

The natural history of cervical HPV infection: unresolved issues

Ciaran B. J. Woodman*, Stuart I. Collins† and Lawrence S. Young*

Abstract | The identification of high-risk human papillomavirus (HPV) types as a necessary cause of cervical cancer offers the prospect of effective primary prevention and the possibility of improving the efficiency of cervical screening programmes. However, for these opportunities to be realized, a more complete understanding of the natural history of HPV infection, and its relationship to the development of epithelial abnormalities of the cervix, is required. We discuss areas of uncertainty, and their possible effect on disease prevention strategies.

Adenocarcinoma

A malignant tumour originating in glandular tissue.

Cross-sectional study

A study examining the association between disease and exposure at one point in time (a prevalence study). The temporal sequence of cause and effect cannot be determined with this study design.

Episome

A piece of hereditary material that can exist as free, autonomously replicating DNA.

Cervical cancer is the second most common cancer among women worldwide¹. The majority of cases occur in the developing world, where, in most countries, it is the leading cause of cancer mortality in women². In many developed countries, the incidence of squamous cell carcinoma of the cervix has been falling for some time, although that of adenocarcinoma of the cervix is now rising^{1,3,4}.

Over 100 human papillomavirus (HPV) types have been identified, of which 40 infect the genital tract⁵. Cervical HPV infection is a common sexually transmitted infection. Most women are infected shortly after beginning their first sexual relationship⁶, with the highest prevalence seen in women under 25 years of age^{7,8}. Thereafter, prevalence decreases rapidly. In young and middle-aged women, HPV infections are usually transient, at least when their duration is measured by how long the virus can be detected in cytological samples^{9–11}. Virus might be detected only intermittently; and the concurrent or sequential detection of different HPV types is common^{12–18}. Cross-sectional studies indicate a second peak of infection in older women close to the age when the incidence of cervical cancer is maximum^{7,8}.

A stream of epidemiological and laboratory-based research has identified infection with any one of 15 high-risk, or oncogenic, HPV types as a necessary but not sufficient cause of cervical cancer^{19–21}. Whereas HPV18 is the type most strongly associated with adenocarcinoma of the cervix, HPV16, followed by HPV18, are the types most frequently detected when squamous cell carcinoma is diagnosed^{22,23}. The frequency with which HPV16 is found in integrated forms increases with the severity of cervical neoplasia,

although in some women with invasive disease only episomal forms are detected. By contrast, HPV18 is almost always found in only integrated forms in women with high-grade cervical intraepithelial neoplasia (HCIN) and invasive disease (FIG. 1). A bivalent HPV (types 16 and 18) and a quadrivalent HPV (types 6, 11, 16 and 18) vaccine are being evaluated in phase III clinical trials. Preliminary results indicate that these prophylactic HPV virus-like particle vaccines are effective in preventing infections with, and epithelial abnormalities caused by, the targeted HPV types^{20,24}.

Cervical cancer is characterized by a well-defined pre-malignant phase that can be suspected on cytological examination of exfoliated cervical cells and confirmed on histological examination of cervical material. These pre-malignant changes represent a spectrum of histological abnormalities ranging from CIN1 (mild dysplasia) to CIN2 (moderate dysplasia) to CIN3 (severe dysplasia/carcinoma *in situ*). Although the treatment of cervical pre-malignant changes is therapeutically efficacious, it is also procedurally inefficient. This situation has arisen because of uncertainties surrounding the natural history of CIN. Cytological and histological examinations cannot reliably distinguish the few women with abnormal smears who will progress to invasive cancer from the vast majority of those whose abnormalities will spontaneously regress. Were a population-based prophylactic immunization programme introduced using either of the vaccines under consideration, and were it to achieve widespread coverage, then not only could it prevent up to 70% of all cervical cancers, but it could also reduce the costs of cervical screening programmes. However, it is unlikely that these screening programmes could be

*Cancer Research UK
Institute for Cancer Studies,
University of Birmingham,
Edgbaston, Birmingham
B15 2TT, UK.

†Cancer Research UK Clinical
Trials Unit, University of
Birmingham, Edgbaston,
Birmingham B15 2TT, UK.
Correspondence to C.B.J.W.
e-mail:
c.b.woodman@bham.ac.uk
doi:10.1038/nrc2050

At a glance

- The most frequently detected human papillomavirus (HPV) type at the time of diagnosis of squamous cell carcinoma (SCC) is HPV16, followed by HPV18. HPV18 is the type most strongly associated with adenocarcinoma of the cervix, which is increasing in incidence at the same time as the incidence of SCC is falling.
- A bivalent HPV (types 16 and 18) and a quadrivalent HPV (types 6, 11, 16 and 18) vaccine are now being evaluated in phase III clinical trials, and have the potential to prevent about 70% of all cervical cancers. The quadrivalent HPV vaccine (Gardasil, Merck) has recently gained FDA and European Commission approval for use in women between the ages of 9–26.
- Although most women will at some time be infected with HPV, very few will progress to invasive disease. The identification of more robust markers of disease progression requires a more complete understanding of the natural history of type-specific HPV infections.
- It is unknown whether persistent HPV infections are characterized by the continuing detection of HPV, or by a state of viral latency during which the virus remains undetectable only to reappear later. A clearer understanding of these issues is essential for the effective implementation of screening strategies that include testing for HPV.
- Integration of HPV into the host genome results in a loss of negative-feedback control of oncogene expression, following disruption of the viral regulatory early gene *E2*. Whether the integration event itself is crucial to carcinogenesis is the subject of continuing debate.
- The prevalence of integrated forms varies with the infecting HPV type. Unlike HPV16, HPV18 integration seems virtually complete in women with cervical intraepithelial neoplasia grade 3 (CIN3) or invasive disease.
- The association between viral load and cervical disease varies with the HPV type, the physical state of the virus and the heterogeneity of the cervical lesion. The complexity of these relationships indicates that a measurement of viral load is not clinically useful.
- The concurrent or sequential detection of more than one HPV type is common. There is some evidence to indicate that the life cycles of different HPV types are not independent of each other, as has previously been assumed.
- HPV oncogenes can activate the cellular methylation machinery. The pattern of HPV gene methylation varies with the viral life cycle, the presence of disease and possibly the HPV type.
- Aberrant methylation of CpG islands in the promoter regions of tumour suppressor genes is one of several epigenetic changes that can contribute to carcinogenesis. The detection of these epigenetic changes in exfoliated cervical cells could improve the effectiveness of cervical screening programmes.

Cervical intraepithelial neoplasia

(CIN) A disease characterized by precancerous changes in, and confined to, the epithelial cells lining the cervix.

Dysplasia

An epithelial abnormality in which the cells become disorganized, which is characterized by developmental changes in cell growth, shape and organization.

Natural history

The course of disease or infection from onset to resolution.

Epigenetic

Inherited changes in gene expression resulting from altered chromatin structure or DNA modification rather than changes in DNA sequence.

discontinued or even scaled-down for several reasons: immunization will only protect against HPV types that are targeted by the vaccine; protection will not be absolute and its longevity is uncertain; as yet, the possibility of genotype replacement cannot be excluded; and older women not covered by vaccination programmes will continue to be at risk.

Although most women will at some time have been infected with HPV, few will progress to invasive disease. Therefore, there is a continuing need for more robust markers of disease progression than those provided by morphological examination, or testing for the presence of high-risk HPV types. The identification of viral and host factors that modulate the risk of disease progression in women infected with HPV requires a more complete understanding of the natural history of HPV infection, and its relationship to the acquisition of epithelial abnormalities. Failure to do so runs the risk of compounding our imperfect understanding of the disease process with an even less perfect understanding of the natural history of HPV infection, therefore

presenting the clinician with another set of management dilemmas.

The contribution of HPV to the pathogenesis of cervical cancer, and issues relating to vaccination have been discussed in previous reviews^{19,20}. Here, we first consider the conceptual and methodological impediments to a clearer understanding of HPV as a persistent viral infection, and how this might affect the efficacy of HPV-based screening strategies. Before considering the possible synergy between different HPV types, we explore how the exposure–disease relationship varies with viral load, viral integration status and the infecting HPV type; these considerations are important in the context of prophylactic vaccines that might not achieve sterilizing immunity. Because progression to invasive disease requires an accumulation of genetic and epigenetic events, we consider how HPV can activate the cellular DNA methylation machinery and therefore epigenetically regulate both viral and cellular genes. Finally, we consider how the detection of these epigenetic changes in exfoliated cervical cells might improve the effectiveness of cervical screening programmes.

Persistent HPV infection

Many viruses establish persistent infections that are characterized by continuous low or high levels of viral replication (for example human immunodeficiency virus and hepatitis B virus) or by periodic reactivation of a latent infection following apparently disease-free intervals (for example herpes simplex virus)²⁵ (FIG. 2).

