Ménage à trois: an evolutionary interplay between human papilloma virus, a tumor and a woman

The Faculty of Oregon State University has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Shulzhenko, N., Lyng, H., Sanson, G. F., &amp; Morgun, A. (2014). Menage a trois: An evolutionary interplay between human papillomavirus, a tumor, and a woman. Trends in Microbiology, 22(6), 345-353. doi:10.1016/j.tim.2014.02.009</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOI</td>
<td>10.1016/j.tim.2014.02.009</td>
</tr>
<tr>
<td>Publisher</td>
<td>Elsevier/Cell Press</td>
</tr>
<tr>
<td>Version</td>
<td>Accepted Manuscript</td>
</tr>
<tr>
<td>Terms of Use</td>
<td><a href="http://cdss.library.oregonstate.edu/sa-termsofuse">http://cdss.library.oregonstate.edu/sa-termsofuse</a></td>
</tr>
</tbody>
</table>
Ménage à trois: an evolutionary interplay between human papilloma virus, a tumor and a woman

Natalia Shulzhenko1, Heidi Lyng3, Gerdine F. Sanson4, Andrey Morgun2

Author affiliations: 1College of Veterinary Medicine and 2College of Pharmacy, Oregon State University, Corvallis, OR, USA; 3Department of Radiation Biology, Norwegian Radium Hospital, Oslo University Hospital, Oslo, Norway; 4Institute of Health Sciences, Federal University of Mato Grosso, Sinop, MT, Brazil

Corresponding author: Morgun, A (anemorgun@hotmail.com; andriy.morgun@oregonstate.edu)

Keywords: cervical cancer; human papillomavirus; immune response; gene expression; genomic aberrations; network analysis; driver genes
Abstract

Cervical cancer is the third most common cancer in women with Human papillomavirus (HPV) being a key etiologic factor of this devastating disease. In this article, we describe modern advances in the genomics and transcriptomics of cervical cancer that led to uncovering the key gene drivers. We also introduce, herein, a Model of Cervical Carcinogenesis that explains how the interplay between virus, tumor and woman results in the selection of clones that simultaneously harbor genomic amplifications for genes that drive cell cycle, antiviral response, and inhibit cell differentiation. The new model may help understanding controversies in antiviral therapy and immunogenetics of this cancer and may provide a basis for future research directions in early diagnostics and personalization of therapy.

Glossary

- **High-risk HPV**: oncogenic types of HPV which can cause cancer; HPV16 and 18 are the most frequent oncogenic types.
- **Low-grade and high-grade CIN**: stages of pre-cancerous lesions, also called dysplasia, in the uterine cervix characterized by abnormal epithelial cell growth ranging from mild to severe degrees.
- **Hallmarks of Cancer**: proposed by Hanahan and Weinberg [ref], these include common properties of cancers in sustaining proliferative signaling, resisting cell death, evading growth suppressors, activating invasion and metastasis, enabling replicative immortality, inducing angiogenesis.
Cervical cancer and human papillomavirus

Cervical cancer is the third most common cancer in women, with an estimated 454,000 new cases and 200,000 deaths each year [1]. More than 85% of the global burden occurs in developing countries, where it accounts for 13% of all female cancers and where most deaths occur. Human papillomavirus (HPV) infection is necessary, although not sufficient, to trigger the disease and it is estimated that almost three hundred million women worldwide will have a HPV infection of the cervix at any given point of their lives. The efforts to develop prophylactic and therapeutic vaccines against HPV infection resulted in the two vaccines (Cervarix and Gardasil) that are already in use, offering protection against the most prevalent HPV types associated with cancer, HPV-16 and HPV-18. Although both vaccines confer near-complete protection against infection for these HPV types, the high cost and difficult vaccination schemes - 3 injected doses over 6 months – make cost effectiveness an enormous drawback in less developed countries, where mass vaccination is more critical. Furthermore, despite significant progress, it is recognized that the story of HPV is still being written and there are gaps in the knowledge of how HPV causes cervical cancer [2-6]. In this article, we describe how recent meta-analyses of gene expression and genomic aberrations aided by regulatory network reconstruction advanced our understanding of cervical carcinogenesis.

