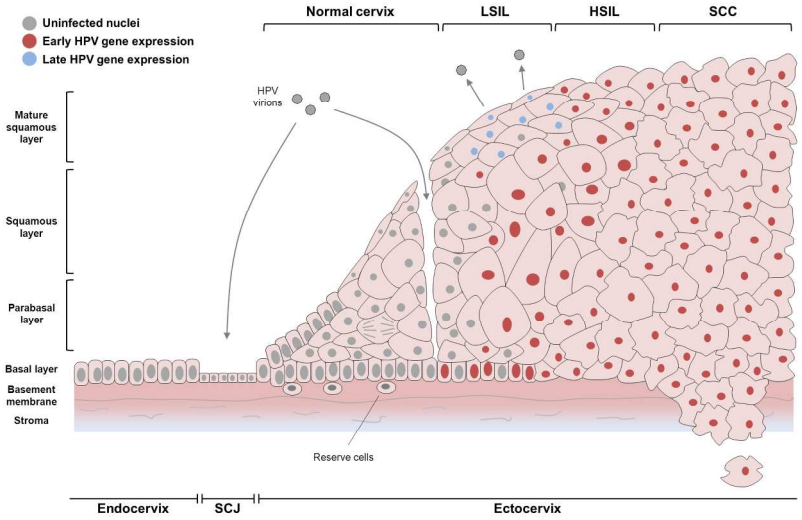
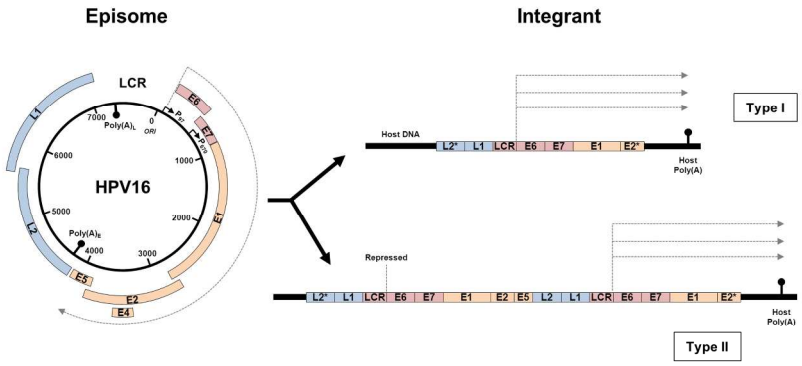
THE JOURNAL OF Pathology
The Journal of The Pathological Society
Understanding Disease

Pathogenesis of human papillomavirus-associated mucosal disease.

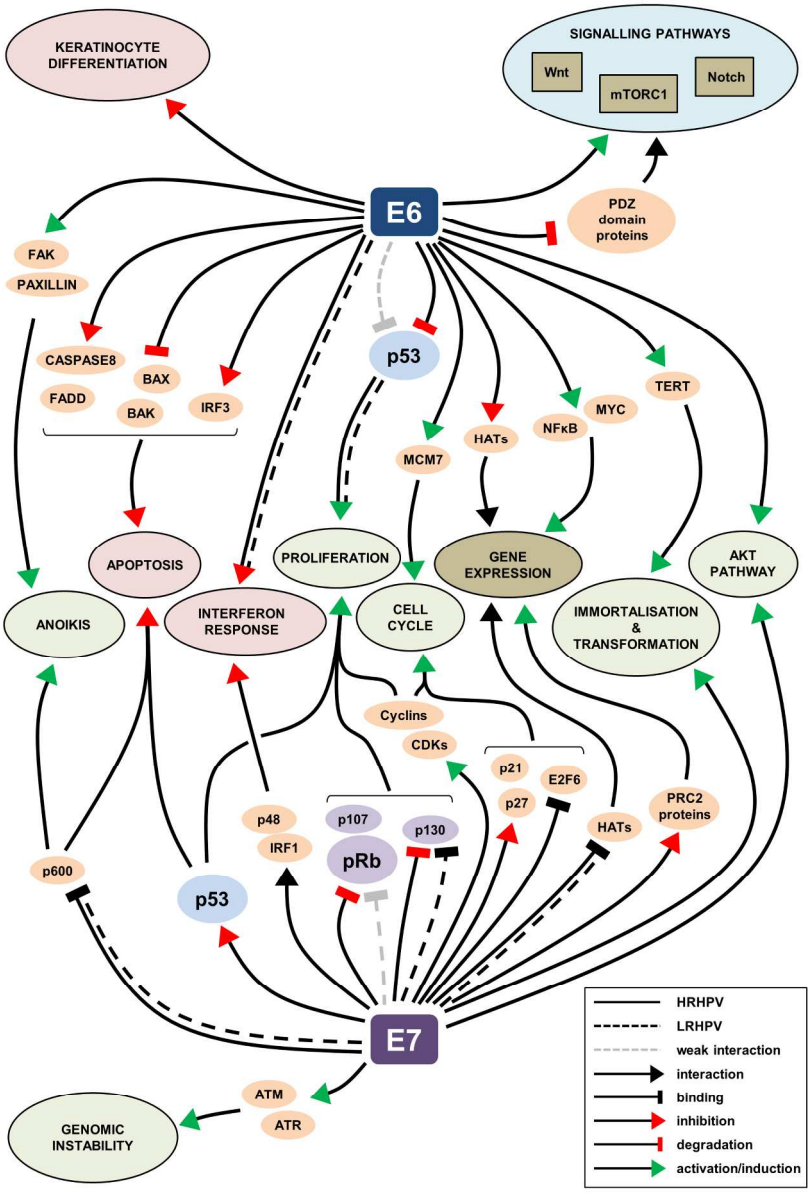
Journal:	*The Journal of Pathology*
Manuscript ID:	Draft
Wiley - Manuscript type:	Invited Review
Date Submitted by the Author:	n/a
Complete List of Authors:	Groves, Ian; University of Cambridge, Department of Pathology Coleman, Nicholas; University of Cambridge, Department of Pathology;
Key Words :	Human papillomavirus, Mucosa, Oncogene, E6/E7, Epigenetics, Integration, Squamous cell carcinoma
File Designations:	
Tissue:	
Pathology:	
Technique:	



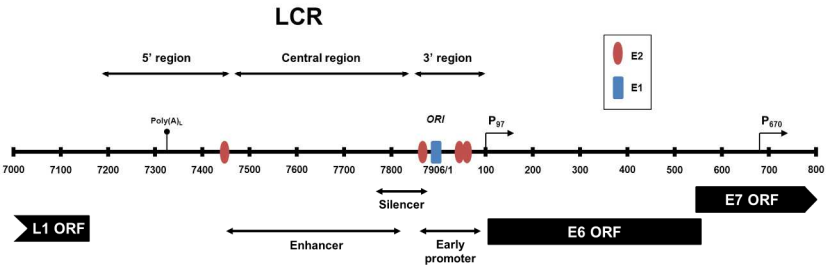
254x338mm (300 x 300 DPI)



254x338mm (300 x 300 DPI)



254x338mm (300 x 300 DPI)



254x338mm (300 x 300 DPI)

Pathogenesis of human papillomavirus-associated mucosal disease.

Ian J Groves & Nicholas Coleman*

University of Cambridge, Department of Pathology, UK

Correspondence: *Nicholas Coleman, University of Cambridge, Department of Pathology, Tennis Court Road, Cambridge, CB2 1QP, UK. e-mail: nc109@cam.ac.uk.

Conflicts of Interest Statement: There are no potential conflicts of interest.

Running title: Mucosal HPV pathogenesis.