Although it is now widely believed that a persistent infection with a high-risk HPV type is necessary for the development of HGCI and invasive disease, the term ‘persistence’ has often been loosely defined when testing this hypothesis. In many of these studies, the occurrence of disease in women who test positive for HPV on two or more occasions (persistent infection) has been compared with that in women who test positive only once (transient infection). There are numerous conceptual problems with this approach. When defined in this way, the duration of a persistent infection is not a constant, but will vary depending on the interval between tests used in each study. In studies using this approach, the interval between tests ranged from 2 months to 7 years, with a median of 6 months^{12,26–49}.

A more fundamental problem relates to inferences drawn from observations made at indeterminate points during the natural history of the infection. In these circumstances, the distinction between a persistent and transient infection is arbitrary to the extent that it is dependent on both the timing of the samples in relation to the natural history of the infection, and the interval between samples (FIG. 3a–c). In particular, it is impossible to determine how long a woman, who tests positive in her first sample, has been infected before that sample was taken (FIG. 3d). A more informative analysis is one that includes only women with incident HPV infections, for whom the date of onset of infection is readily available. However, even this study design does not guarantee clarity. For example, a study that included women with

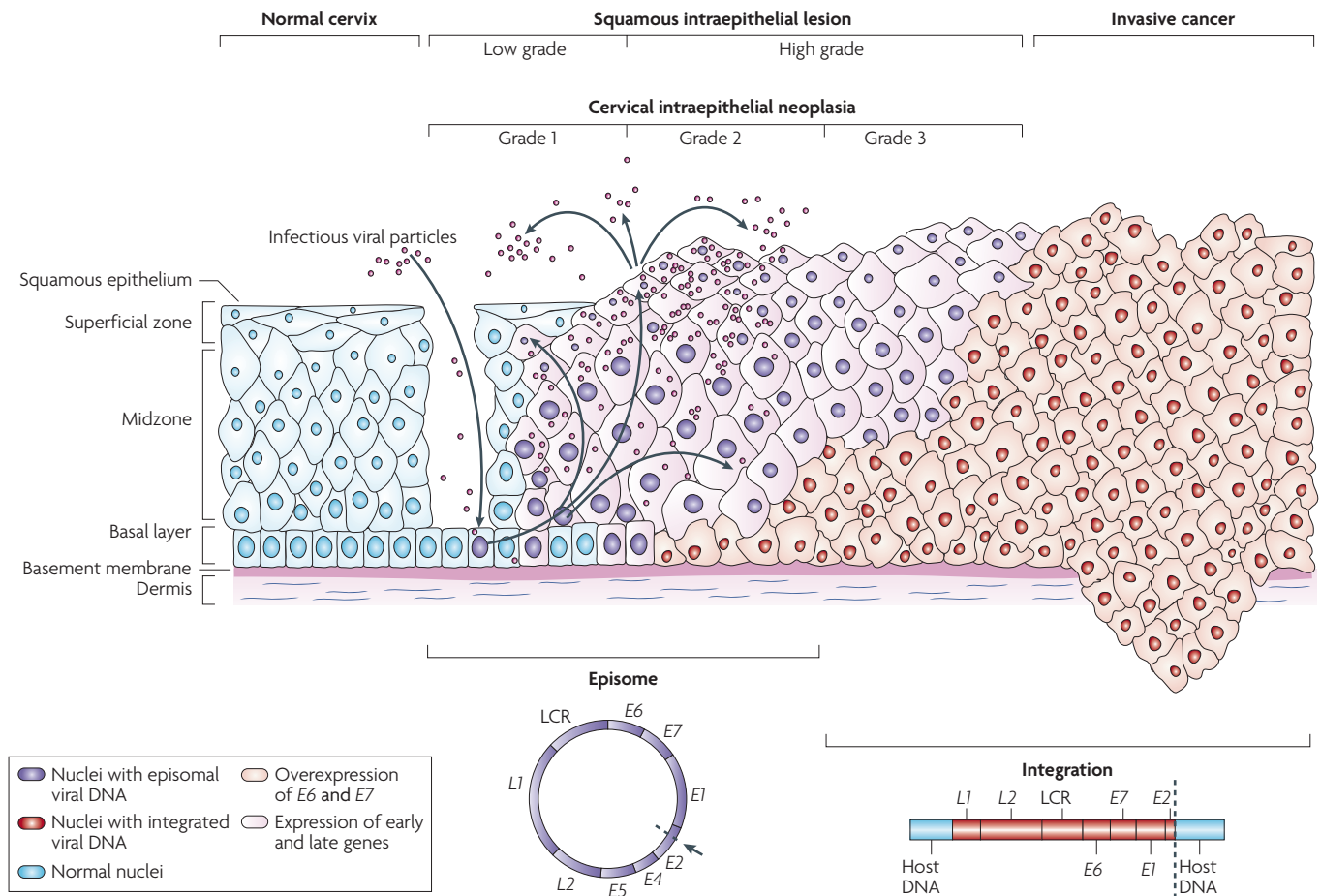


Figure 1 | HPV-mediated progression to cervical cancer. Basal cells in the cervical epithelium rest on the basement membrane, which is supported by the dermis. Human papillomavirus (HPV) is thought to access the basal cells through micro-abrasions in the cervical epithelium. Following infection, the early HPV genes *E1*, *E2*, *E4*, *E5*, *E6* and *E7* are expressed and the viral DNA replicates from episomal DNA (purple nuclei). In the upper layers of epithelium (the midzone and superficial zone) the viral genome is replicated further, and the late genes *L1* and *L2*, and *E4* are expressed. *L1* and *L2* encapsidate the viral genomes to form progeny virions in the nucleus. The shed virus can then initiate a new infection. Low-grade intraepithelial lesions support productive viral replication. An unknown number of high-risk HPV infections progress to high-grade intraepithelial neoplasia (HGGIN). The progression of untreated lesions to microinvasive and invasive cancer is associated with the integration of the HPV genome into the host chromosomes (red nuclei), with associated loss or disruption of *E2*, and subsequent upregulation of *E6* and *E7* oncogene expression. LCR, long control region.

incident HPV infections defined a persistent infection as an infection of 6 months or more. This was less than the median duration of HPV infection in all cohort members, thereby classifying most women as having a persistent infection²⁹. Although several investigators have reported the distribution of exposure times necessary for disease to occur^{9,50}, some still characterize exposure levels according to the number of positive tests. This has led to apparently conflicting interpretations of the same data; the shorter the interval between tests, the more likely an infection will be defined as persistent^{26,30,32,51} (FIG. 3e).

Many studies that have concluded that a persistent infection is a prerequisite for the development of cervical neoplasia can also be criticized on methodological grounds, because the sample taken at diagnosis is used to provide the second of the two consecutive positive

samples that are necessary to define an infection as persistent^{30,32,33} (FIG. 3f). Observations on exposure status made at or after the time of diagnosis are uninformative with respect to determining the exposure necessary for that disease to occur; an outcome cannot be attributed to a given level of exposure until that period of exposure has been completed⁵².

It is now clear that epithelial abnormalities of the cervix can be evident shortly after the first detection of HPV^{9,53}. This is not to deny that HPV can establish a persistent infection, nor that a persistent infection is necessary for the development of invasive disease. However, it remains to be determined whether persistent infections are characterized by the continuing detection of HPV, or by a state of latency during which the virus remains undetectable only to reappear later. For example, one longitudinal study shows that HPV DNA can only

Latent infection

The persistence of an infection in a host without symptoms and/or without being detectable.

Incident

A new occurrence of disease or infection in someone previously free of disease or infection.