High-risk HPV and stages of infection

HPVs are a family of non-enveloped double strand DNA viruses of more than 180 types [7] that have been grouped in five genera based on DNA [8]. Genome organization of most viral genotypes are similar, comprising a circular DNA of approximately 7900 bp with three functional coding regions: a region coding for early viral function (E) representing genes involved in viral genome regulation, replication and modification of host cell processes; a region of late viral function (L) encoding
capsid proteins; and a long control region (LCR) which lies between them [9, 10]. All eight virus genes (E1, E2, E4-E7, L1, L2) are transcribed from the same DNA strand as two polycistronic entities (for early and late genes) and transcripts are processed through numerous alternative splicing events, rendering many more than eight gene products (reviewed in [11]).

Over 14 high-risk or oncogenic HPVs have been identified, among which HPV16 and HPV18 are most frequently found in cervical cancer, contributing to over 70% of cases (HPV16, 54.4%; HPV18, 16.5%). The five next most frequent high-risk types include HPV58 (5.1%), HPV33 (4.7%), HPV45 (4.4%), HPV31 (3.6%), HPV52 (3.4%), with others contributing to less than 2% of cases, individually [3]. First, virus infects proliferating basal cells of stratified squamous epithelia, where viral DNA is released from the capsid and transported into the nucleus as free genetic material or extrachromosomal episomes. Next, the early HPV promoter is activated within the basal cells, whereas expression of the viral E6 and E7 oncoproteins is repressed through a strict control of the early promoter by mechanisms involving the E2 protein [12]. Only low levels of viral DNA synthesis occur, and the episomal copy number increases to approximately 50-100 genomes per cell. As the basal cells differentiate, cease to divide and migrate towards the suprabasal layers, the differentiation-dependent late HPV promoter is activated [13]. This leads to increased E6 and E7 expression, which reactivates the cellular DNA synthesis and prevents apoptosis [14]. Consequently, the viral genome replicates to hundreds or even thousands of copies, followed by capsid protein L1 and L2 synthesis, virion assembly and release of new viral particles in the upper layers of the epithelium, ready to infect new cells in the basal layer. This is the normal viral life cycle and the productive phase of the infection. It is strongly coupled to the differentiation program of the infected epithelium and lasts about 2-3 weeks. The infection can be seen as mild epithelial dysplasia, characterized
as low-grade cervical intraepithelial neoplasia (CIN1) [15], and is in most cases cleared by the woman's immune system.

In approximately 10-15% of cases, a persistent HPV infection occurs possibly due to a combination of viral mechanisms inhibiting and escaping immune surveillance and certain deficiencies in women's immune response (Figure 1) [16-18]. In such scenarios, infected cells remain proliferative without undergoing apoptosis and the natural viral life cycle is aborted. In this abortive or transformed phase of the infection, a pronounced increase in E6 and E7 expression actively blocks negative regulators of the cell cycle through well-characterized interactions and degradation of the tumors suppressor proteins RB1 and TP53, preventing the cell maturation. This can be identified by an overexpression of p16 (INK4a), a transcriptional target of RB1 and a commonly used marker of transforming HPV infection [15]. At this stage, the infection appears as moderate or severe epithelial dysplasia or high-grade intraepithelial lesion (CIN2/3) [15] (Figure 1). Such deregulation of growth control in the infected cells is a rare by-product of the infection, and a crucial step towards malignancy.

