light blue circles in the URR represent E2 binding sites, and the dark blue square is the E1 binding site in the origin of replication (ori). C. In Type 1 integration, a single viral genome is integrated into the host DNA. In Type 2 integration, multiple genomes are integrated in tandem in a head-to-tail orientation. This often is accompanied by focal rearrangement and amplification of flanking cellular sequences.

https://doi.org/10.1371/journal.ppat.1006211.g001

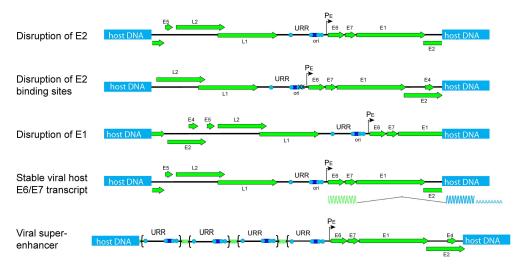


Fig 2. Models of integration events that promote oncogenesis. The five integration models shown in Table 1 are shown in the diagram, as indicated to the left. URR (upstream regulatory region), and P_E (early promoter), are indicated. The light blue circles in the URR represent E2 binding sites, and the dark blue square is the E1 binding site in the origin of replication (ori).

https://doi.org/10.1371/journal.ppat.1006211.g002

E7 promoter and thus driving oncogene expression [13]. E2 regulation can also be disrupted by methylation of the E2 binding sites in the URR [14]. E6/E7 oncogene expression can also be modulated by epigenetic events that do not directly affect E2 DNA binding. Our group recently demonstrated that tandemly integrated repeats of HPV16 DNA could develop into a Brd4-dependent super-enhancer to drive strong expression of the viral oncogenes [15]. Viral genomes are also often integrated in such a way as to disrupt the gene that encodes the E1 replicative helicase, which also disrupts the downstream E2 gene. In HPV infection, the level and nuclear location of the E1 protein are tightly regulated because uncontrolled E1 expression can cause DNA damage and growth arrest [16, 17]. E1 expression can also promote focal genomic instability by inducing overamplification of the integration region [10, 11]. Thus, disruption of the E1 gene could give a selective growth advantage and promote clonal expansion. As

Table 1. Integration events that promote oncogenesis.

Integration event	Effect on oncogenic progression
Disruption of E2 gene	Abrogation of E2-mediated transcriptional repression of the E6/E7 promoter
Disruption of E2 binding sites by methylation	Abrogation of E2-mediated transcriptional repression of the E6/E7 promoter
Disruption of E1 gene	Abrogation of E1 replication activities that can induce DNA damage and growth arrest and promote focal genomic instability at the integration locus
Formation of viral—host fusion transcript	Hybrid E6/E7 transcripts are often more stable than viral E6/E7 transcripts
Tandem repeats of HPV genome/regulatory elements	Internal copies are usually silenced, but sometimes tandem copies of active genomes can function as a transcriptional super-enhancer
Altered regulation of cancer-associated genes in the vicinity of the integration locus	Integrated HPV can disrupt cellular genes and/or their flanking sequences, altering their expression and/or the expression of nearby genes; infrequently, HPV integration alters the expression of nearby cancer genes, thus promoting carcinogenesis

https://doi.org/10.1371/journal.ppat.1006211.t001

mentioned above, most integration events result in expression of a spliced viral—cellular transcript. Jeon and Lambert demonstrated that these fusion transcripts are very often more stable than their viral counterparts, yet again increasing HPV oncogene expression [18]. There are also cases in which an oncogenic HPV has integrated in the vicinity of a cellular oncogene or tumor suppressor gene [2, 19], but this is not thought to be a universal way in which HPV promotes oncogenesis.

Do all HPV-associated cancers contain integrated HPV DNA?

Although many HPV-associated cancers contain integrated viral DNA, it is not universal. HPV-associated cancers can contain either integrated HPV DNA, extrachromosomal viral DNA, or a mix of both [20]. However, in tumors with exclusively extrachromosomal viral DNA, the viral genome has usually acquired genetic or epigenetic changes that result in dysregulated E6/E7 gene expression [12, 21, 22]. For example, methylation of the E2 binding sites in the URR can alleviate E2-mediated repression of the viral oncogenes [22–24]. Therefore, while integration is very frequently detected in HPV-associated cancers, it is not absolutely required.