Word count: 3925

Abstract

Human papillomaviruses (HPVs) are a necessary cause of carcinoma of the cervix and other mucosal epithelia. Key events in high-risk HPV (HRHPV)-associated neoplastic progression include persistent infection, deregulated expression of virus early genes in basal epithelial cells and genomic instability causing secondary host genomic imbalances. There are multiple mechanisms by which deregulated virus early gene expression may be achieved. Integration of virus DNA into host chromosomes is observed in the majority of cervical squamous cell carcinomas (SCC), although in ~15% of cases the virus remains extra-chromosomal (episomal). Interestingly, not all integration events provide a growth advantage to basal cervical epithelial cells, nor lead to increased levels of the virus oncogenes E6 and E7, when compared with episome-containing basal cells. The factors that provide a competitive advantage to some integrants, but not others, are complex and include virus and host contributions. Gene expression from integrated and episomal HRHPV is regulated through host epigenetic mechanisms affecting the virus long control region (LCR), which appear to be of functional importance. New approaches to treating HRHPV-associated mucosal neoplasia include knockout of integrated HRHPV DNA, depletion of virus transcripts and inhibition of virus early gene transcription through targeting or use of epigenetic modifiers.

Keywords: Human papillomavirus; mucosa; oncogene; E6/E7; epigenetics; integration; squamous cell carcinoma

Human papillomavirus infection and global disease

Human papillomavirus (HPV) infection poses a significant risk to morbidity and mortality worldwide, being associated with ~4.8% of all human cancers [1]. Papillomaviruses are a family of small, non-enveloped viruses with a double-stranded DNA (dsDNA) genome of, in the case of human papillomavirus (HPV), approximately 7.9 kilobases (kb) [2]. To date, over 180 individual types of HPV have been sequenced [3, 4] and all infect epithelial cells, usually with a preference for either cutaneous or mucosal surfaces. The proportion that infects the genital tract (30-40 types from the alpha genus) can be sub-divided into low and high-risk types, based on their oncogenic potential. Low-risk HPV types, including HPV6 and 11, are associated more with benign ano-genital warts or condylomata, whereas at least twelve high-risk HPV (HRHPV) types, HPV16, 18, 31, 33, 35, 39, 45, 51 52, 56, 58 and 59, are associated with ano-genital cancers and precursor neoplastic lesions [5, 6].

Although HPV infection is nearly ubiquitous, the virus does not cause cancer in the large majority of cases [7]. Most infections are inapparent and cleared by the host immune system within 18 months [8]. However, ~10-15% of women do not clear HPV infection, with persistence of high-risk HPV being the major risk factor for development of ano-genital cancers [9, 10].

The life cycle of HPV

In stratified epithelia HPV infects cells in the basal layer, most likely via epithelial wounding or micro-fissures [11], through an entry mechanism that is thought to require active cell division [12, 13]. The cervical transformation zone at the squamo-columnar junction may be susceptible to malignancy due to the heightened accessibility of epithelial reserve cells or stem cells in this region [14, 15]. The ability of HPV to target basal stem cells is also likely to

provide one mechanism by which persistent infection is established in some individuals (Figure 1) [16].

Initial infection of basal cells is usually associated with low level amplification of the HPV episome, to a copy number of ~100 per cell. The concerted expression of virus early genes E1 (virus DNA helicase) and E2 (virus transcription factor and genome tether) allow replication of virus DNA from the origin (Figure 2) [17, 18]. Virus copy number is then maintained within the basal proliferating compartment of the epithelium, with the E2 protein, in conjunction with various host binding partners, directing partitioning of virus genomes via tethering to the host chromosomes [19-21]. This maintenance phase of the virus life cycle also involves expression of the oncoproteins E6 and E7 from the virus early promoter (p97 in the major HRHPV, HPV16) [22, 23]. HPV E6 and E7 interact with various host proteins and carry out many modulatory functions within the infected cell (Figure 3) [24]. The constrained level to which E6 and E7 proteins are expressed in basal cells is likely to aid immune evasion and ultimately persistence of infection in the host [16].

During the HPV life cycle, the programme of virus gene expression is dependent on the differentiation profile of the infected cells (Figure 1). HPV proteins are able to drive differentiating suprabasal squamous epithelial cells back into the cell cycle, in order to reactivate DNA synthesis and thereby maintain an elevated virus genome copy number [25, 26]. Virus RNA splicing shifts from early to late polyadenylation sites, controlled by E2 and the availability of splicing factors, which are variably expressed according to the differentiation state of the infected cell [27, 28]. Similar mechanisms are responsible for the production of late gene products L1 and L2, the virus capsid proteins, in the superficial cell layers [29]. The productive life cycle is completed when virions self-assemble, package their

replicated genomes and are released non-lytically from the cell into the immediate external environment [26].

Cervical cancer: the paradigm of HPV-associated oncogenesis

HPV has been determined as a definite risk factor for numerous human cancers of mucosal surfaces, including penile, vulval, vaginal, anal, oropharyngeal and cervical carcinomas [1, 2, 6]. Cervical cancer is the fourth most common malignancy in women worldwide, with ~528K cases (~12% of all female cancers) and ~266K deaths (~50% of cases) per annum [6].

The majority of cervical cancers represent squamous cell carcinomas (SCCs), although adenocarcinomas and adeno-squamous carcinomas are also seen. SCCs arise from precursor lesions that may be classified using the three-tier cervical intraepithelial neoplasia (CIN) or two-tier squamous intraepithelial lesion (SIL) systems [30]. Low-grade SILs (LSILs) broadly correspond to CIN1 and generally represent non-neoplastic productive HPV infections that have a low risk of progression to malignancy [31]. In contrast, high-grade SILs (HSILs), broadly corresponding to CIN2/3, comprise abortive virus infections in which there is deregulated expression of HPV early genes in basal epithelial cells and a greater risk of progression to invasive disease (Figure 1) [32]. The cervical squamous cell carcinomas (SCCs) that arise from these precursor SILs are usually clonal, due to the emergence of cells with the greatest competitive growth advantage [33, 34].

Deregulation of HPV oncogene expression in neoplastic progression

An important event in HPV-associated neoplastic progression is deregulation of normal patterns of virus gene expression. Increased expression of E6 and E7 in basal epithelium leads to pro-malignant effects in the proliferating cell compartment, resulting in increased

cell cycle entry [35] and loss of differentiation across the epithelium [36]. There appear to be multiple causes of deregulated HPV gene expression, occurring at both genetic and epigenetic levels. The most common event in cervical SCC is integration of the virus genome into host chromosomes (Figure 2) [36]. However, a proportion of cervical SCCs are associated with retention of episomes, in which the E2 open reading frame is maintained and expressed throughout progression to malignancy [37].

HRHPV integration

In cervical SCC tissue samples, ~50-80% of HPV16-positive and almost all HPV18-positive cases are associated with integration of virus genomes [38-42]. Interestingly, different HPV types appear to integrate at different frequencies. In one study of 835 cervical samples, integration was detected in 55% of HPV16-positive cases, 92% of HPV18-positive cases, 14% of HPV31-positive cases, 37% of HPV33-positive cases and 83% of HPV45-positive cases [41]. The frequency of HPV integration in SILs has been widely debated. It has been proposed that integration can either be an early event associated with LSIL to HSIL progression [43-46] or a later event that accompanies progression from HSIL to SCC [47, 48].

Such contrasting perspectives can at least partly be explained by technical differences between the methods used to identify integrated HPV. Protocols designed to detect virus-host fusion transcripts, such as RNA *in situ* hybridisation (ISH) [49] or 3' RACE-PCR (termed amplification of papillomavirus oncogene transcripts (APOT)) [50] will only identify transcriptionally active integrants. However, methods that target the virus DNA, including DNA ISH [47, 51, 52], Southern blotting [38, 53], restriction-site PCR [54] and quantitative PCR [44, 55], will detect all integrated virus genomes, regardless of their transcriptional

activity [56]. Previous *in vitro* studies have suggested that HPV integration events occur in cells that initially retain non-integrated episomes. Expression of the virus E2 transcriptional regulator from such episomes is able to repress integrant-derived transcription [57, 58]. Episome clearance (for example, through a host anti-virus response) [59] is an important step in overcoming this transcriptional inhibition and represents a key additional step in cervical neoplastic progression following initial virus integration. Accordingly, identification of integrated virus DNA does not necessarily indicate transcriptionally active, and therefore selected, integration events, for which RNA-based detection methods are required.