A further step in the neoplastic progression is integration of HPV DNA into the host genome, the mechanisms for which are not quite clear [19, 20] [21, 22]. Increased genetic instability following the transforming infection through enhanced expression and stabilization of viral transcripts may play a role in the integration process [23], and further enhance the chromosomal instability, as discussed below. At integration, the E2 open reading frame is usually disrupted, abolishing E2 expression in the integrant[24], [25]. Expression of integrated E6 and E7 oncogenes is, however, controlled by the episomal E2, which selectively reduces expression of the integrated but not the episomal proteins [26]. It was suggested that a more open conformation of the integrated HPV genome could contribute to the differential effect of E2 on the integrated vs. episomal E6/E7 expression [22]. Thus integration of HPV into human genome sets a stage for E6 and E7 overexpression, which was shown to be abolished.
experimentally by re-introduction of the E2 protein into cancer cells leading to reduced cell proliferation and activation of apoptosis[27, 28].

Hence, the transcriptional regulatory effects of E2 appear to depend on the physical state of the virus, and the background levels of episomal HPV in the cells can hinder E6 and E7 expression by the integrant. The integrated HPV can therefore exist in a minority of cells as a relatively silent integrant for long periods with no selective growth advantage.

The change in the physical state of the virus from the episomal to integrated form is strongly associated with the severity of the neoplastic lesion. CIN1 lesions almost exclusively contain episomal HPV, whereas in CIN2/3 lesions the mixed (episomal plus integrated) forms start to emerge and can be seen in up to 75% of the cases depending on the study (Figure 1). At invasive stages, the pure episomal form is hardly seen. The integrated form is the most common (45-80%), although several cases contain mixed forms [29-36]. These observations strongly suggest a selective growth advantage of cells with integrated HPV.

Changes in the human genome

It has become clear that integration of the HPV DNA is associated with high level chromosomal instability [37], probably due to stabilization of E6 and E7 mRNA levels [19], increased protein levels and a high proliferation rate of the integrant [20, 37]. Important insight into the onset of chromosomal instability has been derived from studies on the W12 cervical keratinocyte cell line, which is an excellent model system for exploring changes in the human genome in relation to the physical state of the virus. The cells are infected with high risk HPV16 and contain episomes at approximately 100 copies per cell at early passages and integrated HPV genome after long term cultivation [37]. By use of this model system, chromosomal instability in the form of amplifications, gains, and losses of large chromosomal regions have been
shown to emerge in the presence of both integrated and episomal genomes but not in cells with pure episomes [37]. A study reviewing almost 200 HPV integration sites did not identify preferential integration into cancer-related genes but rather into chromosomal fragile sites [38]. However, a recent high resolution global sequencing of cell lines [39] and primary tumors [40] revealed that HPV integrants were frequently located adjacent to host genomic aberrations and, in some cases, to genes involved in oncogenesis.

In accordance with the cell line results, studies on clinical samples show a pronounced increase in chromosomal aberrations when comparing low-grade and high-grade CIN lesions. The clinical data may also pinpoint aberrations that could be crucial for the malignant progression of the lesions in an environment influenced by the host. A meta-analysis based on 12 published studies identified gain of 3q as the most common chromosomal aberration in CIN2/3 lesions and was seen in about 30% of the cases, but not in CIN1 lesions [41]. The 3q gains were also the most common aberration at invasive stages (60%), as reported previously [42]. 3q gain was particularly frequent in cancers infected with the HPV-16 virus type (84%) [41], which is associated with the highest risk of progression and shortest progression time towards high-grade dysplasia [43]. Common aberrations in both high-grade CIN lesions (CIN2/3) and cervical cancers included also gain of chromosome 1 and 20 and loss of 2q and 4. Another meta-analysis [44], involving only studies with invasive cervical carcinoma demonstrated similar results (Figure 2) with 3q gain being the most frequent followed by gains in 1q, 1p, and 20q. Importantly, most tumors had gains of four or more chromosomal regions suggesting their synergistic effect on disease progression [44]. Altogether, these studies propose that specific chromosomal gains and losses that may provide a selection advantage during the progression from low-grade to high-grade dysplasia and invasive cancer.