Analysis of samples from The Cancer Genome Atlas study shows that HPV integration occurs in >80% of HPV-positive cervical cancers [25]. Of these, 76% of HPV16-positive samples have integrated HPV, whereas integration is evident in all HPV18-positive samples. This confirms early observations of different frequencies of integration between HPV16 and HPV18 [26]. In HPV-positive oropharyngeal squamous cell carcinomas, the incidence of viral integration is lower, and many tumors have either extrachromosomal or mixed extrachromosomal and integrated viral DNA [2, 27–29]. The rate of HPV integration in other anogenital cancers is not as well documented. One study reports that almost 80% of anal carcinomas contain integrated HPV; however, the vast majority of these samples also contain extrachromosomal genomes [30]. Many different methods have been used to detect integrated viral genomes, and more research is needed to clarify the significance of the observed differential rates of integration and to determine whether HPVs promote oncogenesis by different mechanisms at different anatomical locations.

Are there specific integration target sequences in the human and viral genomes?

Genome-wide efforts to elucidate a genomic signature of HPV integration events have identified only a handful of recurrent loci. These so-called genomic "hotspots" are highly correlated with common fragile sites [31] and transcriptionally active regions of the genome [19, 32]. Regions of microhomology (1–10 bp) between viral and human genomic sequences are sometimes found at integration breakpoints [2], as have AT-rich regions of the genome that have the potential to form stem—loop structures that promote the formation of stalled replication forks during replication stress [28]. There are also examples of HPV integration resulting in insertional mutagenesis and/or potential regulatory effects on neighbouring genes [33]. Increased integration within cancer-associated genes or pathways, including the MYC gene locus, have also been reported [2, 34]. However, this is not a universal phenomenon associated with HPV integration [4].

The best-characterized HPV integration sites are those responsible for HPV cervical cancers. In these sites, the E6 and E7 genes are invariably expressed from viral promoters in the URR, and the viral genome is most often disrupted in the E1 and E2 genes, leading to the models described above. With the advent of more sensitive techniques, it is becoming apparent that there are often multiple integration sites, and many are disrupted at positions throughout the viral genome [35]. However, it is likely that these are background/silent integration events [36], resulting from overall increased genetic instability, and that most of these genomes are passengers rather than drivers of oncogenesis. Likely, non-oncogenic HPVs also occasionally become integrated, but without oncogene-driven clonal expansion, these rare events are almost never detected.

It is becoming clear that there is very often rearrangement and amplification of cellular sequences flanking the HPV integration site [2, 35, 37]. The initial integration event likely occurs in a cell that also harbors extrachromosomal genomes that express the E1 and E2 replication proteins. In this scenario, the E1 and E2 proteins cause overamplification, or onion skin replication, of the integrated HPV genome and adjacent sequences [10, 11]. This results in heterogeneous replication intermediates that serve as substrates for recombination and repair, resulting in rearrangements, deletions, and amplification of viral and host sequences [10, 11]. A looping model also describes how viral—host DNA concatemers can arise from the formation of a transient loop that acts as a substrate for rolling circle replication [35]. Additional rearrangements could be further exacerbated by the inherent genomic instability of the host region (e.g., common fragile sites), from the instability of the tandemly repeated locus, and from continuing E6/E7-mediated genetic instability.

What promotes HPV integration?

There are several stages in the development of an HPV integration site that eventually drives oncogenic progression. The initial integration event likely takes place in regions where the viral and host DNA are in close proximity in the proliferating basal cells of a lesion. HPV hijacks the host DNA damage response to replicate its own DNA in certain phases of the life cycle [26, 27], and this occurs adjacent to regions of the host DNA susceptible to replication stress (e.g., common fragile sites) [38]. This could explain the preferential integration in these regions [28]. Secondly, the target region must be in a transcriptionally competent region of host chromatin that can support viral oncogene expression. In most cases, E6/E7 mRNA is expressed as a viral—host fusion transcript, and this necessitates the presence of a nearby cellular splice acceptor and polyadenylation site (Often cryptic). Finally, it is likely that there is epigenetic modulation of the integration site (DNA methylation and chromatin modifications) that further determine whether the integration site is active or silenced [12, 39]. Therefore, many events and processes contribute to the development of an HPV integration event that is a strong driver of oncogenesis; there are likely many dead-end integration events that fail to produce sufficient E6/E7 oncoproteins to drive clonal expansion of the host cell.

References

- 1. Doorbar J, Egawa N, Griffin H, Kranjec C, Murakami I. Human papillomavirus molecular biology and disease association. Rev Med Virol. 2015;25 Suppl 1(S1):2–23.
- 2. 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(43):15544–9. Epub 2014/10/15. https://doi.org/10.1073/pnas.1416074111 PMID: 25313082
- Schwarz E, Freese UK, Gissmann L, Mayer W, Roggenbuck B, Stremlau A, et al. Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. Nature. 1985; 314 (6006):111–4. PMID: 2983228
- 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(11):3878–84. Epub 2004/06/03. https://doi.org/10.1158/0008-5472.CAN-04-0009 PMID: 15172997
- Shukla S, Mahata S, Shishodia G, Pande S, Verma G, Hedau S, et al. Physical state & copy number of high risk human papillomavirus type 16 DNA in progression of cervical cancer. The Indian journal of medical research. 2014; 139(4):531–43. Epub 2014/06/14. PMID: 24927339