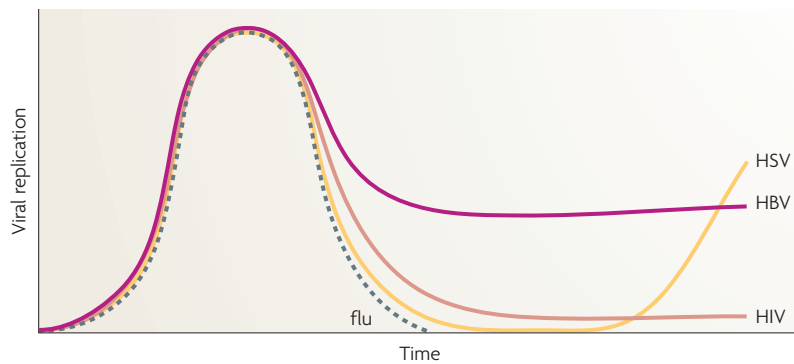


Figure 2 | General patterns of infection. To illustrate different patterns of persistent infection, the replication of herpes simplex virus (HSV), hepatitis B virus (HBV) and human immunodeficiency virus (HIV) is plotted as a function of time after infection. Acute infection (dashed grey line) is associated with clinical symptoms and the release of an infectious virus, such as influenza virus (flu). Persistent infection is associated with the production of an infectious virus, HBV for example (magenta line), for the lifetime of the host. Latent infection as seen in HSV infection (orange line) is a variation of persistent infection in which the acute infection is followed by a quiescent phase in which the virus productive cycle is absent or significantly reduced. The viral genome remains in a ‘silent’ state but can be intermittently reactivated into bouts of productive infection. Slow virus infection, as seen with HIV (pink line) is yet another version of persistent infection typified by long periods (years) between primary infection and the development of fatal symptoms: production of an infectious virus is either continuous at a low level or absent until failing immune control results in overwhelming virus production.

be detected in cytologically normal samples for a short time before cervical cancer is diagnosed³². A mechanism for latency has not been established so far, nor is it clear whether the differences between a latent and active cervical infection are qualitative or quantitative. However, a woman cannot be defined as having a persistent infection in any meaningful virological sense just because she tests positive for HPV on two occasions, some months apart, and therefore she should not, on the basis of this evidence, be considered to have a higher risk of cervical cancer. Nor can a woman, who tests positive for a specific HPV type, be considered to have cleared her infection when she first tests negative for that type. A clearer understanding of these issues is essential for the effective implementation of screening strategies, which include HPV testing.

HPV viral load

Among women who test positive for high-risk HPV types, cytological abnormality is more common in those with a high viral load^{49,54–58}. The consistency of this finding, irrespective of the method used to measure viral load, persuaded many that the inclusion of such a measurement could improve the effectiveness of HPV-based screening and triage strategies. However, it is now clear that the relationship between viral load and disease is more complex than was previously thought. Whereas many cross-sectional studies reported an increase in viral load with increasing disease severity, others found either no association, or a higher viral load in women with low-grade squamous intraepithelial lesion (LSIL) than in those with high-grade squamous intraepithelial lesion (HSIL)^{49,54,55,57,58}

(TABLE 1; references 1–68 from **Supplementary information S1** (table)). Longitudinal studies have also failed to find a consistent association between a baseline measurement of viral load and duration of infection, clearance of disease, and subsequent risk of acquisition or progression of disease^{33,42,59}.

There are several possible explanations for these inconsistencies. Whereas the prevalence of integrated forms of HPV increases with increasing disease severity, integration itself is followed by a decrease in viral load. However, in almost all cross-sectional and longitudinal surveys that have measured viral load, integration status is undefined and might, of course, change over time^{60–63}. In women with HGCIN, viral load is higher when low-grade CIN (LGCIN) is also present; in almost all studies, only the most severe histological abnormality is reported⁶⁴. The acquisition of new HPV types is associated with both changes in viral load and with the development of new CIN lesions; therefore, measures of association might be unreliable in longitudinal studies that rely on a single baseline measurement of exposure^{51,59}.

Finally, whereas many of the studies already cited used an assay that provided a measure of viral load aggregated across a panel of high-risk HPV types, it is now clear that the relationship between viral load and disease varies with HPV type. For example, cross-sectional studies show that HPV16 viral load increases with increasing disease severity, whereas that of HPV18 does not^{57,65,66}. If the cytopathic effect observed in exfoliated cervical cells is a reflection of the viral load, then this might explain why the cytological changes detected after HPV18 infection underestimate the severity of the underlying histological abnormality, unlike those detected after HPV16 infection⁶⁷. This is an important consideration because the benefits of cervical screening follow from the detection, investigation and treatment of epithelial abnormality, with the decision to refer for assessment by colposcopy usually based on the severity of the cytological abnormality. If HPV16 infection is more likely to be followed by a severe cytological abnormality than HPV18 infection, then screening programmes are more likely to interrupt the natural history of an HPV16 infection than that of an HPV18 infection. In many countries, screening programmes have not prevented an increase in the incidence of adenocarcinoma of the cervix, despite their success in preventing squamous cell carcinoma^{1,4,48,68}. This failure has been attributed to the inaccessibility of these lesions to cytological sampling, but could also be explained in part by the strong association of adenocarcinoma with HPV18 infection, which in turn is associated with only minor cytological changes^{22,23,67,69}. Although the initial optimism regarding the clinical value of HPV viral load testing now seems misplaced, robust measurements of type-specific viral load in samples in which the integration status is also defined, could provide useful insights into the natural history of HPV infections and their relationship to disease.

Longitudinal study

A study in which subjects are followed over a period of time so that the temporal sequence of potential cause and effect can be established.

Triage

The classification of patients according to priority of need.

Squamous intraepithelial lesion

(SIL). A disease characterized by the abnormal growth of squamous cells on the surface of the cervix. It is classified cytologically as low-grade (LSIL) or high-grade (HSIL) according to how much of the cervix is affected and how abnormal the cells are.

Colposcopy

The visual examination of the uterine cervix with a magnifying lens to detect abnormal cells.

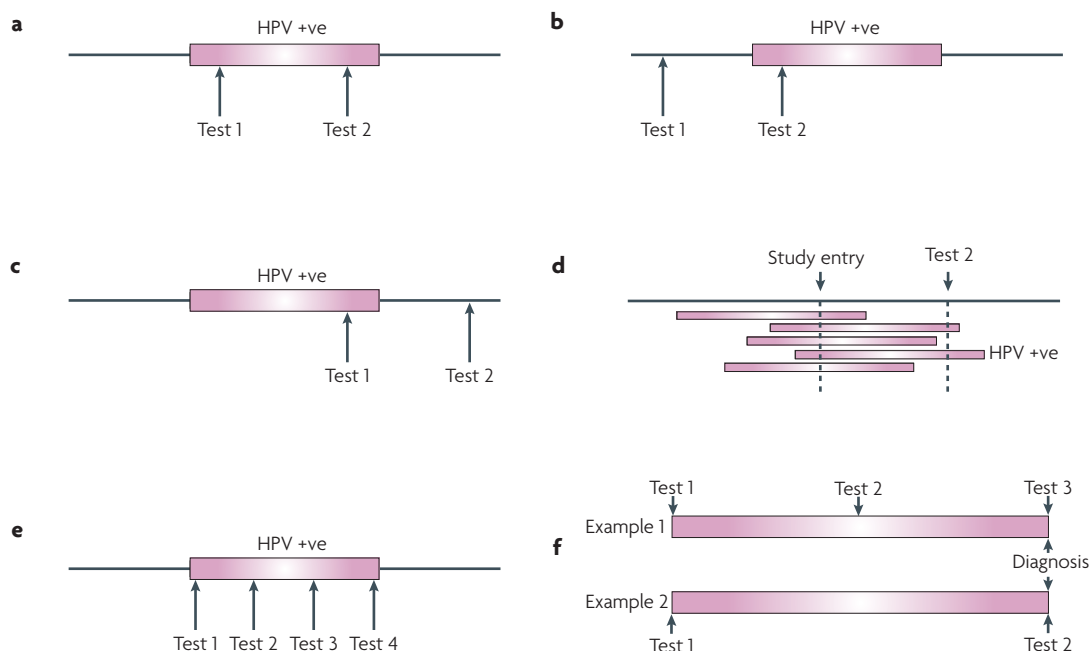


Figure 3 | Persistent HPV infection. The rectangles denote a period of time during which human papillomavirus (HPV) DNA sequences could be detected. In many studies, a woman is considered to have a persistent infection if she is HPV-DNA-positive in two or more consecutive tests, and a transient infection if she is positive only once. Panels **a–c** show how the same infection might be considered persistent or transient, depending on when the samples were taken. **a** | The infection is characterized as persistent. **b, c** | The same infection is now characterized as transient merely by changing the sampling times. **d** | These episodes of HPV infection are of identical duration, but began at different, undetermined times before entry to the study. Two of these infections would have been considered persistent, and three transient, on the basis of when the second test was performed. **e** | With a definition based on the number of positive tests, an infection is more likely to be considered persistent if a woman is tested more frequently. **f** | Example 1 fulfils the definition of a persistent infection as being positive in two or more consecutive tests and, in so far as it is defined before the detection of disease, it could be considered necessary for that disease to occur. Example 2 could also be considered a persistent infection, but the definition of a persistent infection was not fulfilled before the visit at which disease was detected, and therefore cannot be considered necessary for disease to occur.

HPV integration

Human papillomavirus can be found in cervical material in episomal forms, integrated forms or in mixed forms that contain both. Viral integration into the host-cell genome occurs downstream of the early genes *E6* and *E7*, often in the *E1* or *E2* region; this disruption results in a loss of negative-feedback control of oncogene expression by the viral regulatory *E2* protein (FIG. 1). Integrant-derived transcripts are more stable than those derived from episomal viral DNA, and HPV16 integration has been associated with a selective growth advantage for affected cells^{62,70}.