**Gene expression in tumors**
Evaluation of global gene expression became a commonly used strategy to understand pathways operating in cancer as well as to identify potential biomarkers and drug targets [45]. Cervical cancer has also been explored using this approach by multiple groups throughout the globe [46-51]. Unsurprisingly, one common observation for most of those studies was a detection of dysregulation in cell cycle and in epithelial cytoskeleton genes (reviewed by [52]). A handful of studies reported a few individual immune genes as overexpressed in cervical carcinoma [53, 54]. Despite consistency on the pathway level, reported individual genes varied considerably from one study to another. This potentially could be explained by two factors: inconsistency in detection/analysis and heterogeneity of disease. Independent of the actual reason for these discrepancies, it is clear that meta-analysis of different studies could provide a better overview on expression phenotype of cervical cancer. Indeed, in one gene expression meta-analysis a robust meta-signature of cervical cancer was established consisting of 742 up- and 546 downregulated genes [44].

Although identifying a reproducible gene signature of disease is a critical piece of analysis it is only a first step in the understanding of a gene expression phenotype. The reconstruction of regulatory networks emerged recently as an efficient method to aid the interpretation of gene expression data from different diseases including cancer [55-57]. In fact, by reconstructing and analyzing regulatory networks from differentially expressed genes revealed by meta-analysis and performing next level meta-analysis for regulatory networks, three major pathways defining an expression phenotype of cervical cancers have been established including: upregulated cell cycle, downregulated epithelial cell differentiation and upregulated antiviral response. While the first two pathways were to some degree positive controls validating earlier analyses, antiviral response was more unexpected. Despite HPV being the main etiological factor of cervical carcinoma and antiviral response per se being an issue of a continuous intensive research in this field, the common notion in the literature is that antiviral response is dampened in women who develop cervical cancer [58, 59].
The antiviral immune response in the course of HPV infection seems to be more complex though. Indeed, only a small proportion of women (~10%) do not eliminate HPV after getting infected. And a plausible explanation for this is that immune system of this relatively rare subpopulation of women might have a certain deficiency that precludes virus elimination [18]. But paradoxically, transition into invasive cancer is accompanied by a dramatic decrease in episomal virus (Figure 1), which was proposed to be caused by host antiviral immune response [60]. In an elegant study of in vitro carcinogenesis in W12 cells, a group from the UK showed that loss of episomal virus (E2) and elevation of integrated (E6, E7) HPV accompanying oncogenic transformation coincides with an increase in the expression of host antiviral genes [60]. Moreover, treatment with interferon beta accelerated the process of malignant transformation, with faster emergence of cells that lost episomes but contained integrated HPV-16 [61, 62]. In another in vitro model, the involvement of cell cycle and antiviral genes has been also noticed in the course of infection [63].

There were two missing pieces for the proposed model [58] to be fully valid: first, there was no clear demonstration of antiviral responses in tumors in vivo; and second, the potential trigger of antiviral response was not evident. In other words, why would a woman, who was unable to eliminate episomal virus for a long time (sometimes decades), acquires the ability to abolish HPV in the tumor?

Mine et al. has solved both problems [44]. First, a distinct meta-signature was observed that was validated in an independent dataset of almost 100 women consisting of several antiviral genes. This signature was characteristic of antiviral, but not of antibacterial immune activation with prominent dependence on both types of interferon. Even though there is support for the antiviral response in vivo, it was still unclear what causes this antiviral response.

**Key drivers of carcinogenesis**
A common notion that chromosomal aberrations provide competitive advantage for tumor cells led to the merging of transcriptional network with results of meta-analysis for genomic aberrations. Of the cervical cancer signature, ~9% (119 genes) were directly regulated by frequent gains or losses of chromosomes (i.e. gene expression corresponds to numbers of copies of a gene). Furthermore, causal inference analysis demonstrated that 36 key driver genes from frequent chromosomal gains could regulate the majority (~1000 genes) from the signature that were not located in regions of frequent aberrations. On average about half of the key drivers were present in each given tumor, most frequently within chromosomal gains at 3q, 1p, 1q, 20q [44].