In cervical SCCs, HRHPV may integrate as a single copy or as multiple copies. In the former (so called type-I integrants), complete loss or truncation of the virus early 3' region leads to deletion or disruption of integrant-derived E2, together with loss of the virus early gene polyadenylation site [36]. Thus, the transcribing polymerase adopts the nearest host polyadenylation site, producing virus-host fusion transcripts thought to be more stable than those of the virus alone. In the cells that undergo selection during cervical carcinogenesis, these changes lead to increased expression and stability of transcripts encoding HRHPV E6 and E7 [34, 60]. The second type of integrant (type-II integrants) involves concatemers of full length HPV (including the E2 gene), often with interspersed host sequences. In cells containing multiple integrated HPV copies, there appears to be selection of cells containing only a few transcriptionally active sites [49]. In concatemerised virus DNA, epigenetic silencing of full length copies, for example through DNA methylation, would prevent E2 expression [61], with transcription being limited to the virus-host junction sequences where the E2 gene is disrupted or deleted. Such epigenetic regulation could also restrict E6 and E7 transcription, to prevent deleterious genomic instability caused by high-level gene expression [37, 49].

Until recently, it has been unclear whether all integration events, when de-repressed following episome loss, lead to increased levels of virus oncogenes and/or a cell growth advantage. Selection of integrated HRHPV occurs relatively early in cervical carcinogenesis and determinants of selection have been difficult to investigate adequately using clinical samples [36]. However, the W12 model of HPV16-associated cervical carcinogenesis has enabled the generation of a useful panel of cell clones that were derived from an identical background under non-competitive conditions and differed only by the genomic site of HPV16 integration [62]. Interestingly, when compared with the episome-containing cells from which they were generated, only ~50% of these clones showed significantly greater growth rates and only ~50% showed significantly greater expression of E6/E7 [63], indicating that HRHPV integration *per se* does not necessarily lead to increased oncogene expression or a cell growth advantage. It remains unclear what provides some virus integrants with a selective advantage compared with others, although the epigenetic environment at the integrated virus LCR, and a contribution from host genes appear to be involved [63]. Interestingly, the site of virus integration in the host genome may also affect cell responsiveness to steroid hormones [64-67], through as yet unidentified mechanisms.

A large majority of HRHPV integration sites in cervical SCC have been mapped at, or near to, chromosome fragile sites (CFS), with around 50% of selected HPV16 and HPV18 integration sites found in the same chromosomal band as a CFS [62, 68]. The E6/E7 proteins are likely to play a role in the process of integration [69], although major genomic instability is typically seen following the emergence of intergant-containing cells [70, 71]. While HRHPV integration sites are widely distributed across the human genome, it appears that integration is more likely to occur at certain genomic sites (so-called integration hotspots),

such as 3q28, 4q13.3, 8q24.21, 13q22.1 and 17q21 [54, 62, 72, 73]. Some integration sites share sequence homology with the HPV16 genome, most commonly with the E5 and L2 genes, albeit only around 50 nucleotides maximally [73]. HPV integration sites also have a propensity for occurring near to clusters of host microRNAs (miRNAs). In one study, over two thirds of integrants mapped within 3MB of a miRNA-coding locus, potentially leading to deregulation of miRNA expression [73]. It has been proposed that the mechanism by which HPV genomes are tethered onto host chromosomes, involving interaction of the E2:Brd4 complex with acetylated histones that are found most commonly at active sites of transcription or CFSs, allows HPV genomes to integrate more often at CFS when genomic instability increases [74].

As well as the question of whether integration of HPV genomes at particular host sites can lead to specific changes in virus gene expression, there is contention as to whether integration can lead to insertional mutagenesis of host genes, possibly leading to greater selectability of the affected cells. Early work assessing HPV integration sites in cervical cancers found evidence for insertion at sites of both host tumour suppressor genes and host oncogenes [75-77]. There have been reports of repeated integration, for example near to the *MYC* gene at 8q24 [77-79]. Recent genome-wide studies of HPV-positive cell lines and tumour material that have led to a much greater insight into the extent of amplification of virus sequences and local rearrangement of host DNA at sites of integration [80-82]. Significant changes in expression of genes at HPV integration sites were suggested from genome wide cross-sectional studies of cervical carcinomas [83, 84]. Despite other evidence supporting potential HPV insertional mutagenesis in ano-genital neoplasia [56, 85-87], it should be noted that there are few functional data that support this possibility, and thus no causal association with cervical carcinogenesis has been verified. Indeed, a recent study of oropharyngeal SCC

(OSCC) found that integration of HPV16 genomes had no significant effect on the expression of virus-disrupted host genes [88].

Epigenetic modification of the HPV LCR: DNA methylation

Deregulation of HPV oncogene expression under conditions of episome maintenance may occur via diverse mechanisms [37]. There are reports of mutations to the LCR of episomes, which modify the interaction of several transcription factors and result in altered virus transcript levels [89-92]. However, the relationship of these changes with clinical disease is still poorly defined. Of greater importance may be epigenetic changes affecting the virus LCR (Figure 4).

One mechanism of abrogating binding of the virus regulatory protein E2 to the episomal or integrated HPV genome is methylation of CpG dinucleotides within the E2 binding sites (E2BSs) of the LCR [93]. However, a clear association between DNA methylation of the LCR and disease progression is still not resolved. While it was initially suggested that DNA methylation at the LCR decreased with progression from normal cervix to SILs and SCC following HPV16 infection [94-97], other studies have shown that increases in LCR DNA methylation can occur [98-100]. There are further discrepancies across studies that focussed on the E2BSs in the LCR, when considering both disease progression and cell differentiation. In cells showing episome maintenance, E2BSs of genomes in a basal-like monolayer culture showed high levels of methylation that were lost during cell differentiation [101, 102]. However one study, while also showing such differentiation-related reductions at the promoter-distal E2BS, showed an increase in DNA methylation at the three promoter-proximal E2BSs (Figure 4) [103], changes that may facilitate episome replication in

differentiated cells. For integrated virus genomes there is little DNA methylation at the LCR unless multiple copies have integrated concatemerically [61].

Epigenetic modification of the HPV LCR: chromatin modification

Gene expression from virus genomes, both episomal and integrated, is also associated with post-translational modifications to histone tails of nucleosomes at the HPV LCR, and with subsequent transcription factor occupancy. The association of nucleosomes with the HPV genome, whether episomal [104] or integrated [105], has been known for nearly forty years. Functional association of histone proteins with the HPV LCR was shown to occur at the enhancer and early promoter (Figure 4), depending on specific DNA sequences, such that nucleosome occupancy appeared distinct for different HPV types [106]. Further evidence from the use of histone deacetylase (HDAC) inhibitors, which caused increased levels of early transcripts, pointed towards modification of episome-associated histones and nucleosome remodelling as a method of virus transcriptional regulation [107]. This early finding has been verified by work showing the necessary involvement of some histone acetyltransferases (HATs), including p300, in efficient activation of HPV gene expression [108, 109] and the presence of acetylated histones H3 and H4 (H3ac and H4ac, respectively) at the LCR in transcriptionally active episomal [37, 110] and integrated HPV genomes [63]. Consistent with these observations, the recruitment of HDACs by various host proteins is able to repress HPV transcription [111, 112].

The level of association of histone post-translational modifications at the LCR in cervical squamous cells has been shown to change as cells differentiate and as disease progresses. Upon cell differentiation, activation of virus late gene expression is associated with stimulation of the late promoter (Figure 4), coinciding with increases in levels of the active

marks H3ac and dimethylated lysine 4 of H3 (H3K4me2) and with increased transcription factor binding at the enhancer and promoter [110, 113]. Throughout the course of *in vitro* neoplastic progression associated with episomal HPV16, H3ac and H4ac were present at the virus enhancer and at both early and late promoters [37]. Acetylation of histones initially increased across the HPV16 genome as the cells progressed phenotypically to SCC, subsiding in association with decreasing oncogene expression in late-stage cells (a change capable of limiting excessive genomic instability) [37]. In cells containing integrated HPV16, high levels of virus transcription per DNA template were associated with greater abundance of marks of active transcription, H3ac and H3K4me3 across the LCR and early gene open reading frames [63], together with reduced association of H3K9me2 and H3K27me2. The latter are marks of transcriptional repression that have also been observed at the numerous heterochromatinised genomes in CaSki cervical SCC cells [114].