The prevalence in exfoliated cervical cells or cervical tissue of episomal or integrated forms of HPV, or both, varies with the severity of disease, the infecting HPV type and the method used to determine the physical state of the virus^{60,61,71–115}.

For example, HPV integration was once considered a late event in cervical carcinogenesis, because early studies rarely found integrated forms in women with CIN. However, when the physical state of the virus is determined by the failure to amplify full-length *E2* using PCR, the frequency of detection of integrated forms of HPV16 alone in women with CIN3 and in

those with invasive disease is the same, and is comparable to that found in women with invasive disease using Southern blot hybridization (TABLE 2). Because the *E2* approach cannot distinguish integrated forms in the presence of episomal forms, the Southern blot detection method provides the most robust estimate of the overall prevalence of integrated forms. However, it might fail to detect integrated forms in the presence of a low viral load, which is more likely to be found following viral integration.

Integrated forms have also been detected in disease-free women and in those with LGCIN using *E2* PCR, although much less frequently. Unlike HPV16, HPV18 integration seems virtually complete in women with CIN3 or invasive disease, in whom episomal forms are rarely detected. Recently, real-time PCR assays, which simultaneously measure the HPV *E2* and *E6* copy numbers, have been used to determine the integration status. Although this assay seems ideally suited to exfoliated cervical cells and small cervical biopsy specimens, reconstitution experiments using different ratios of episomal and integrated forms of HPV16 show that, using this method, integrated forms can only be distinguished

Table 1 | Association between high viral load and cervical neoplasia

Outcome variable	Positive association (number of studies)	
	Yes	No
Longitudinal study design		
Duration of HPV infection*	4	4
Acquisition of disease	8	4
Duration of disease	3	2
Progression of disease	3	2
Cross-sectional study design		
Presence of disease	29	1
Severity of disease	17	25
HSIL > LSIL	14 [‡]	–
LSIL > HSIL	6 [‡]	–

*Based on a single baseline measurement. †In two studies, the exposure–disease relationship varied with HPV type. This table was compiled from references 1–68 in [Supplementary information S1](#). HPV, human papillomavirus; HSIL, high-grade squamous intraepithelial lesion; LSIL, low-grade squamous intraepithelial lesion.

in the presence of episomal forms when they are at least 100 times as common¹¹⁰. Furthermore, the size of the *E2* product that can be detected using this method is constrained by the need to optimize amplification efficiency, and therefore it is not surprising that integrated forms alone are detected less frequently than when full-length *E2* is amplified.

In almost all cervical cancers examined, integration is found at only one chromosomal site in the tumour cells, consistent with the idea that cervical cancer is a clonal disease¹¹⁶. Whether the integration event itself is crucial to carcinogenesis is the subject of ongoing debate. There are two schools of thought and both agree that, in women with invasive disease, integration preferentially occurs at common fragile sites; however, they disagree as to whether these sites are otherwise randomly distributed throughout the genome.

One school of thought argues that the overrepresentation of fragile sites is due to their greater susceptibility to integration-induced chromosomal alterations^{117–123}; whereas the other argues that it is just a reflection of their greater accessibility^{124–126}. One highlights the clustering of integration sites in certain chromosomal regions; whereas the other points to the absence of a specific cellular sequence motif in the transition sequence between the viral and cellular genome, or of recurrent integrations in a given region at the same locus. One emphasizes the number of cellular genes that are disrupted by HPV integration and that have been linked to carcinogenesis at other sites; whereas the other emphasizes the lack of evidence linking changes in expression of many of these genes to an integration event.

Although it is reasonable to suspect that integration will preferentially occur at those sites that confer a growth or survival advantage on the affected cells, the sites of HPV integration into the human genome have only been systematically studied in women with invasive cancer. There is no evidence that the location of integration sites in women who progressed to

invasive disease differs from those who did not. So far, the most compelling evidence for insertional mutagenesis is provided by a report that describes transcriptional and protein overexpression of the proto-oncogene *MYC* in five cervical cell lines with HPV16 or HPV18 DNA sequences integrated at 8q24, but not in four cell lines in which integration occurred at other sites¹²⁷. Integration at 8q24 has been found in approximately 10% of women with invasive disease and is reported to be more common in women with adenocarcinoma of the cervix and in those who test positive for HPV18 (REF. 120).

It has been proposed that the identification of integrated forms of HPV could be a useful biomarker for progressive disease. There are several problems with this proposal. First, the identification of small numbers of integrated forms in a background of mainly episomal forms will always be a technical challenge when only exfoliated cells are available for analysis. Second, if integrated genomes are often transcriptionally silent, or become so shortly after integration, then their detection might have limited prognostic usefulness. Although integrated forms are detected in over 40% of women with CIN3, active transcription of integrated forms has only been reported in 15% of such women^{93,128} (TABLE 2). The detection of integrant-derived transcripts could provide more useful prognostic information. However, it has been shown that cervical keratinocytes containing integrated forms will only emerge following a reduction in the number of *E2*-expressing episomes¹²⁹. This loss of episomal *E2*, which is associated with the endogenous activation of anti-viral genes and is accelerated by exogenous interferon- β (*IFN* β), increases the expression of viral oncogenes in cells containing integrated forms^{129–131}. Therefore, testing for the absence of *E2* in exfoliated cells, a technically much simpler proposition, merits evaluation as a marker of disease progression in HPV-positive women.

Multiple HPV types: competition or synergy?

Longitudinal studies in both men and women show that the concurrent or sequential detection of more than one HPV type is common, and that this occurs more frequently than would be expected by chance^{12–15,18,132,133}. If HPV types compete to colonize the cervical epithelium, then the prevalence of competing HPV types that are not targeted by vaccination could increase. Natural history studies offer no evidence of competition between HPV types, at least in so far as the risk of acquiring a new HPV type seems unrelated to a previous infection with another type^{14,15}. However, publication of the post-vaccination HPV-type-distributions in women who have already been vaccinated would provide more compelling evidence that this is the case. Natural history studies show an increased risk of acquisition of new HPV types in women already infected, compared with those who are HPV negative^{14,15}. For example, the risk of acquiring HPV58 is up to seven times higher in women with an incident HPV16 or HPV18 infection compared with those who are not infected with these types¹⁸.

Fragile site

A site on a chromosome that tends to break more often than other sites.

Insertional mutagenesis

The occurrence of a mutation that is caused by the introduction of foreign DNA sequences into a gene.

Keratinocyte

An epidermal cell that produces the protein keratin.

Table 2 | Frequency of HPV DNA forms detected according to HPV type, disease severity and detection method

Disease severity	Integrated forms alone					Episomal forms alone			Integrated ± episomal forms		
	Full-length E1 PCR	Full-length E2 PCR	SB ± 2D GEL	E2:E6 ratio	ISH	SB ± 2D GEL	E2:E6 ratio	ISH	SB ± 2D GEL	E2:E6 ratio	ISH
HPV16											
Invasive disease	28% (4; 267)	43% (14; 643)	49% (15; 469)	23% (3; 142)	67% (6; 147)	25% (15; 469)	17% (3; 142)	9% (6; 147)	66% (15; 469)	70% (3; 142)	91% (6; 147)
CIN3/HSIL	–	42% (7; 212)	23% (4; 90)	13% (8; 243)	32% (2; 34)	56% (4; 90)	32% (8; 243)	12% (2; 34)	44% (4; 90)	62% (8; 243)	88% (2; 34)
CIN2	–	0% (2; 44)	–	50% (2; 29)	–	–	15% (2; 29)	–	–	83% (2; 29)	–
CIN1/LSIL	–	5% (3; 98)	0% (1; 12)	6% (3; 54)	–	100% (1; 12)	29% (3; 54)	–	0% (1; 12)	72% (3; 54)	–
Disease-free	–	11% (2; 37)	–	11% (1; 127)	–	–	20% (1; 127)	–	–	80% (1; 127)	–
HPV18											
Invasive disease	–	100% (2; 53)	98% (3; 51)	50% (1; 10)	97% (2; 31)	0% (3; 51)	0% (1; 10)	0% (2; 31)	100% (3; 51)	100% (1; 10)	100% (2; 31)
CIN3/HSIL	–	–	94% (1; 17)	–	–	0% (1; 17)	–	–	100% (1; 17)	–	–

Pooled analysis of studies reporting HPV16 or HPV18 integration status in cervical material. Studies were only included if they contributed 10 or more cases to at least one disease category. Median values are presented as a percentage followed by the number of studies contributing to the estimate, and the total number of cases in brackets. Episomal and integrated forms can be distinguished by the presence or absence of HPV–human DNA junction fragments using Southern blot (SB) analysis; by their different mobility using two-dimensional gel electrophoresis (2D gel); by their pattern of nuclear signalling using *in situ* hybridization (ISH); or by the E2:E6 copy number ratio measured using real-time PCR. In real-time PCR, HPV is considered present in only episomal forms when E2 and E6 copy number are equivalent; in only integrated forms when E2 is absent; and in mixed forms when E2 is present but its copy number is less than that of E6. Failure to amplify full-length HPV E1 or E2 in the presence of HPV E6 or E7 is consistent with the presence of integrated forms in the absence of episomal forms. This assay cannot identify integrated forms in the presence of episomal forms. CIN, cervical intraepithelial neoplasia; HPV, human papillomavirus; HSIL, high-grade squamous intraepithelial lesion; LSIL, low-grade squamous intraepithelial lesion.