Surprisingly, the most frequent chromosomal aberrations simultaneously harbored drivers for antiviral response and cell cycle [44]. For example, in the most frequent gain at 3q, six cell cycle drivers (NAT13, MCM2, TOPBP1, CEP70, GMPS, and RFC4) and one antiviral driver LAMP3 were found. These cell cycle genes are known to mostly participate in different stages of mitosis guiding chromatid cohesion (NAT13), spindle assembly (CEP70) and DNA replication (MCM2, RFC4, TOPBP1) potentially controlling these processes in cancer cells. For example, RFC4 knockdown demonstrated that it was essential for liver cancer cell proliferation and survival [64].

While several of the identified cell cycle drivers have been previously reported to perform this function in normal tissues and other cancers, this was the first report to show LAMP3 driving expression of antiviral genes such as STAT1, IRF7, HERC5, ISG20, OAS1. Interestingly, among these genes, in vitro overexpression of STAT1 has been associated with the inhibition of episomal and rise in integrated HPV genomes [65]. In agreement with the finding of potential antiviral capacity of LAMP3, other work has reported that high LAMP3 expression is correlated with ability of hepatitis C patients to respond to antiviral therapy [66]. Interestingly, the role of LAMP3 in cervical cancer is not limited to orchestration of HPV elimination. It was also shown that experimental overexpression of this gene leads to increased metastasis in an animal
model of cervical cancer [67]. Thus, this possible dual function of LAMP3 might be a part of the explanation of why it is so frequently amplified in cervical cancer. Besides LAMP3, additional antiviral gene drivers were found to be present in frequent gains on other chromosomes (ADAR, AIM2, RFX5 on 1q; IFI44L, IFI44, ISG15 on 1p; MMP9 on 20q; and TYROBP on 19q) [44]. Unlike LAMP3, most of these genes are well known to induce antiviral responses. For example, AIM2 has been shown to be protective against DNA viruses [68].

Although genomic aberrations would cause induction of cell cycle and antiviral pathways, there was no answer to the question of what downregulates the epithelial differentiation pathway. Because the meta-analysis of gene expression was missing an important group of regulatory molecules called miRNAs, the same causal inference analysis employed for protein coding genes to interrogate the role of miRNAs could not be used. Instead, a bioinformatics analysis that allows prediction of miRNA targets based on DNA sequences was used [44]. mir-15b and mir-16-2 located on the 3q gain were suggested to be the regulators of the epithelial differentiation pathway as they were predicted to potentially target two genes (NUAK2, SLURP1) within this subnetwork, which were highly connected with others in that group. In line with this hypothesis, expression of mir-15b and mir-16-2 as well as mir-9 (1q23.2), mir-205 (1q32.2) and mir-28-5p (3q27.3) has been found to be elevated in high grade CIN lesions or cervix cancers compared to normal epithelium [69-73] and a direct relationship between expression and chromosomal amplification has been demonstrated for mir-15b, mir-28-5p and mir-9 [72].

Thus, it appears that chromosomal aberrations, together with HPV integrants, regulate the key set of genes driving most of the functional pathways that were proposed as Hallmarks of Cancer [74]. The essential role of HPV in cervical carcinogenesis, however, gives somewhat a different perspective (as discussed next) on the immune
alterations in women that are on their way to developing a cervical tumor than the host–immune changes and tumor adaptations proposed in the Hallmarks of Cancer.