- 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(5):2989–97. Epub 1995/ 05/01. PMID: 7707525
- Goodwin EC, Yang E, Lee CJ, Lee HW, DiMaio D, Hwang ES. Rapid induction of senescence in human cervical carcinoma cells. Proc Natl Acad Sci U S A. 2000; 97(20):10978–83. Epub 2000/09/27. PMID: 11005870
- Wentzensen N, Ridder R, Klaes R, Vinokurova S, Schaefer U, Doeberitz M. Characterization of viralcellular fusion transcripts in a large series of HPV16 and 18 positive anogenital lesions. Oncogene. 2002; 21(3):419–26. https://doi.org/10.1038/sj.onc.1205104 PMID: 11821954
- 9. Van Tine BA, Kappes JC, Banerjee NS, Knops J, Lai L, Steenbergen RD, et al. Clonal selection for transcriptionally active viral oncogenes during progression to cancer. JVirol. 2004; 78(20):11172–86.
- Kadaja M, Isok-Paas H, Laos T, Ustav E, Ustav M. Mechanism of genomic instability in cells infected with the high-risk human papillomaviruses. PLoS Pathog. 2009; 5(4):e1000397. Epub 2009/04/25. https://doi.org/10.1371/journal.ppat.1000397 PMID: 19390600
- Kadaja M, Sumerina A, Verst T, Ojarand M, Ustav E, Ustav M. Genomic instability of the host cell induced by the human papillomavirus replication machinery. EMBO J. 2007; 26(8):2180–91. https://doi. org/10.1038/sj.emboj.7601665 PMID: 17396148
- Chaiwongkot A, Vinokurova S, Pientong C, Ekalaksananan T, Kongyingyoes B, Kleebkaow P, et al. Differential methylation of E2 binding sites in episomal and integrated HPV 16 genomes in preinvasive and invasive cervical lesions. Int J Cancer. 2013; 132(9):2087–94. <u>https://doi.org/10.1002/ijc.27906</u> PMID: 23065631
- 13. Thierry F, Yaniv M. The Bpv1-E2 Trans-Acting Protein Can Be Either an Activator or a Repressor of the Hpv18 Regulatory Region. Embo Journal. 1987; 6(11):3391–7. PMID: 2828029
- Leung TW, Liu SS, Leung RC, Chu MM, Cheung AN, Ngan HY. HPV 16 E2 binding sites 1 and 2 become more methylated than E2 binding site 4 during cervical carcinogenesis. J Med Virol. 2015; 87(6):1022–33. Epub 2015/02/05. https://doi.org/10.1002/jmv.24129 PMID: 25648229
- 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(5).
- Sakakibara N, Mitra R, McBride AA. The papillomavirus E1 helicase activates a cellular DNA damage response in viral replication foci. J Virol. 2011; 85(17):8981–95. https://doi.org/10.1128/JVI.00541-11 PMID: 21734054
- Fradet-Turcotte A, Bergeron-Labrecque F, Moody CA, Lehoux M, Laimins LA, Archambault J. Nuclear accumulation of the papillomavirus E1 helicase blocks S-phase progression and triggers an ATMdependent DNA damage response. J Virol. 2011; 85(17):8996–9012. https://doi.org/10.1128/JVI. 00542-11 PMID: 21734051
- Jeon S, Lambert PF. Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis. ProcNatlAcadSciUSA. 1995; 92:1654–8.
- Bodelon C, Untereiner ME, Machiela MJ, Vinokurova S, Wentzensen N. Genomic characterization of viral integration sites in HPV-related cancers. Int J Cancer. 2016; 139(9):2001–11. Epub 2016/06/28. https://doi.org/10.1002/ijc.30243 PMID: 27343048
- Kristiansen E, Jenkins A, Holm R. Coexistence of episomal and integrated HPV16 DNA in squamous cell carcinoma of the cervix. J Clin Pathol. 1994; 47(3):253–6. Epub 1994/03/01. PMID: 7677803
- Dong XP, Stubenrauch F, Beyer-Finkler E, Pfister H. Prevalence of deletions of YY1-binding sites in episomal HPV 16 DNA from cervical cancers. IntJCancer. 1994; 58(6):803–8.
- 22. Bhattacharjee B, Sengupta S. CpG methylation of HPV 16 LCR at E2 binding site proximal to P97 is associated with cervical cancer in presence of intact E2. Virology. 2006; 354(2):280–5. <u>https://doi.org/10.1016/j.virol.2006.06.018 PMID</u>: 16905170
- Kim K, Garner-Hamrick PA, Fisher C, Lee D, Lambert PF. Methylation patterns of papillomavirus DNA, its influence on E2 function, and implications in viral infection. JVirol. 2003; 77(23):12450–9.
- Reuschenbach M, Huebbers CU, Prigge ES, Bermejo JL, Kalteis MS, Preuss SF, et al. Methylation status of HPV16 E2-binding sites classifies subtypes of HPV-associated oropharyngeal cancers. Cancer. 2015; 121(12):1966–76. Epub 2015/03/04. https://doi.org/10.1002/cncr.29315 PMID: 25731880
- 25. Integrated genomic and molecular characterization of cervical cancer. Nature. 2017. Epub 2017/01/24.
- Cullen AP, Reid R, Campion M, Lorincz AT. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. J Virol. 1991; 65(2):606–12. Epub 1991/ 02/01. PMID: 1846186