The sequential detection of new HPV types could be the result of their sequential acquisition from different sexual partners. Whereas the first detection of HPV is clearly associated with the acquisition of a new sexual partner, the evidence linking the detection of new HPV types with a change in sexual partner in women who are already infected is not compelling^{18,134,135}. Sequential transmission of several HPV types from the same partner cannot be excluded, but it seems unlikely. An alternative interpretation of these findings might be that more than one HPV type is transmitted simultaneously, and that their sequential detection after that is a consequence of replicative life cycles that are asynchronous and only occasionally overlapping. These life cycles might be interdependent. For example, in women with HSIL, HPV16 viral load is higher when other HPV types are present than when HPV16 is present alone¹³⁶. In women with invasive disease, episomal forms of HPV16 are significantly more common in those who test positive for both HPV16 and HPV18 than in those who test positive for HPV16 alone^{85,137}. If co-infection confers some mutual type-specific survival benefit, then the elimination of one HPV type could have an unexpected beneficial effect on the natural history of others.

Epigenetic changes in cervical neoplasia

Both viral and host genes can be targeted by the cellular DNA methylation machinery. The pattern of methylation of HPV genes varies with the viral life

cycle, the presence of disease and possibly the viral type^{138–141}. The *de novo* methylation of HPV DNA could be a host defence mechanism for suppressing transcription of foreign DNA or a strategy that the virus uses to maintain a long-term infection, or both^{142,143}. Treatment with the demethylating agent 5'-azacytidine has been shown to reactivate transcriptionally silent HPV integration sites in cervical cancer cell lines¹⁴⁴. This raises the possibility, for which there is no empirical evidence so far, that the detection of HPV in older women might follow the epigenetic reactivation of previously silenced integration sites.

Aberrant methylation of CpG islands in the promoter regions of tumour suppressor genes (TSGs) is one of several epigenetic changes that can contribute to carcinogenesis¹⁴⁵. TSG-promoter methylation varies qualitatively and quantitatively between Epstein–Barr virus (EBV)-positive and negative gastric cancers; hepatitis C-positive and negative hepatocellular carcinomas; and simian virus 40 (SV40)-positive and negative mesotheliomas^{146–157}. It also varies between HPV-positive and negative vulval cancers, and between HPV-positive and negative head and neck cancers, but such comparisons are non-informative in cervical cancers, almost all of which are HPV-positive at diagnosis^{158–161}. Cross-sectional comparisons have also failed to show HPV type-specific epigenetic changes, but have included only a small number of patients, and are likely to have been confounded by differences in histological type and disease severity.

CpG island

A region of genomic DNA in which the frequency of the CG sequence is higher than in other regions.

Table 3 | Epigenetic changes in cervical neoplasia

No.	Gene	Disease-free controls	LG CIN	HG CIN	Squamous cell carcinoma	Cervical carcinoma	Adeno-carcinoma
1	APC	18% (90; 4)	32% (37; 1)	34% (38; 1)	24% (238; 5)	32%* (88; 1)	54%* (65; 4)
2	CCNA1	0% (25; 1)	0% (13; 1)	36%* (11; 1)		93%* (30; 1)	
3	CDH1	0% (53; 4)	7% (42; 3)	22%* (60; 3)	61%* (170; 3)	47%* (135; 4)	33%* (57; 3)
4	CDKN2A	3% (254; 7)	12%* (120; 6)	29%* (237; 9)	32%* (407; 8)	22%* (372; 7)	20%* (110; 7)
5	DAPK1	1% (184; 5)	6% (69; 3)	30%* (88; 3)	64%* (299; 6)	52%* (180; 3)	39%* (89; 5)
6	HIC1	2% (43; 3)	52%* (54; 2)	70%* (91; 3)	20%* (108; 2)	71%* (79; 1)	63%* (27; 2)
7	IGSF4	0% (25; 3)	0% (29; 2)	39%* (31; 2)	58%* (52; 1)	65%* (23; 1)	
8	RARB	0% (47; 4)	5% (83; 3)	15%* (61; 3)	30%* (117; 3)	40%* (121; 3)	15%* (13; 2)
9	ROBO1	0% (51; 1)	7% (62; 1)	8%* (48; 1)		46%* (119; 1)	
10	SLIT1	0% (40; 1)	0% (48; 1)	10%* (39; 1)		53%* (119; 1)	
11	SLIT2	0% (51; 1)	2% (62; 1)	25%* (48; 1)		64%* (119; 1)	
12	FANCF	0% (18; 1)		0% (37; 1)		30%* (91; 1)	
13	FHIT	0% (50; 4)	3% (76; 2)	2% (63; 2)	12%* (77; 1)	24%* (189; 4)	0% (5; 1)
14	MGMT	3% (206; 6)	4% (93; 3)	7% (74; 3)	11%* (217; 4)	14%* (109; 2)	12%* (51; 3)
15	PTEN	0% (11; 1)	15% (27; 2)	0% (11; 1)	58%* (62; 1)		
16	RASSF1	3% (29; 3)	3% (58; 3)	1% (73; 3)	17%* (299; 7)	5% (110; 3)	26%* (132; 6)
17	SLIT3	0% (40; 1)	4% (48; 1)	2% (42; 1)		49%* (118; 1)	
18	TERT	0% (14; 1)	0% (13; 1)	0% (31; 1)		62%* (76; 2)	0% (9; 1)
19	TIMP3	0% (8; 1)	0% (13; 1)	16% (31; 1)		11% (171; 3)	55%* (38; 2)
20	C15orf48	0% (21; 1)				36%* (22; 1)	
21	MT1G	5% (21; 1)				55%* (22; 1)	
22	POU2F3	0% (7; 1)			41%* (32; 1)		36% (14; 1)
23	SFRP1	5% (21; 1)				58%* (22; 1)	
24	SPARC	5% (21; 1)				91%* (22; 1)	
25	TFPI2	38% (21; 1)				82%* (22; 1)	
26	TNFRSF10C	0% (12; 1)				100%* (50; 1)	
27	HSPA2		0% (13; 1)	3% (31; 1)		73%* (11; 1)	
28	SOCS1		0% (13; 1)	7% (31; 1)		50%* (11; 1)	
29	TWIST1		0% (23; 1)	14% (22; 1)		43%* (56; 1)	
30	SOCS2		23% (13; 1)	45% (31; 1)		64%* (11; 1)	
31	CDH13		5% (41; 2)	14% (63; 2)		46%* (89; 1)	

Pooled analysis reporting as a percentage the detection of specific methylated genes in cervical material. Listed are those genes for which methylated forms in tissue or exfoliated cells, taken from women with invasive disease, were significantly more frequent ($P < 0.05$) than in disease-free controls or when these were unavailable, women with low-grade cervical intraepithelial neoplasia (LGCIN). Genes numbered 1–11 are also more likely to be methylated in high-grade cervical intraepithelial neoplasia (HGICIN), whereas those numbered 12–19 are not. The presence of methylated forms in HGICIN and in disease-free control tissues has yet to be established for genes numbered 20–26 and 27–31, respectively. *Estimates that are significantly different from the baseline. The table includes the number of cases contributing to each estimate and the number of studies from which they are derived in brackets. 'Cervical carcinoma' refers to cases of invasive disease with unspecified histological type. This table was compiled from references 69–110 in [Supplementary information S1](#). APC, adenomatous polyposis coli; C15orf48, chromosome 15 open reading frame 48; CCNA1, cyclin A1; CDH1, E-cadherin; CDH13, H-cadherin; CDKN2A, cyclin-dependent kinase inhibitor 2A; DAPK1, death-associated protein kinase 1; FANCF, Fanconi anaemia, complementation group F; FHIT, fragile histidine triad gene; HIC1, hypermethylated in cancer 1; HSPA2, heat-shock 70kDa protein 2; IGSF4, immunoglobulin superfamily, member 4; MGMT, O-6-methylguanine-DNA methyltransferase; MT1G, metallothionein 1G; POU2F3, POU domain, class 2, transcription factor 3; PTEN, phosphatase and tensin homologue, mutated in multiple advanced cancers 1; RARB, retinoic acid receptor β ; RASSF1, Ras association (RalGDS/AF6) domain family 1; ROBO1, roundabout, axon guidance receptor, homologue 1; SFRP1, secreted frizzled-related protein 1; SLIT, slit homologue; SOCS, suppressor of cytokine signalling 1; SPARC, secreted protein, acidic, cysteine-rich (osteonectin); TERT, telomerase reverse transcriptase; TFPI2, tissue factor pathway inhibitor 2; TIMP3, TIMP metalloproteinase inhibitor 3; TNFRSF10C, tumour necrosis factor receptor superfamily, member 10c, decoy without an intracellular domain; TWIST1, twist homologue 1.