Together, these data indicate that levels of HPV16 oncogene transcription in cervical squamous cells are directly associated with chromatin structure at the HPV LCR and promoters. Such findings are consistent with published ChIP-seq data of histone modifications observed in populations of HPV18 integrant containing HeLa cervical adenocarcinoma cells [115]. Strikingly, assessment of transcription factor association across the HPV18 genome in HeLa showed binding of all interacting proteins at the LCR. Despite the length of the LCR and the presence of known binding motifs, it is unlikely that all transcription factors can bind this region at the same time; hence the associations would be dynamic [115]. Furthermore, it is possible that transcription factor binding might not occur in a regulated, sequence specific manner, at least in HeLa cells. This is supported by the finding that certain specific locations in the human genome, known as high-occupancy target (HOT) regions and first found in *C. elegans*, are able to bind numerous unrelated transcription

factors [116]. Importantly, such regions have been found in HeLa cells [117], although the relevance for the integrated virus genome has not yet been demonstrated. It has been proposed that a majority of transcription factors could be recruited to the HOT regions by a fewer number of common co-factors, thus allowing the possibility of some sequence-specific transcription factor binding. It will therefore be interesting to establish how numerous transcription factors are able to bind a relatively short genomic regulatory sequence like the virus LCR and whether there is any biological relevance to these associations, for example during differentiation of HPV-infected cells.

Genetic and epigenetic modification of host genes

Chromatin modification and DNA methylation. In addition to the above host effects on HPV transcription, HRHPV oncoproteins are able to modify host gene transcription through epigenetic mechanisms. The E7 protein interacts with and inhibits polycomb group proteins, which act in concert to repress gene expression [118, 119]. In conjunction with E7-mediated up-regulation of lysine demethylases (KDM) 6A and 6B, which remove the facultative heterochromatin mark H3K27me3, inactivation of the polycomb repressive complex 2 (PRC2) histone methyltransferase EZH2 and down-regulation of PRC1 protein BMI1 by HRHPV proteins are able to cause reactivation of transcription from numerous normally repressed genes, including members of the HOX group and the cyclin-dependent kinase inhibitor p16INK4A [120, 121]. Interestingly, in SCCs HPV16 has also been shown to decrease the level of RBBP4 (RbAp48) [122], a nucleosome-binding recruitment protein for PRC2 and HDAC containing complexes including NuRD, Sin3 and CoREST, thereby further indirectly modulating global host gene expression. Additionally, E7 is able to bind and sequester HDACs allowing both transcriptional activation of target genes, such as that mediated by the normally HDAC-repressed transcription factor HIF1α [123], and repression

of immune genes such as *TLR9* via retargeting of deacetylase (HDAC1) and demethylase (JARID1B) enzymes [124]. Indeed, E7 was first shown to gain HDAC activity through interaction with Mi-2 (CHD4), a member of the nucleosome remodelling deacetylase (NuRD) complex, thus promoting growth of HPV infected cells [125]. Furthermore, E7 can bind chromatin remodelling enzyme BRG1 (SMARCA4) and, alongside gained HDAC activity, repress host gene promoters including that of *C-FOS*, in contrast to the activating effect of BRG1 at the integrated HPV18 LCR [126-128]. The HPV oncoproteins, especially E7, therefore balance a global cellular inhibition of chromatin modification complexes by retargeting enzymatic activity toward specific host genes, probably including the virus genome, to initiate and maintain infection of keratinocytes.

Of possible clinical significance has been the accumulation of data showing changes in host gene DNA methylation throughout the progression of HPV-associated disease. High-risk HPV types have been associated with up-regulation at the protein level of DNA methyltransferases (DNMTs) 1, 3A and 3B [129-132]. The overall effect, whether direct or indirect, is an increase in DNA methylation of the host genome, often occurring at tumour suppressor genes and inhibiting their transcription. Studies have led to numerous potential biomarkers of disease progression that may in due course prove clinically useful [133, 134].

Host DNA mutations. In HPV-associated SCCs, deregulation of virus gene expression is typically associated with abnormalities of host genes, usually in a manner that activates oncogenes or inactivates tumour suppressor genes. Recent genome-wide next generation sequencing studies have allowed in-depth analysis of the frequencies and associations of somatic mutations in SCCs. Since HRHPV functionally inactivates p53 and pRB (Figure 3), relatively few examples of somatic mutations of the *TP53* and *RB1* genes have been found in

cervical SCCs (5% and 3%, respectively) [135]. Many of the most frequently mutated genes are associated with signalling pathways; recurrent mutations have been found in SCCs of the cervix and head and neck at the *PIK3CA*, *MAPK1* and *EGFR* genes, with strong relationships between the NOTCH signalling pathway and head and neck SCC [84, 136-138]. Other relatively commonly mutated genes in SCCs include *CDKN2A*, *PTEN*, *HRAS* and *FBXW7* [84, 136, 137]. Interestingly, a small number of mutations have now been found in genes encoding proteins associated with chromatin modification and remodelling, including repressive (*EZH2*) and activating (*MLL2, p300*) enzymes [84, 136]. Intriguingly, the *ERBB2* gene, whose encoded protein binds ligand-activated EGFR and enhances kinase-mediated activation of downstream signalling pathways including MAPK and PI3CA, is not just mutated in cervical SCCs but is also a reported site of HPV integration, with concomitant up-regulation in gene expression [83, 84, 135].

MicroRNAs

Post-transcriptional control of host gene expression in cervical cancer is an area of current interest. Cervical carcinoma is associated with changes to host microRNAs (miRNA) profiles [134, 139] and there are data supporting direct HPV effects on miRNA levels. Host miRNAs appear to be modulated by the major HRHPV oncoproteins E6 and E7 [140] and also by E5 [141]. Together, such changes affect cell pathways associated with apoptosis, differentiation and proliferation [140, 142] and in turn may regulate differentiation-dependent virus gene expression [140]. Altered expression of host miRNAs appears to occur at an early stage in HPV-associated cancer development and the changes may be prove to be diagnostically useful [134]. Modulation may be due to chromosomal copy number alterations, either directly at the miRNA locus [143] or indirectly via mechanisms such as upregulation of Drosha expression following 5p gain [139]. Further indirect down-regulation of some miRNAs might

also be due to global changes in DNA methylation patterns at promoter regions in cervical SCCs [134], while integration of the HPV genome at miRNA loci may also be able to modify expression from such loci [73]. Interestingly, a small number of studies have reported the possible expression of HPV-encoded miRNAs that could target virus and host transcripts [144, 145]. However, these findings do not appear to apply across all papillomavirus types [146] and require further biological verification. Host and/or virus gene expression may also be controlled by other non-coding RNAs, although these have yet to be identified.

Conclusions

It is becoming increasingly clear from *in vivo* and *in vitro* data that there are multiple routes by which neoplasia develops at mucosal surfaces following HRHPV infection. Events of fundamental importance in these processes include virus persistence, deregulation of virus early gene expression and host genomic instability. Although there are prospects for improved control of HRHPV-related disease through vaccination, worldwide vaccine coverage is still very low and current vaccines target a restricted range of virus types. There are still important opportunities for identifying new strategies for treating HRHPV-associated disease. Multiple approaches are under consideration, including knockout of integrated virus sequences through gene editing, or depletion of virus transcripts using small interfering RNA (siRNA) [147, 148]. Increased understanding of epigenetic regulation of gene expression from episomal and integrated HRHPV DNA, and of host gene expression through effects of HRHPV early proteins, may allow the development of epigenetic therapies that will permit selective inhibition of virus transcription without deleterious host effects. The availability of accurate *in vitro* models for detailed functional characterisation of the cell selection processes that occur during HRHPV-associated neoplastic progression will be of value in identifying rational therapeutic strategies for cervical and other carcinomas.

Acknowledgements

We thank Cancer Research UK and the Medical Research Council for funding our research, some of which is described in this review. We apologise to colleagues whose work we were unable to cite due to space constraints.

Author contributions

IJG and NC conceived the article and wrote the manuscript. Both authors had final approval of the submitted and published versions.