- Vojtechova Z, Sabol I, Salakova M, Turek L, Grega M, Smahelova J, et al. Analysis of the integration of human papillomaviruses in head and neck tumours in relation to patients' prognosis. Int J Cancer. 2016; 138(2):386–95. https://doi.org/10.1002/ijc.29712 PMID: 26239888
- Gao G, Johnson SH, Vasmatzis G, Pauley CE, Tombers NM, Kasperbauer JL, et al. Common Fragile Sites and Extremely Large are Targets for Human Papillomavirus Integrations and Chromosome Rearrangements in Oropharyngeal Squamous Cell Carcinoma. Genes Chromosomes Cancer. 2016. Epub 2016/09/17.
- 29. 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(2):e88718. https://doi.org/10.1371/journal.pone.0088718 PMID: 24586376
- Valmary-Degano S, Jacquin E, Pretet JL, Monnien F, Girardo B, Arbez-Gindre F, et al. Signature patterns of human papillomavirus type 16 in invasive anal carcinoma. Human pathology. 2013; 44(6):992–1002. Epub 2012/12/26. https://doi.org/10.1016/j.humpath.2012.08.019 PMID: 23266444
- Thorland EC, Myers SL, Persing DH, Sarkar G, McGovern RM, Gostout BS, et al. Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites. Cancer Res. 2000; 60(21):5916–21. Epub 2000/11/21. PMID: 11085503
- Christiansen IK, Sandve GK, Schmitz M, Durst M, Hovig E. Transcriptionally active regions are the preferred targets for chromosomal HPV integration in cervical carcinogenesis. PLoS ONE. 2015; 10(3): e0119566. Epub 2015/03/21. https://doi.org/10.1371/journal.pone.0119566 PMID: 25793388
- Durst M, Croce CM, Gissmann L, Schwarz E, Huebner K. Papillomavirus Sequences Integrate near Cellular Oncogenes in Some Cervical Carcinomas. Proceedings of the National Academy of Sciences of the United States of America. 1987; 84(4):1070–4. PMID: 3029760
- Ferber MJ, Thorland EC, Brink AA, Rapp AK, Phillips LA, McGovern R, et al. Preferential integration of human papillomavirus type 18 near the c-myc locus in cervical carcinoma. Oncogene. 2003; 22 (46):7233–42. https://doi.org/10.1038/sj.onc.1207006 PMID: 14562053
- Akagi K, Li J, Broutian TR, Padilla-Nash H, Xiao W, Jiang B, et al. Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. Genome Res. 2014; 24(2):185– 99. https://doi.org/10.1101/gr.164806.113 PMID: 24201445
- 36. Xu B, Chotewutmontri S, Wolf S, Klos U, Schmitz M, Durst M, et al. Multiplex Identification of Human Papillomavirus 16 DNA Integration Sites in Cervical Carcinomas. PLoS ONE. 2013; 8(6):e66693. https://doi.org/10.1371/journal.pone.0066693 PMID: 23824673
- Peter M, Stransky N, Couturier J, Hupe P, Barillot E, de Cremoux P, et al. Frequent genomic structural alterations at HPV insertion sites in cervical carcinoma. J Pathol. 2010; 221(3):320–30. Epub 2010/06/ 09. https://doi.org/10.1002/path.2713 PMID: 20527025
- Jang MK, Shen K, McBride AA. Papillomavirus genomes associate with BRD4 to replicate at fragile sites in the host genome. PLoS Pathog. 2014; 10(5):e1004117. <u>https://doi.org/10.1371/journal.ppat.</u> 1004117 PMID: 24832099
- **39.** Groves IJ, Knight EL, Ang QY, Scarpini CG, Coleman N. HPV16 oncogene expression levels during early cervical carcinogenesis are determined by the balance of epigenetic chromatin modifications at the integrated virus genome. Oncogene. 2016.