The DNA methyltransferases (**DNMT1**, **DNMT3A** and **DNMT3B**) are responsible for the initiation and maintenance of methylation, and are overexpressed in several solid and haematological malignancies¹⁴⁵. Gene expression profiling shows increased expression of DNMT1 in short-term primary cervical cancer cultures, compared with normal cervical keratinocyte cultures. Genome-wide microarray-based comparative genomic hybridization shows an increase in DNMT3B copy number that is associated with changes in mRNA expression in all HPV-immortalized cell lines, and in most cervical cancers analysed^{162,163}.

Viral oncogenes can induce TSG methylation following activation of DNA methyltransferases. For example, the TSG cadherin 1 (**CDH1**) is methylated following activation of DNA methyltransferases by the EBV oncogene latent membrane protein 1 (**LMP1**), and by the hepatitis B oncogene **HBxAg** (hepatitis B virus X antigen). DNMT1 is activated by LMP1 through the AP1–JNK (activated protein 1–JUN N-terminal kinase) signalling pathway^{164–168}. Although there is no evidence so far for HPV-induced methylation of TSGs, HPV16 E7 has been shown to bind DNMT1, and to stimulate its enzymatic activity¹⁶⁹. Moreover, similar to the BK virus large T antigen and the adenovirus E1A protein, E7 might also activate transcription of **DNMT1**. **DNMT1** is a target of the transcription factor E2F1 and E7 can inactivate members of the pocket protein family that inhibit E2F family members. Therefore E7 could stimulate E2F activity and **DNMT1** transcription^{170–172}.

TABLE 3 lists those genes that are commonly found in methylated forms in exfoliated cervical cells or cervical tissue¹⁵² (see also references 69–110 in **Supplementary information S1**). For some genes, the prevalence of methylated forms increases with disease severity; for others, methylated forms are only detected in women with invasive disease. For some of those genes, an association with HPV and cervical carcinogenesis has already been described. For example, **SFRP1** (secreted frizzled-related protein 1) is a negative regulator of the Wnt signalling pathway, the activation of which is required for the transformation of HPV-expressing human keratinocytes; and methylation of **SFRP4**, another negative regulator of this pathway, is associated with the detection of HPV16 in head and neck cancers^{161,173}.

Other TSGs have been associated with HPV and carcinogenesis. **SOC1** (suppressor of cytokine signalling 1) interacts with HPV16 E7 protein and induces its ubiquitylation and degradation; **PTEN** (phosphatase and tensin homologue) inactivates **STAT3** (signal transducer and activator of transcription 3) in HPV-infected cells; and **POU2F3** (POU domain, class 2, transcription factor 3) might be involved in differentiation-dependent regulation of HPV transcription^{174–177}. A screening test based on the detection of methylated forms of any one of these TSGs would lack sensitivity. However, the remarkable absence of methylated forms from disease-free women indicates that a test including more than one TSG, could improve sensitivity without the usual loss of specificity that follows the application of screening tests in parallel. Preliminary

studies using this strategy so far have confirmed the impression of high specificity without achieving the required sensitivity^{178,179}.

Future directions

A clearer understanding of HPV persistence is essential. Ultimately, cohort studies with long-term follow-up, using sensitive measures of type-specific viral load, will determine the incidence of true persistent HPV infections, the characteristics of that type of infection and its contribution to carcinogenesis. A more immediate goal would be to determine how often, in which epithelial cell subpopulations and in what form, HPV DNA sequences can be detected in cervical tissue after they can no longer be detected in cytological samples. Also of interest are age-related changes in the patterns of viral genome methylation, given the possibility that HPV persists in an epigenetically regulated latent state.

There has been a trend towards conflating all HPV types associated with cervical neoplasia, with a view to establishing whether testing for the presence of any one of a panel of high-risk types improves the effectiveness of cervical screening programmes. This has distracted attention from type-specific differences in the natural history of HPV infection that affect the exposure–disease relationship. For example, it will be important to determine for different HPV types how often and how soon viral integration follows an incident HPV infection; for how long following an integration event can integrant-derived transcripts be detected; and, once silenced, how often are transcription centres reactivated.

The contribution of TSG methylation to the initiation or progression of CIN has not been defined. Indeed, for no site of cancer has the risk of incident disease in disease-free individuals, or the risk of disease progression in those who already have pre-malignant changes, been related to the presence of TSG methylation in baseline material. Nor has the extent to which epigenetic changes explain the role of known carcinogens in the oncogenic process, been adequately explored. For example, the successful eradication from the stomach of another infectious agent, *Helicobacter pylori*, is followed by a reduction in the methylation density of **CDH1**, a TSG inactivated during gastric carcinogenesis¹⁸⁰. Only longitudinal studies in women with incident HPV infections can distinguish those epigenetic events directly attributable to HPV infection or to co-factors such as smoking¹⁶¹, from those secondary to the disease process. Longitudinal studies in women with prevalent HPV infections will show how often epigenetic changes persist after HPV DNA sequences can no longer be detected, or after the successful treatment of CIN. The ease with which the cervix can be sampled, unlike almost all other sites of cancer, makes it an ideal model to explore how progression to a pre-malignant phenotype *in vivo* can be explained by genetic and epigenetic changes, and how these changes are linked to the acquisition of an oncogenic virus. The insights provided could be relevant to other models of viral carcinogenesis.

DNA methyltransferase

An enzyme that transfers a methyl group to DNA. DNMT1 is the most abundant methyltransferase and is the main maintenance methyltransferase. DNMT3A and DNMT3B are the main *de novo* methyltransferases.

Pocket protein

The pocket protein family includes three proteins, RB (retinoblastoma), RBL1 (retinoblastoma-like 1) and RBL2. They have a crucial role in cell-cycle regulation through interaction with the E2F transcription factors.