Figure Legends

Figure 1. ***HPV-associated neoplastic progression in cervical epithelium.***

HPV may infect basal squamous epithelial cells following micro-trauma or may target cells at or near the squamo-columnar junction (SCJ), including epithelial reserve cells. In the normal virus life cycle, HPV early gene expression is initiated (red nuclei) while the HPV genome is maintained at low copy number as an episome. As infected basal cells replicate and daughter cells move into the parabasal layer, expression of virus oncoproteins E6 and E7 allow cells that would normally differentiate to re-enter the cell cycle and produce an expanded epithelium. Migration of cells to the upper layers causes increased replication of the virus genome to high copy number, with expression of the virus E4 gene and often the structural proteins L1 and L2. These events allow encapsidation of the episomes into infectious virions which are then shed from/with the cornified surface. Such completion of the virus life cycle is supported in a low-grade squamous intraepithelial lesion (LSIL) but not in a high-grade SIL (HSIL), where disease progression is associated with deregulation of virus early gene expression and loss of late gene expression. Squamous cell carcinoma arises when cells gain the ability to penetrate the epithelial basement membrane and invade the underlying stroma.

Figure 2. ***Physical states of the HPV16 genome.***

The left hand panel shows the genomic organisation of the major HRHPV, HPV16, demonstrating the early (E) region, the late (L) region and the long control region (LCR). The integrated forms of HPV observed in cervical neoplasia (i.e. following a selection process) may represent fragments of the virus genome, with retention of E6/E7 and deletion or disruption of E2 (type-I integrants) or may represent concatamerised full length copies (type-

II integrants), in which transcriptional activity appears to be restricted to downstream sequences. (ORI = origin of replication; * = disrupted open reading frame.)

Figure 3. ***Important functions of high- and low-risk HPV E6 and E7 proteins.***

The Figure gives an overview of important direct and indirect effects of the alpha genus HPV E6 and E7 proteins on cellular pathways and processes. Some of the areas of overlap and interplay between the functions of the two oncoproteins are shown. Important roles of E6 and E7 include inhibition (low-risk HPV) and degradation (high-risk HPV) of cellular p53 and pRb, respectively. Loss of p53 has multiple consequences (some not illustrated), including effects on proliferation, DNA repair, senescence and apoptosis. The diagram does not take account of dose dependent effects of the virus oncoproteins nor their level of expression throughout progression. (Red oval = general downregulation of cellular process or pathway. Green oval = general upregulation of cellular process or pathway. Brown oval = modulation of cellular process or pathway. HATs = histone acetyltransferases, including p300, CBP, pCAF and TIP60. CDKs = cyclin-dependent kinases.)

Figure 4. ***The HPV16 long control region (LCR).***

The LCR is positioned between the late and early virus gene regions. The virus replication protein E1 binds as a dimer of hexamers at the origin of replication (ORI), while the virus transcription factor E2 associates as a dimer at four E2 binding sites (E2BSs) which broadly define the 5', central and 3' regions. Transcription of early genes occurs from the early promoter (P97) and is dictated by the binding of numerous host transcription factors and the virus E2 dimer across the enhancer, silencer and promoter regions. Activation of transcription from the late promoter (P670) is dependent upon cell differentiation and binding of

differentiation-associated transcription factors. (Poly(A)L = Late polyadenylation site; ORF = open reading frame.)

References

1. Forman D, de Martel C, Lacey CJ, *et al.* Global burden of human papillomavirus and related diseases. *Vaccine* 2012; **30 Suppl 5**: F12-23.
2. Doorbar J, Quint W, Banks L, *et al.* The biology and life-cycle of human papillomaviruses. *Vaccine* 2012; **30 Suppl 5**: F55-70.
3. Bernard HU, Burk RD, Chen Z, *et al.* Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 2010; **401**: 70-79.
4. de Villiers EM. Cross-roads in the classification of papillomaviruses. *Virology* 2013; **445**: 2-10.
5. Stanley MA, Pett MR, Coleman N. HPV: from infection to cancer. *Biochem Soc Trans* 2007; **35**: 1456-1460.
6. IARC. Biological agents. Volume 100 B. A review of human carcinogens. *IARC Monogr Eval Carcinog Risks Hum* 2012; **100**: 1-441.
7. zur Hausen H. Papillomaviruses in the causation of human cancers - a brief historical account. *Virology* 2009; **384**: 260-265.
8. Stanley M. Immunobiology of HPV and HPV vaccines. *Gynecol Oncol* 2008; **109**: S15-21.
9. Liaw KL, Hildesheim A, Burk RD, *et al.* A prospective study of human papillomavirus (HPV) type 16 DNA detection by polymerase chain reaction and its association with acquisition and persistence of other HPV types. *J Infect Dis* 2001; **183**: 8-15.
10. Woodman CB, Collins S, Winter H, *et al.* Natural history of cervical human papillomavirus infection in young women: a longitudinal cohort study. *Lancet* 2001; **357**: 1831-1836.
11. Schiller JT, Day PM, Kines RC. Current understanding of the mechanism of HPV infection. *Gynecol Oncol* 2010; **118**: S12-17.
12. Pyeon D, Pearce SM, Lank SM, *et al.* Establishment of human papillomavirus infection requires cell cycle progression. *PLoS Pathog* 2009; **5**: e1000318.
13. Aydin I, Weber S, Snijder B, *et al.* Large Scale RNAi Reveals the Requirement of Nuclear Envelope Breakdown for Nuclear Import of Human Papillomaviruses. *PLoS Pathog* 2014; **10**: e1004162.
14. Lopez J, Ruiz G, Organista-Nava J, *et al.* Human papillomavirus infections and cancer stem cells of tumors from the uterine cervix. *Open Virol J* 2012; **6**: 232-240.
15. Herfs M, Yamamoto Y, Laury A, *et al.* A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer. *Proc Natl Acad Sci U S A* 2012; **109**: 10516-10521.
16. Doorbar J. Molecular biology of human papillomavirus infection and cervical cancer. *Clin Sci (Lond)* 2006; **110**: 525-541.
17. Stanley MA, Browne HM, Appleby M, *et al.* Properties of a non-tumorigenic human cervical keratinocyte cell line. *Int J Cancer* 1989; **43**: 672-676.
18. Bedell MA, Hudson JB, Golub TR, *et al.* Amplification of human papillomavirus genomes in vitro is dependent on epithelial differentiation. *J Virol* 1991; **65**: 2254-2260.
19. Parish JL, Bean AM, Park RB, *et al.* ChlR1 is required for loading papillomavirus E2 onto mitotic chromosomes and viral genome maintenance. *Mol Cell* 2006; **24**: 867-876.
20. McBride AA. Replication and partitioning of papillomavirus genomes. *Adv Virus Res* 2008; **72**: 155-205.
21. Van Tine BA, Dao LD, Wu SY, *et al.* Human papillomavirus (HPV) origin-binding protein associates with mitotic spindles to enable viral DNA partitioning. *Proc Natl Acad Sci U S A* 2004; **101**: 4030-4035.
22. Thomas JT, Hubert WG, Ruesch MN, *et al.* Human papillomavirus type 31 oncoproteins E6 and E7 are required for the maintenance of episomes during the viral life cycle in normal human keratinocytes. *Proc Natl Acad Sci U S A* 1999; **96**: 8449-8454.
23. Park RB, Androphy EJ. Genetic analysis of high-risk e6 in episomal maintenance of human papillomavirus genomes in primary human keratinocytes. *J Virol* 2002; **76**: 11359-11364.

24. Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer* 2010; **10**: 550-560.
25. Munger K, Howley PM. Human papillomavirus immortalization and transformation functions. *Virus Res* 2002; **89**: 213-228.
26. Chow LT, Broker TR, Steinberg BM. The natural history of human papillomavirus infections of the mucosal epithelia. *APMIS* 2010; **118**: 422-449.
27. Mole S, McFarlane M, Chuen-Im T*, et al.* RNA splicing factors regulated by HPV16 during cervical tumour progression. *J Pathol* 2009; **219**: 383-391.
28. Johansson C, Somberg M, Li X*, et al.* HPV-16 E2 contributes to induction of HPV-16 late gene expression by inhibiting early polyadenylation. *EMBO J* 2012; **31**: 3212-3227.
29. Johansson C, Schwartz S. Regulation of human papillomavirus gene expression by splicing and polyadenylation. *Nat Rev Microbiol* 2013; **11**: 239-251.
30. Baldwin P, Laskey R, Coleman N. Translational approaches to improving cervical screening. *Nat Rev Cancer* 2003; **3**: 217-226.
31. Middleton K, Peh W, Southern S*, et al.* Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. *J Virol* 2003; **77**: 10186-10201.
32. McIndoe WA, McLean MR, Jones RW*, et al.* The invasive potential of carcinoma in situ of the cervix. *Obstet Gynecol* 1984; **64**: 451-458.
33. Park TW, Richart RM, Sun XW*, et al.* Association between human papillomavirus type and clonal status of cervical squamous intraepithelial lesions. *J Natl Cancer Inst* 1996; **88**: 355-358.
34. Jeon S, Allen-Hoffmann BL, Lambert PF. Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *J Virol* 1995; **69**: 2989-2997.
35. Williams GH, Romanowski P, Morris L*, et al.* Improved cervical smear assessment using antibodies against proteins that regulate DNA replication. *Proc Natl Acad Sci U S A* 1998; **95**: 14932-14937.
36. Pett M, Coleman N. Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? *J Pathol* 2007; **212**: 356-367.
37. Gray E, Pett MR, Ward D*, et al.* In vitro progression of human papillomavirus 16 episome-associated cervical neoplasia displays fundamental similarities to integrant-associated carcinogenesis. *Cancer Res* 2010; **70**: 4081-4091.
38. Cullen AP, Reid R, Campion M*, et al.* Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. *J Virol* 1991; **65**: 606-612.
39. Pirami L, Giache V, Becciolini A. Analysis of HPV16, 18, 31, and 35 DNA in pre-invasive and invasive lesions of the uterine cervix. *J Clin Pathol* 1997; **50**: 600-604.
40. Badaracco G, Venuti A, Sedati A*, et al.* HPV16 and HPV18 in genital tumors: Significantly different levels of viral integration and correlation to tumor invasiveness. *J Med Virol* 2002; **67**: 574-582.
41. Vinokurova S, Wentzensen N, Kraus I*, et al.* Type-dependent integration frequency of human papillomavirus genomes in cervical lesions. *Cancer Res* 2008; **68**: 307-313.
42. Kalantari M, Blennow E, Hagmar B*, et al.* Physical state of HPV16 and chromosomal mapping of the integrated form in cervical carcinomas. *Diagn Mol Pathol* 2001; **10**: 46-54.
43. Daniel B, Mukherjee G, Seshadri L*, et al.* Changes in the physical state and expression of human papillomavirus type 16 in the progression of cervical intraepithelial neoplasia lesions analysed by PCR. *J Gen Virol* 1995; **76 (Pt 10)**: 2589-2593.
44. Peitsaro P, Johansson B, Syrjanen S. Integrated human papillomavirus type 16 is frequently found in cervical cancer precursors as demonstrated by a novel quantitative real-time PCR technique. *J Clin Microbiol* 2002; **40**: 886-891.

45. Hudelist G, Manavi M, Pischinger KI, *et al.* 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.
46. Kulmala SM, Syrjanen SM, Gyllensten UB, *et al.* Early integration of high copy HPV16 detectable in women with normal and low grade cervical cytology and histology. *J Clin Pathol* 2006; **59**: 513-517.
47. Hopman AH, Smedts F, Dignef W, *et al.* Transition of high-grade cervical intraepithelial neoplasia to micro-invasive carcinoma is characterized by integration of HPV 16/18 and numerical chromosome abnormalities. *J Pathol* 2004; **202**: 23-33.
48. Arias-Pulido H, Peyton CL, Joste NE, *et al.* Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. *J Clin Microbiol* 2006; **44**: 1755-1762.
49. Van Tine BA, Kappes JC, Banerjee NS, *et al.* Clonal selection for transcriptionally active viral oncogenes during progression to cancer. *J Virol* 2004; **78**: 11172-11186.
50. Klaes R, Woerner SM, Ridder R, *et al.* Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes. *Cancer Res* 1999; **59**: 6132-6136.
51. Evans MF, Cooper K. Human papillomavirus integration: detection by in situ hybridization and potential clinical application. *J Pathol* 2004; **202**: 1-4.
52. Cooper K, Herrington CS, Stickland JE, *et al.* Episomal and integrated human papillomavirus in cervical neoplasia shown by non-isotopic in situ hybridisation. *J Clin Pathol* 1991; **44**: 990-996.
53. Durst M, Kleinheinz A, Hotz M, *et al.* The physical state of human papillomavirus type 16 DNA in benign and malignant genital tumours. *J Gen Virol* 1985; **66 (Pt 7)**: 1515-1522.
54. Thorland EC, Myers SL, Persing DH, *et al.* Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites. *Cancer Res* 2000; **60**: 5916-5921.
55. Nagao S, Yoshinouchi M, Miyagi Y, *et al.* Rapid and sensitive detection of physical status of human papillomavirus type 16 DNA by quantitative real-time PCR. *J Clin Microbiol* 2002; **40**: 863-867.
56. Ziegert C, Wentzensen N, Vinokurova S, *et al.* A comprehensive analysis of HPV integration loci in anogenital lesions combining transcript and genome-based amplification techniques. *Oncogene* 2003; **22**: 3977-3984.
57. Bechtold V, Beard P, Raj K. Human papillomavirus type 16 E2 protein has no effect on transcription from episomal viral DNA. *J Virol* 2003; **77**: 2021-2028.
58. Pett MR, Herdman MT, Palmer RD, *et al.* Selection of cervical keratinocytes containing integrated HPV16 associates with episome loss and an endogenous antiviral response. *Proc Natl Acad Sci U S A* 2006; **103**: 3822-3827.
59. Herdman MT, Pett MR, Roberts I, *et al.* Interferon-beta treatment of cervical keratinocytes naturally infected with human papillomavirus 16 episomes promotes rapid reduction in episome numbers and emergence of latent integrants. *Carcinogenesis* 2006; **27**: 2341-2353.
60. 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. *Proc Natl Acad Sci U S A* 1995; **92**: 1654-1658.
61. Chaiwongkot A, Vinokurova S, Pientong C, *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**: 2087-2094.
62. Dall KL, Scarpini CG, Roberts I, *et al.* Characterization of naturally occurring HPV16 integration sites isolated from cervical keratinocytes under noncompetitive conditions. *Cancer Res* 2008; **68**: 8249-8259.
63. Scarpini CG, Groves IJ, Pett MR, *et al.* Virus transcript levels and cell growth rates after naturally occurring HPV16 integration events in basal cervical keratinocytes. *J Pathol* 2014; **233**: 281-293.