**Model of carcinogenesis and its implications**

The natural course HPV infection rarely ends with invasive carcinoma as the estimated lifetime risk of developing cervical cancer is around 1% (Figure 1) [75]. Thus, there are several events, some of them genetically and environmentally determined and some of them random, which are needed for development of cancer. In a model of carcinogenesis, the combination of HPV infection with a weak antiviral response at first results in chronic infection. Next, chronic infection with high-risk HPV leads to genomic instability resulting in an increased rate of chromosomal aberrations. Simultaneously, HPV integrates into the human genome, although it is still present in the episomal form with E2 keeping low expression of integrated E6/E7 oncogenes (Figure 3a). While genomic instability might affect random loci, only those cells that harbor advantageous for tumor growth aberrations would be further selected. Such chromosomal amplifications (e.g. 3q, 1p, 1q, 20q) contain drivers of cell cycle and interferon-related antiviral genes. On one hand, cell cycle drivers sustain continuous cell proliferation and growth. On the other hand, the antiviral drivers, overexpressed in the tumor, trigger an immune response helping the woman’s immune system which was unable to eliminate the virus alone. The reduction in episomal virus and consequently in E2 activity in tumors would release E6/E7 expression which would block cell cycle controlling proteins (p53 and retinoblastoma) synergizing with the direct effect of cell cycle drivers (Figure 3b). The enhanced cell cycle and block of apoptosis operating in cells expressing higher levels of E6/E7 might in turn contribute to the resistance to killing by the ongoing immune response. Remarkably, the selected chromosomal gains simultaneously contain drivers of both processes indicating that it might be cost-effective to select one aberration that affects both functions at once (i.e. killing two birds with one stone). In addition, another type of drivers (miRNAs) located in the
frequent gains may also contribute to cervical carcinogenesis by inhibiting epithelial differentiation (Figure 3b). The proposed model directly refers to the great majority of cervical carcinomas as they contain HPV integration [76, 77]. The more rare episome-associated carcinomas might have some similarities in the pathogenesis [78] but their exact molecular mechanisms remain to be elucidated.

This novel perspective on disease pathogenesis gives us an opportunity to re-evaluate some puzzling observations about cervical cancer genetics and patients’ response to antiviral treatment. Seemingly contradictory results were reported in two studies of genetic association between alleles of two immune genes and cervical cancer. First, the CD28(TT), IFNG(AA) single nucleotide polymorphism (SNP) genotype combination (both genes expressed by T lymphocytes) was associated with susceptibility to invasive cervical cancer in three patient cohorts [79]. Subsequently, opposite results (protection against disease) for the interaction of the same alleles were reported [80]. Notably, the probability of finding by chance the interaction effect of two loci associated with the same disease in two independent studies is infinitely close to zero. Hence, these opposite effects would look enigmatic if another difference, hidden from a first glance, would not exist between two studies. While first study analyzed retrospective cases of invasive cervical cancer, there were extremely few (3.6%) of those in the cohort analyzed by the second study as most of these cases were carcinoma in situ, frequently regarded as a pre-cancerous state [81]. Armed with a new model, we may reconcile the results of both studies. According to the model, different disease stages represent opposite poles of the disease in terms of antiviral immune response. At the first pre-cancer stage of disease the selection is directed to women who are poor responders to the virus and therefore develop chronic infection. To proceed to the second (cancer) stage, however, women who are more capable in eliminating the episomal virus would be preferentially selected as this would work in concert with genomic aberrations in antiviral genes. Therefore, now selection pressure is for good
antiviral responders and consequently tumors would be more likely to progress in women carrying gene variants for strong antiviral response.