1. Parkin, D. M. & Bray, F. Chapter 2: the burden of HPV-related cancers. *Vaccine* **24** (Suppl. 3), S11–S25 (2006).
2. World Health Organization. International Agency for Research on Cancer. GLOBOCAN database 2002. *CANCERmondial* [online]. <http://www-dep.iarc.fr/> (2002).
3. Wang, S. S., Sherman, M. E., Hildesheim, A., Lacey, J. V. Jr & Devesa, S. Cervical adenocarcinoma and squamous cell carcinoma incidence trends among white women and black women in the United States for 1976–2000. *Cancer* **100**, 1035–1044 (2004).
4. Bray, F. *et al.* Incidence trends of adenocarcinoma of the cervix in 13 European countries. *Cancer Epidemiol. Biomarkers Prev.* **14**, 2191–2199 (2005).
5. de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H. U. & zur Hausen, H. Classification of papillomaviruses. *Virology* **324**, 17–27 (2004).
6. Collins, S. *et al.* High incidence of cervical human papillomavirus infection in women during their first sexual relationship. *BJOG* **109**, 96–98 (2002).
7. Castle, P. E. *et al.* A prospective study of age trends in cervical human papillomavirus acquisition and persistence in Guanacaste, Costa Rica. *J. Infect. Dis.* **191**, 1808–1816 (2005).
8. Franceschi, S. *et al.* Variations in the age-specific curves of human papillomavirus prevalence in women worldwide. *Int. J. Cancer* **119**, 2677–2684 (2006). **This is a cross-sectional study of sexually active women from the general population of 15 areas in four continents; it describes age-specific curves of HPV prevalence, and draws attention to the finding in some societies of a second peak of HPV infection in older women.**
9. Woodman, C. B. *et al.* Natural history of cervical human papillomavirus infection in young women: a longitudinal cohort study. *Lancet* **357**, 1831–1836 (2001).
10. de Sanjose, S. *et al.* Clearance of HPV infection in middle aged men and women after 9 years' follow up. *Sex. Transm. Infect.* **79**, 348 (2003).
11. Richardson, H. *et al.* The natural history of type-specific human papillomavirus infections in female university students. *Cancer Epidemiol. Biomarkers Prev.* **12**, 485–490 (2003).
12. Ho, G. Y. *et al.* Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. *J. Natl Cancer Inst.* **87**, 1365–1371 (1995).
13. Thomas, K. K. *et al.* Concurrent and sequential acquisition of different genital human papillomavirus types. *J. Infect. Dis.* **182**, 1097–1102 (2000).
14. Liaw, K. L. *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.* **183**, 8–15 (2001).
15. Rousseau, M. C. *et al.* Cervical coinfection with human papillomavirus (HPV) types as a predictor of acquisition and persistence of HPV infection. *J. Infect. Dis.* **184**, 1508–1517 (2001).
16. Ho, G. Y. *et al.* Risk factors for subsequent cervicovaginal human papillomavirus (HPV) infection and the protective role of antibodies to HPV-16 virus-like particles. *J. Infect. Dis.* **186**, 737–742 (2002).
17. Brown, D. R. *et al.* A longitudinal study of genital human papillomavirus infection in a cohort of closely followed adolescent women. *J. Infect. Dis.* **191**, 182–192 (2005).
18. Mendez, F. *et al.* Cervical coinfection with human papillomavirus (HPV) types and possible implications for the prevention of cervical cancer by HPV vaccines. *J. Infect. Dis.* **192**, 1158–1165 (2005). **This study includes women with incident infections and shows that the concurrent acquisition of multiple HPV types occurs more frequently than would be expected by chance.**
19. zur Hausen, H. Papillomaviruses and cancer: from basic studies to clinical application. *Nature Rev. Cancer* **2**, 342–350 (2002). **This review details the contribution of HPV to the pathogenesis of cervical neoplasia.**
20. Roden, R. & Wu, T. C. How will HPV vaccines affect cervical cancer? *Nature Rev. Cancer* **6**, 753–763 (2006). **This recent review considers the ramifications of prophylactic HPV vaccines.**
21. Doorbar, J. Molecular biology of human papillomavirus infection and cervical cancer. *Clin. Sci.* **110**, 525–541 (2006).
22. Clifford, G. M., Smith, J. S., Plummer, M., Munoz, N. & Franceschi, S. Human papillomavirus types in invasive cervical cancer worldwide: a meta-analysis. *Br. J. Cancer* **88**, 63–73 (2003).
23. Munoz, N. *et al.* Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N. Engl. J. Med.* **348**, 518–527 (2003).
24. Koutsky, L. A. & Harper, D. M. Chapter 13: current findings from prophylactic HPV vaccine trials. *Vaccine* **24** (Suppl. 3), S114–S121 (2006).
25. Tyler, K. and Nathanson, N. in *Fields Virology* 4th edn (eds Fields BN, Knipe, D.M., and Howley, P.M.), 220–222 (Lippincott Williams & Wilkins, Philadelphia, 2001). **This chapter reviews the characteristics of persistent viral infections.**
26. Koutsky, L. A. *et al.* A cohort study of the risk of cervical intraepithelial neoplasia grade 2 or 3 in relation to papillomavirus infection. *N. Engl. J. Med.* **327**, 1272–1278 (1992).
27. Evander, M. *et al.* Human papillomavirus infection is transient in young women: a population-based cohort study. *J. Infect. Dis.* **171**, 1026–1030 (1995).
28. Giuliano, A. R. *et al.* Antioxidant nutrients: associations with persistent human papillomavirus infection. *Cancer Epidemiol. Biomarkers Prev.* **6**, 917–923 (1997).
29. Ho, G. Y., Bierman, R., Beardsley, L., Chang, C. J. & Burk, R. D. Natural history of cervicovaginal papillomavirus infection in young women. *N. Engl. J. Med.* **338**, 423–428 (1998).
30. Liaw, K. L. *et al.* Detection of human papillomavirus DNA in cytologically normal women and subsequent cervical squamous intraepithelial lesions. *J. Natl Cancer Inst.* **91**, 954–960 (1999).
31. Nobbenhuis, M. A. *et al.* Relation of human papillomavirus status to cervical lesions and consequences for cervical-cancer screening: a prospective study. *Lancet* **354**, 20–25 (1999).
32. Wallin, K. L. *et al.* Type-specific persistence of human papillomavirus DNA before the development of invasive cervical cancer. *N. Engl. J. Med.* **341**, 1633–1638 (1999). **This article shows that a single positive finding of HPV DNA in a Pap smear confers increased risk of future invasive cervical cancer that is positive for the same type of virus as identified earlier; it also shows that HPV DNA can only be detected in cytologically normal samples for a short time before cervical cancer is diagnosed.**
33. Clavel, C. *et al.* Human papillomavirus detection by the hybrid capture II assay: a reliable test to select women with normal cervical smears at risk for developing cervical lesions. *Diagn. Mol. Pathol.* **9**, 145–150 (2000).
34. Elfegren, K., Kalantari, M., Moberger, B., Hagmar, B. & Dillner, J. A population-based five-year follow-up study of cervical human papillomavirus infection. *Am. J. Obstet. Gynecol.* **183**, 561–567 (2000).
35. Hopman, E. H. *et al.* High risk human papillomavirus in women with normal cervical cytology prior to the development of abnormal cytology and colposcopy. *BJOG* **107**, 600–604 (2000).
36. Ylitalo, N. *et al.* A prospective study showing long-term infection with human papillomavirus 16 before the development of cervical carcinoma *in situ*. *Cancer Res.* **60**, 6027–6032 (2000).
37. Ahdieh, L. *et al.* Prevalence, incidence, and type-specific persistence of human papillomavirus in human immunodeficiency virus (HIV)-positive and HIV-negative women. *J. Infect. Dis.* **184**, 682–690 (2001).
38. Ghaderi, M. *et al.* Risk of invasive cervical cancer associated with polymorphic HLA DR/DQ haplotypes. *Int. J. Cancer* **100**, 698–701 (2002).
39. Kjaer, S. K. *et al.* Type specific persistence of high risk human papillomavirus (HPV) as indicator of high grade cervical squamous intraepithelial lesions in young women: population based prospective follow up study. *BMJ* **325**, 572 (2002).
40. Sedjo, R. L. *et al.* Human papillomavirus persistence and nutrients involved in the methylation pathway among a cohort of young women. *Cancer Epidemiol. Biomarkers Prev.* **11**, 353–359 (2002).
41. van der Graaf, Y. *et al.* Human papillomavirus and the long-term risk of cervical neoplasia. *Am. J. Epidemiol.* **156**, 158–164 (2002).
42. van Duin, M. *et al.* Human papillomavirus 16 load in normal and abnormal cervical scrapes: an indicator of CIN II/III and viral clearance. *Int. J. Cancer* **98**, 590–595 (2002).
43. Chan, J. K. *et al.* HPV infection and number of lifetime sexual partners are strong predictors for 'natural' regression of CIN 2 and 3. *Br. J. Cancer* **89**, 1062–1066 (2003).
44. Giuliano, A. R. *et al.* Dietary intake and risk of persistent human papillomavirus (HPV) infection: the Ludwig–McGill HPV natural history study. *J. Infect. Dis.* **188**, 1508–1516 (2003).
45. Cuschieri, K. S., Whitley, M. J. & Cubie, H. A. Human papillomavirus type specific DNA and RNA persistence—implications for cervical disease progression and monitoring. *J. Med. Virol.* **73**, 65–70 (2004).
46. Peto, J. *et al.* Cervical HPV infection and neoplasia in a large population-based prospective study: the Manchester cohort. *Br. J. Cancer* **91**, 942–953 (2004).
47. Piyathilake, C. J. *et al.* Folate is associated with the natural history of high-risk human papillomaviruses. *Cancer Res.* **64**, 8788–8793 (2004).
48. Wang, S. S. *et al.* Validation of p16INK4a as a marker of oncogenic human papillomavirus infection in cervical biopsies from a population-based cohort in Costa Rica. *Cancer Epidemiol. Biomarkers Prev.* **13**, 1355–1360 (2004).
49. Lillo, F. B. *et al.* Determination of human papillomavirus (HPV) load and type in high-grade cervical lesions surgically resected from HIV-infected women during follow-up of HPV infection. *Clin. Infect. Dis.* **40**, 451–457 (2005).
50. Moscicki, A. B. *et al.* Risks for incident human papillomavirus infection and low-grade squamous intraepithelial lesion development in young females. *JAMA* **285**, 2995–3002 (2001).
51. Woodman, C. B. & Collins, S. A critique of cohort studies examining the role of human papillomavirus infection in cervical neoplasia. *BJOG* **109**, 1311–1318 (2002).
52. Rothman, K. J. & Greenland, S. in *Modern Epidemiology* 2nd edn (eds Rothman, K.J., Greenland, S.), 79–91 (Lippincott–Raven, Philadelphia, 1998).
53. Winer, R. L. *et al.* Genital human papillomavirus infection: incidence and risk factors in a cohort of female university students. *Am. J. Epidemiol.* **157**, 218–226 (2003).
54. Hall, S. *et al.* Human papillomavirus DNA detection in cervical specimens by hybrid capture: correlation with cytologic and histologic diagnoses of squamous intraepithelial lesions of the cervix. *Gynecol. Oncol.* **62**, 353–359 (1996).
55. Nindl, I. *et al.* Human papillomavirus distribution in cervical tissues of different morphology as determined by hybrid capture assay and PCR. *Int. J. Gynecol. Pathol.* **16**, 197–204 (1997).
56. Ho, G. Y. *et al.* Viral characteristics of human papillomavirus infection and antioxidant levels as risk factors for cervical dysplasia. *Int. J. Cancer* **78**, 594–599 (1998).
57. Swan, D. C. *et al.* Human papillomavirus (HPV) DNA copy number is dependent on grade of cervical disease and HPV type. *J. Clin. Microbiol.* **37**, 1030–1034 (1999).
58. Heard, I. *et al.* Increased risk of cervical disease among human immunodeficiency virus-infected women with severe immunosuppression and high human papillomavirus load(1). *Obstet. Gynecol.* **96**, 403–409 (2000).
59. Crum, C. P. *et al.* Dynamics of human papillomavirus infection between biopsy and excision of cervical intraepithelial neoplasia: results from the ZYC101a protocol. *J. Infect. Dis.* **189**, 1348–1354 (2004).
60. Cricca, M. *et al.* HPV DNA patterns and disease implications in the follow-up of patients treated for HPV16 high-grade carcinoma *in situ*. *J. Med. Virol.* **78**, 494–500 (2006).
61. Berumen, J., Unger, E. R., Casas, L. & Figueroa, P. Amplification of human papillomavirus types 16 and 18 in invasive cervical cancer. *Hum. Pathol.* **26**, 676–681 (1995).
62. Jeon, S., Allen-Hoffmann, B. L. & Lambert, P. F. Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *J. Virol.* **69**, 2989–2997 (1995). **This is an important paper in that it is consistent with the hypothesis that integration might provide a selective growth advantage for the epithelial precursors of cervical cancer.**
63. Spartz, H., Lehr, E., Zhang, B., Roman, A. & Brown, D. R. Progression from productive infection to integration and oncogenic transformation in human papillomavirus type 59-immortalized foreskin keratinocytes. *Virology* **336**, 11–25 (2005).