64. Pater MM, Hughes GA, Hyslop DE*, et al.* Glucocorticoid-dependent oncogenic transformation by type 16 but not type 11 human papilloma virus DNA. *Nature* 1988; **335**: 832-835.
65. Piccini A, Storey A, Romanos M*, et al.* Regulation of human papillomavirus type 16 DNA replication by E2, glucocorticoid hormone and epidermal growth factor. *J Gen Virol* 1997; **78 (Pt 8)**: 1963-1970.
66. Arbeit JM, Howley PM, Hanahan D. Chronic estrogen-induced cervical and vaginal squamous carcinogenesis in human papillomavirus type 16 transgenic mice. *Proc Natl Acad Sci U S A* 1996; **93**: 2930-2935.
67. von Knebel Doeberitz M, Bauknecht T, Bartsch D*, et al.* Influence of chromosomal integration on glucocorticoid-regulated transcription of growth-stimulating papillomavirus genes E6 and E7 in cervical carcinoma cells. *Proc Natl Acad Sci U S A* 1991; **88**: 1411-1415.
68. Yu T, Ferber MJ, Cheung TH*, et al.* The role of viral integration in the development of cervical cancer. *Cancer Genet Cytogenet* 2005; **158**: 27-34.
69. Kessis TD, Connolly DC, Hedrick L*, et al.* Expression of HPV16 E6 or E7 increases integration of foreign DNA. *Oncogene* 1996; **13**: 427-431.
70. Pett MR, Alazawi WO, Roberts I*, et al.* Acquisition of high-level chromosomal instability is associated with integration of human papillomavirus type 16 in cervical keratinocytes. *Cancer Res* 2004; **64**: 1359-1368.
71. Winder DM, Pett MR, Foster N*, et al.* An increase in DNA double-strand breaks, induced by Ku70 depletion, is associated with human papillomavirus 16 episome loss and de novo viral integration events. *J Pathol* 2007; **213**: 27-34.
72. Wentzensen N, Vinokurova S, von Knebel Doeberitz M. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. *Cancer Res* 2004; **64**: 3878-3884.
73. Schmitz M, Driesch C, Jansen L*, et al.* Non-random integration of the HPV genome in cervical cancer. *PLoS One* 2012; **7**: e39632.
74. Jang MK, Shen K, McBride AA. Papillomavirus Genomes Associate with BRD4 to Replicate at Fragile Sites in the Host Genome. *PLoS Pathog* 2014; **10**: e1004117.
75. Reuter S, Bartelmann M, Vogt M*, et al.* APM-1, a novel human gene, identified by aberrant co-transcription with papillomavirus oncogenes in a cervical carcinoma cell line, encodes a BTB/POZ-zinc finger protein with growth inhibitory activity. *EMBO J* 1998; **17**: 215-222.
76. Thorland EC, Myers SL, Gostout BS*, et al.* Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. *Oncogene* 2003; **22**: 1225-1237.
77. Ferber MJ, Thorland EC, Brink AA*, et al.* Preferential integration of human papillomavirus type 18 near the c-myc locus in cervical carcinoma. *Oncogene* 2003; **22**: 7233-7242.
78. Peter M, Rosty C, Couturier J*, et al.* MYC activation associated with the integration of HPV DNA at the MYC locus in genital tumors. *Oncogene* 2006; **25**: 5985-5993.
79. Couturier J, Sastre-Garau X, Schneider-Maunoury S*, et al.* Integration of papillomavirus DNA near myc genes in genital carcinomas and its consequences for proto-oncogene expression. *J Virol* 1991; **65**: 4534-4538.
80. Adey A, Burton JN, Kitzman JO*, et al.* The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line. *Nature* 2013; **500**: 207-211.
81. Landry JJ, Pyl PT, Rausch T*, et al.* The genomic and transcriptomic landscape of a HeLa cell line. *G3 (Bethesda)* 2013; **3**: 1213-1224.
82. Akagi K, Li J, Broutian TR*, et al.* Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. *Genome Res* 2014; **24**: 185-199.
83. Tang KW, Alaei-Mahabadi B, Samuelsson T*, et al.* The landscape of viral expression and host gene fusion and adaptation in human cancer. *Nat Commun* 2013; **4**: 2513.
84. Ojesina AI, Lichtenstein L, Freeman SS*, et al.* Landscape of genomic alterations in cervical carcinomas. *Nature* 2014; **506**: 371-375.
85. Wentzensen N, Ridder R, Klaes R*, et al.* Characterization of viral-cellular fusion transcripts in a large series of HPV16 and 18 positive anogenital lesions. *Oncogene* 2002; **21**: 419-426.

86. Schmitz M, Driesch C, Beer-Grondke K, *et al.* Loss of gene function as a consequence of human papillomavirus DNA integration. *Int J Cancer* 2012; **131**: E593-602.
87. Xu B, Chotewutmontri S, Wolf S, *et al.* Multiplex Identification of Human Papillomavirus 16 DNA Integration Sites in Cervical Carcinomas. *PLoS One* 2013; **8**: e66693.
88. Olthof NC, Huebbers CU, Kolligs J, *et al.* Viral load, gene expression and mapping of viral integration sites in HPV16-associated HNSCC cell lines. *Int J Cancer* 2014.
89. Dong XP, Stubenrauch F, Beyer-Finkler E, *et al.* Prevalence of deletions of YY1-binding sites in episomal HPV 16 DNA from cervical cancers. *Int J Cancer* 1994; **58**: 803-808.
90. May M, Dong XP, Beyer-Finkler E, *et al.* The E6/E7 promoter of extrachromosomal HPV16 DNA in cervical cancers escapes from cellular repression by mutation of target sequences for YY1. *EMBO J* 1994; **13**: 1460-1466.
91. Veress G, Szarka K, Dong XP, *et al.* Functional significance of sequence variation in the E2 gene and the long control region of human papillomavirus type 16. *J Gen Virol* 1999; **80 (Pt 4)**: 1035-1043.
92. Lace MJ, Isacson C, Anson JR, *et al.* Upstream regulatory region alterations found in human papillomavirus type 16 (HPV-16) isolates from cervical carcinomas increase transcription, ori function, and HPV immortalization capacity in culture. *J Virol* 2009; **83**: 7457-7466.
93. Thain A, Jenkins O, Clarke AR, *et al.* CpG methylation directly inhibits binding of the human papillomavirus type 16 E2 protein to specific DNA sequences. *J Virol* 1996; **70**: 7233-7235.
94. Badal V, Chuang LS, Tan EH, *et al.* CpG methylation of human papillomavirus type 16 DNA in cervical cancer cell lines and in clinical specimens: genomic hypomethylation correlates with carcinogenic progression. *J Virol* 2003; **77**: 6227-6234.
95. Hublarova P, Hrstka R, Rotterova P, *et al.* Prediction of human papillomavirus 16 e6 gene expression and cervical intraepithelial neoplasia progression by methylation status. *Int J Gynecol Cancer* 2009; **19**: 321-325.
96. Mazumder Indra D, Singh RK, Mitra S, *et al.* Genetic and epigenetic changes of HPV16 in cervical cancer differentially regulate E6/E7 expression and associate with disease progression. *Gynecol Oncol* 2011; **123**: 597-604.
97. Xi LF, Jiang M, Shen Z, *et al.* Inverse association between methylation of human papillomavirus type 16 DNA and risk of cervical intraepithelial neoplasia grades 2 or 3. *PLoS One* 2011; **6**: e23897.
98. 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**: 280-285.
99. Ding DC, Chiang MH, Lai HC, *et al.* Methylation of the long control region of HPV16 is related to the severity of cervical neoplasia. *Eur J Obstet Gynecol Reprod Biol* 2009; **147**: 215-220.
100. Hong D, Ye F, Lu W, *et al.* Methylation status of the long control region of HPV 16 in clinical cervical specimens. *Mol Med Rep* 2008; **1**: 555-560.
101. Kalantari M, Lee D, Calleja-Macias IE, *et al.* Effects of cellular differentiation, chromosomal integration and 5-aza-2'-deoxycytidine treatment on human papillomavirus-16 DNA methylation in cultured cell lines. *Virology* 2008; **374**: 292-303.
102. Kim K, Garner-Hamrick PA, Fisher C, *et al.* Methylation patterns of papillomavirus DNA, its influence on E2 function, and implications in viral infection. *J Virol* 2003; **77**: 12450-12459.
103. Vinokurova S, von Knebel Doeberitz M. Differential methylation of the HPV 16 upstream regulatory region during epithelial differentiation and neoplastic transformation. *PLoS One* 2011; **6**: e24451.
104. Favre M, Breitburd F, Croissant O, *et al.* Chromatin-like structures obtained after alkaline disruption of bovine and human papillomaviruses. *J Virol* 1977; **21**: 1205-1209.
105. Rosl F, Westphal EM, zur Hausen H. Chromatin structure and transcriptional regulation of human papillomavirus type 18 DNA in HeLa cells. *Mol Carcinog* 1989; **2**: 72-80.
106. Stunkel W, Bernard HU. The chromatin structure of the long control region of human papillomavirus type 16 represses viral oncoprotein expression. *J Virol* 1999; **73**: 1918-1930.