The second topic has a major implication for developing adequate therapeutic strategy, especially considering strong recent enthusiasm for development not only preventive, but also therapeutic vaccines (reviewed in [58]). Interestingly, while preventive vaccines show reasonable efficacy in averting high grade lesions and invasive cancers (reviewed in [58]) the therapeutic approaches, including agents that stimulate antiviral immunity or directly inhibit virus, produced some controversial results [58, 82, 83]. Although we do not have a complete set of evidence, we hypothesize that this discrepancy can be explained by different stages of infection of patients undergoing antiviral therapy. Indeed, women whose infection is prevented (by vaccine) or treated before HPV has integrated into human genome would benefit from virus elimination. However, patients that already have signs of HPV genomic integration will lose episomic virus and consequently the inhibitory effect of E2 on expression of E6 and E7. This might boost the malignization of lesions rather than tumor repression. Therefore, we believe that prospective clinical trials should account for the stage of HPV infection by monitoring its integration into human genome and expression of E2/E6/E7.

Concluding remarks and future perspectives (see also Box 1)

Thus, as has been proposed more than a century ago for tumors in general [84], chromosomal aberrations play a major role in cervical carcinogenesis. Furthermore, recent studies provided data supporting two key ideas: first- that during disease progression there is a dramatic change in relation between the a woman and HPV from insufficient antiviral immunity resulting in chronic infection to enhancement of antiviral immunity in the tumor driven by genes located in chromosomal amplifications; and second, that these chromosomal amplifications simultaneously harbor genes that drive antiviral response, cell cycle, and inhibit cell differentiation.
A novel understanding of the disease evokes testing for diagnostic and therapeutic applications (Box 1). In addition to currently used viral status tests, the simultaneous monitoring of precancerous lesions for genomic and transcriptomic aberrations of the key driver genes should be further explored as a diagnostic tool for selection of women who need immediate therapeutic intervention versus those ones whose lesions either regress or do not progress further. Furthermore, by knowing which tumor harbors which specific group of driver genes we can start personalizing therapy by targeting drugs such as siRNA to specific alterations observed in each given patient.

Although this review is devoted to “ménage a trois” there might be a fourth player – vaginal microbiota whose role have not been explored yet in cervical cancer. There is an explosion of studies demonstrating that the gut microbiota play essential roles in host’s ability to deal with viruses [85], to develop intestinal cancers [86], and to respond to chemotherapy [87]. In the field of cervical cancer, it has been recently shown that HPV status can influence microbial diversity of vaginal microbiota [88]. However, whether alterations in normal vaginal microbiota or in opportunistic pathobionts have any causal role in the development of chronic HPV infection or cervical cancer have not been addressed.

Finally, besides cervical carcinoma there are several other cancers such as oropharyngeal, anal and penile cancers that are caused by or at least are associated with HPV infection [89]. Although gene drivers for those cancers are not well elucidated, 3q gain is one of the most frequent aberrations in oropharyngeal [90] and lung [91] cancers. Therefore, it would be worthwhile to explore whether molecular mechanisms of cervical carcinogenesis discussed in this review can be extended to other HPV-associated malignancies.
Figure Legends

**Figure 1:** Progression of HPV cervical infection to cancer and major changes in the HPV physical state (pure episomal, episomal+integrated (mixed), or pure integrated), expression of viral genes E2, E6, E7 and host chromosomal aberrations. Percentages denote proportions of patients. Numbers have been summarized from several studies: for disease stages progression [92]; for viral state [29-36]; for E2, E6/E7 expression [93-97]; for aberrations [41, 42, 44, 98].

**Figure 2:** Frequency of gain (red) or loss (blue) in the genome detected in the meta-analysis of comparative genomic hybridization studies using cervical cancer samples. Key gene drivers (orange, antiviral; black, cell cycle) are indicated in corresponding chromosomal locations. Data used from [44].