107. del Mar Pena LM, Laimins LA. Differentiation-dependent chromatin rearrangement coincides with activation of human papillomavirus type 31 late gene expression. *J Virol* 2001; **75**: 10005-10013.
108. Kruppel U, Muller-Schiffmann A, Baldus SE*, et al.* E2 and the co-activator p300 can cooperate in activation of the human papillomavirus type 16 early promoter. *Virology* 2008; **377**: 151-159.
109. Wang WM, Wu SY, Lee AY*, et al.* Binding site specificity and factor redundancy in activator protein-1-driven human papillomavirus chromatin-dependent transcription. *J Biol Chem* 2011; **286**: 40974-40986.
110. Wooldridge TR, Laimins LA. Regulation of human papillomavirus type 31 gene expression during the differentiation-dependent life cycle through histone modifications and transcription factor binding. *Virology* 2008; **374**: 371-380.
111. Sichtig N, Korfer N, Steger G. Papillomavirus binding factor binds to SAP30 and represses transcription via recruitment of the HDAC1 co-repressor complex. *Arch Biochem Biophys* 2007; **467**: 67-75.
112. Chakraborty S, Das K, Saha S*, et al.* Nuclear matrix protein SMAR1 represses c-Fos-mediated HPV18 E6 transcription through alteration of chromatin histone de-acetylation. *J Biol Chem* 2014.
113. Carson A, Khan SA. Characterization of transcription factor binding to human papillomavirus type 16 DNA during cellular differentiation. *J Virol* 2006; **80**: 4356-4362.
114. De-Castro Arce J, Gockel-Krzikalla E, Rosl F. Silencing of multi-copy HPV16 by viral self-methylation and chromatin occlusion: a model for epigenetic virus-host interaction. *Hum Mol Genet* 2012; **21**: 1693-1705.
115. Johannsen E, Lambert PF. Epigenetics of human papillomaviruses. *Virology* 2013; **445**: 205-212.
116. Gerstein MB, Lu ZJ, Van Nostrand EL*, et al.* Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project. *Science* 2010; **330**: 1775-1787.
117. Yip KY, Cheng C, Bhardwaj N*, et al.* Classification of human genomic regions based on experimentally determined binding sites of more than 100 transcription-related factors. *Genome Biol* 2012; **13**: R48.
118. McLaughlin-Drubin ME, Huh KW, Munger K. Human papillomavirus type 16 E7 oncoprotein associates with E2F6. *J Virol* 2008; **82**: 8695-8705.
119. McLaughlin-Drubin ME, Munger K. Biochemical and functional interactions of human papillomavirus proteins with polycomb group proteins. *Viruses* 2013; **5**: 1231-1249.
120. McLaughlin-Drubin ME, Crum CP, Munger K. Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming. *Proc Natl Acad Sci U S A* 2011; **108**: 2130-2135.
121. Hyland PL, McDade SS, McCloskey R*, et al.* Evidence for alteration of EZH2, BMI1, and KDM6A and epigenetic reprogramming in human papillomavirus type 16 E6/E7-expressing keratinocytes. *J Virol* 2011; **85**: 10999-11006.
122. Kong L, Yu XP, Bai XH*, et al.* RbAp48 is a critical mediator controlling the transforming activity of human papillomavirus type 16 in cervical cancer. *J Biol Chem* 2007; **282**: 26381-26391.
123. Bodily JM, Mehta KP, Laimins LA. Human papillomavirus E7 enhances hypoxia-inducible factor 1-mediated transcription by inhibiting binding of histone deacetylases. *Cancer Res* 2011; **71**: 1187-1195.
124. Hasan UA, Zannetti C, Parroche P*, et al.* The human papillomavirus type 16 E7 oncoprotein induces a transcriptional repressor complex on the Toll-like receptor 9 promoter. *J Exp Med* 2013; **210**: 1369-1387.
125. Brehm A, Nielsen SJ, Miska EA*, et al.* The E7 oncoprotein associates with Mi2 and histone deacetylase activity to promote cell growth. *EMBO J* 1999; **18**: 2449-2458.

126. Lee D, Lim C, Seo T, *et al.* The viral oncogene human papillomavirus E7 deregulates transcriptional silencing by Brm-related gene 1 via molecular interactions. *J Biol Chem* 2002; **277**: 48842-48848.
127. Lee K, Lee AY, Kwon YK, *et al.* Suppression of HPV E6 and E7 expression by BAF53 depletion in cervical cancer cells. *Biochem Biophys Res Commun* 2011; **412**: 328-333.
128. He H, Luo Y. Brg1 regulates the transcription of human papillomavirus type 18 E6 and E7 genes. *Cell Cycle* 2012; **11**: 617-627.
129. Burgers WA, Blanchon L, Pradhan S, *et al.* Viral oncoproteins target the DNA methyltransferases. *Oncogene* 2007; **26**: 1650-1655.
130. Laurson J, Khan S, Chung R, *et al.* Epigenetic repression of E-cadherin by human papillomavirus 16 E7 protein. *Carcinogenesis* 2010; **31**: 918-926.
131. Au Yeung CL, Tsang WP, Tsang TY, *et al.* HPV-16 E6 upregulation of DNMT1 through repression of tumor suppressor p53. *Oncol Rep* 2010; **24**: 1599-1604.
132. Leonard SM, Wei W, Collins SI, *et al.* Oncogenic human papillomavirus imposes an instructive pattern of DNA methylation changes which parallel the natural history of cervical HPV infection in young women. *Carcinogenesis* 2012; **33**: 1286-1293.
133. Szalmas A, Konya J. Epigenetic alterations in cervical carcinogenesis. *Semin Cancer Biol* 2009; **19**: 144-152.
134. Steenbergen RD, Snijders PJ, Heideman DA, *et al.* Clinical implications of (epi)genetic changes in HPV-induced cervical precancerous lesions. *Nat Rev Cancer* 2014; **14**: 395-405.
135. Forbes SA, Bindal N, Bamford S, *et al.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res* 2011; **39**: D945-950.
136. Stransky N, Egloff AM, Tward AD, *et al.* The mutational landscape of head and neck squamous cell carcinoma. *Science* 2011; **333**: 1157-1160.
137. Agrawal N, Frederick MJ, Pickering CR, *et al.* Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science* 2011; **333**: 1154-1157.
138. Wright AA, Howitt BE, Myers AP, *et al.* Oncogenic mutations in cervical cancer: genomic differences between adenocarcinomas and squamous cell carcinomas of the cervix. *Cancer* 2013; **119**: 3776-3783.
139. Muralidhar B, Winder D, Murray M, *et al.* Functional evidence that Drosha overexpression in cervical squamous cell carcinoma affects cell phenotype and microRNA profiles. *J Pathol* 2011; **224**: 496-507.
140. Zheng ZM, Wang X. Regulation of cellular miRNA expression by human papillomaviruses. *Biochim Biophys Acta* 2011; **1809**: 668-677.
141. Greco D, Kivi N, Qian K, *et al.* Human papillomavirus 16 E5 modulates the expression of host microRNAs. *PLoS One* 2011; **6**: e21646.
142. de Freitas AC, Coimbra EC, Leitao Mda C. Molecular targets of HPV oncoproteins: potential biomarkers for cervical carcinogenesis. *Biochim Biophys Acta* 2014; **1845**: 91-103.
143. Wilting SM, Snijders PJ, Verlaat W, *et al.* Altered microRNA expression associated with chromosomal changes contributes to cervical carcinogenesis. *Oncogene* 2013; **32**: 106-116.
144. Qian K, Pietila T, Ronty M, *et al.* Identification and validation of human papillomavirus encoded microRNAs. *PLoS One* 2013; **8**: e70202.
145. Gu W, An J, Ye P, *et al.* Prediction of conserved microRNAs from skin and mucosal human papillomaviruses. *Arch Virol* 2011; **156**: 1161-1171.
146. Cai X, Li G, Laimins LA, *et al.* Human papillomavirus genotype 31 does not express detectable microRNA levels during latent or productive virus replication. *J Virol* 2006; **80**: 10890-10893.
147. Jiang M, Milner J. Selective silencing of viral gene expression in HPV-positive human cervical carcinoma cells treated with siRNA, a primer of RNA interference. *Oncogene* 2002; **21**: 6041-6048.

148. Hanning JE, Saini HK, Murray MJ, *et al.* Depletion of HPV16 early genes induces autophagy and senescence in a cervical carcinogenesis model, regardless of viral physical state. *J Pathol* 2013; **231**: 354-366.