**Figure 3:** A model of cervical carcinogenesis. a) Persistent high risk HPV infection may result in the integration of virus into host genome upon which E2 is disrupted. The integration leads to the increased genomic instability, however, the expression of E6/E7 oncogenes is still controlled by episomal E2. b) frequent chromosomal aberrations (gains) occur in the regions containing antiviral genes, which will induce the elimination of inhibitory episomal E2, release of E6/E7 that will block suppressors of cell cycle (p53, retinoblastoma, Rb). The same chromosomal gains contain drivers of cell cycle that directly induce cell proliferation, and miRNAs that may inhibit cell differentiation. All three processes act synergistically allowing the dysplastic cell to become a malignant tumor.
References:

10 Danos, O., et al. (1982) Human papillomavirus 1a complete DNA sequence: a novel type of genome organization among papovaviridae. The EMBO journal 1, 231-236
cervix uteri. Clinical cancer research : an official journal of the American Association for Cancer Research 10, 3059-3063
28 Hwang, E.S., et al. (1993) Inhibition of cervical carcinoma cell line proliferation by the introduction of a bovine papillomavirus regulatory gene. Journal of virology 67, 3720-3729
32 Hudelist, G., et al. (2004) Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. Gynecologic oncology 92, 873-880


48 Kitahara, O., et al. (2002) Classification of sensitivity or resistance of cervical cancers to irradiating radiation according to expression profiles of 62 genes selected by cDNA microarray analysis. Neoplasia 4, 295-303


65 Hong, S., et al. (2011) Suppression of STAT-1 expression by human papillomaviruses is necessary for differentiation-dependent genome amplification and plasmid maintenance. Journal of virology 85, 9486-9494
73 Xie, H., et al. (2012) miR-205 expression promotes cell proliferation and migration of human cervical cancer cells. PloS one 7, e46990
83 Roy, S., et al. (2011) Defective dendritic cell generation from monocytes is a potential reason for poor therapeutic efficacy of interferon alpha2b (IFNalpha2b) in cervical cancer. Translational research : the journal of laboratory and clinical medicine 158, 200-213
93 Hafner, N., et al. (2008) Integration of the HPV16 genome does not invariably result in high levels of viral oncogene transcripts. Oncogene 27, 1610-1617
97 Xue, Y., et al. (2010) HPV16 E2 is an immediate early marker of viral infection, preceding E7 expression in precursor structures of cervical carcinoma. Cancer research 70, 5316-5325
Box 1. Outstanding questions

- As multiple chromosomal aberrations are usually present in each tumor, is this a result of mixture of different clones or is there a need for several aberrations within one cell to support malignization?
- What is the role of vaginal microbiota in developing of persistent HPV infection and progression of precancerous lesions into invasive cancer?
- What are the host and viral molecular markers for identification of patients that would benefit from antiviral treatment and therapeutic vaccines?
- Can the knowledge of key drivers of cervical carcinogenesis be implemented into early diagnosis and personalization of treatment of invasive cancers?
- To what extend the novel molecular mechanisms of cervical carcinogenesis are applicable to other HPV associated malignancies?
Figure 1

- **Episomal virus**: 100% for CIN 1, 100% for CIN 2/3, 20-95% for invasive cancer, 0-20% for 5-40%
- **Mixed (episomal+ integrated virus)**: 0% for CIN 1, 0-30% for CIN 2/3, 0-75% for invasive cancer, 5-40% for 0-50%
- **Integrated virus**: 0% for CIN 1, 0% for CIN 2/3, 0-50% for invasive cancer, 45-80% for 0%
- **E2 expression**: Variation indicated by bars
- **E6/E7 expression**: Variation indicated by bars
- **Chromosomal aberrations**: ~0 for CIN 1, 3q gain ~2% for CIN 2/3, 3q gain 30%, 1p gain 30%, 1q gain 30%, 20q gain 20%, 2q loss 20%, 4p loss 20% for invasive cancer, 3q gain 60%, 1p gain 25%, 1q gain 40%, 20q gain 30%, 2q loss 35%, 4p loss 35%
Integration of E6/E7 Persistent HPV infection

Expression of E6/E7 is low

Genomic instability

Increased expression of E6/E7

Selection of chromosomal gains

Antiviral response

loss of viral episomes

Direct stimulation of cell cycle

Epithelial differentiation block

Uncontrolled growth

Invasive cancer

Cell cycle suppressors