64. Sherman, M. E. *et al.* Determinants of human papillomavirus load among women with histological cervical intraepithelial neoplasia 3: dominant impact of surrounding low-grade lesions. *Cancer Epidemiol. Biomarkers Prev.* **12**, 1038–1044 (2003).

65. Gravitt, P. E. *et al.* A comparison between real-time polymerase chain reaction and hybrid capture 2 for human papillomavirus DNA quantitation. *Cancer Epidemiol. Biomarkers Prev.* **12**, 477–484 (2003). **Using real-time PCR this study shows, how the relationship between viral load and disease varies with the infecting HPV type.**

66. Ho, C. M. *et al.* Detection and quantitation of human papillomavirus type 16, 18 and 52 DNA in the peripheral blood of cervical cancer patients. *Gynecol. Oncol.* **99**, 615–621 (2005).

67. Woodman, C. B. *et al.* Human papillomavirus type 18 and rapidly progressing cervical intraepithelial neoplasia. *Lancet* **361**, 40–43 (2003).

68. Mitchell, H., Medley, G., Gordon, I. & Giles, G. Cervical cytology reported as negative and risk of adenocarcinoma of the cervix: no strong evidence of benefit. *Br. J. Cancer* **71**, 894–897 (1995).

69. Kovacic, M. B. *et al.* Relationships of human papillomavirus type, qualitative viral load, and age with cytologic abnormality. *Cancer Res.* **66**, 10112–10119 (2006).

70. Jeon, S. & Lambert, P. F. 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. USA* **92**, 1654–1658 (1995).

71. Andersson, S. *et al.* Type distribution, viral load and integration status of high-risk human papillomaviruses in pre-stages of cervical cancer (CIN). *Br. J. Cancer* **92**, 2195–2200 (2005).

72. Peitsaro, P., Johansson, B. & Syrjänen, 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.* **40**, 886–891 (2002).

73. Choo, K. B., Pan, C. C. & Han, S. H. Integration of human papillomavirus type 16 into cellular DNA of cervical carcinoma: preferential deletion of the E2 gene and invariable retention of the long control region and the E6/E7 open reading frames. *Virology* **161**, 259–261 (1987).

74. Meanwell, C. A., Cox, M. F., Blackledge, G. & Maitland, N. J. HPV 16 DNA in normal and malignant cervical epithelium: implications for the aetiology and behaviour of cervical neoplasia. *Lancet* **1**, 703–707 (1987).

75. Fuchs, P. G., Girardi, F. & Pfister, H. Human papillomavirus 16 DNA in cervical cancers and in lymph nodes of cervical cancer patients: a diagnostic marker for early metastases? *Int. J. Cancer* **43**, 41–44 (1989).

76. Matsukura, T., Koi, S. & Sugase, M. Both episomal and integrated forms of human papillomavirus type 16 are involved in invasive cervical cancers. *Virology* **172**, 63–72 (1989).

77. Cooper, K., Herrington, C. S., Graham, A. K., Evans, M. F. & McGee, J. O. *In situ* evidence for HPV 16, 18, 33 integration in cervical squamous cell cancer in Britain and South Africa. *J. Clin. Pathol.* **44**, 406–409 (1991).

78. Cullen, A. P., Reid, R., Campion, M. & Lorincz, A. T. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasia. *J. Virol.* **65**, 606–612 (1991).

79. Cooper, K., Herrington, C. S., Lo, E. S., Evans, M. F. & McGee, J. O. Integration of human papillomavirus types 16 and 18 in cervical adenocarcinoma. *J. Clin. Pathol.* **45**, 382–384 (1992).

80. Das, B. C., Sharma, J. K., Gopalakrishna, V. & Luthra, U. K. Analysis by polymerase chain reaction of the physical state of human papillomavirus type 16 DNA in cervical preneoplastic and neoplastic lesions. *J. Gen. Virol.* **73**, 2327–2336 (1992).

81. Chen, C. M. *et al.* Analysis of deletion of the integrated human papillomavirus 16 sequence in cervical cancer: a rapid multiplex polymerase chain reaction approach. *J. Med. Virol.* **44**, 206–211 (1994).

82. Kristiansen, E., Jenkins, A. & Holm, R. Coexistence of episomal and integrated HPV16 DNA in squamous cell carcinoma of the cervix. *J. Clin. Pathol.* **47**, 253–256 (1994).

83. Daniel, B., Mukherjee, G., Seshadri, L., Vallikad, E. & Krishna, S. 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.* **76**, 2589–2593 (1995).

84. Unger, E. R. *et al.* Human papillomavirus and disease-free survival in FIGO stage Ib cervical cancer. *J. Infect. Dis.* **172**, 1184–1190 (1995).

85. Park, J. S. *et al.* Physical status and expression of HPV genes in cervical cancers. *Gynecol. Oncol.* **65**, 121–129 (1997).

86. 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.* **50**, 600–604 (1997).

87. Vernon, S. D., Unger, E. R., Miller, D. L., Lee, D. R. & Reeves, W. C. Association of human papillomavirus type 16 integration in the E2 gene with poor disease-free survival from cervical cancer. *Int. J. Cancer* **74**, 50–56 (1997).

88. Kalantari, M. *et al.* Disruption of the E1 and E2 reading frames of HPV 16 in cervical carcinoma is associated with poor prognosis. *Int. J. Gynecol. Pathol.* **17**, 146–153 (1998).

89. Casas, L. *et al.* Asian-american variants of human papillomavirus type 16 have extensive mutations in the E2 gene and are highly amplified in cervical carcinomas. *Int. J. Cancer* **83**, 449–455 (1999).

90. Corden, S. A., Sant-Cassia, L. J., Easton, A. J. & Morris, A. G. The integration of HPV-18 DNA in cervical carcinoma. *Mol. Pathol.* **52**, 275–282 (1999).

91. Eriksson, A., Herron, J. R., Yamada, T. & Wheeler, C. M. Human papillomavirus type 16 variant lineages characterized by nucleotide sequence analysis of the E5 coding segment and the E2 hinge region. *J. Gen. Virol.* **80**, 595–600 (1999).

92. Kim, C. J. *et al.* The antibody response to HPV proteins and the genomic state of HPVs in patients with cervical cancer. *Int. J. Gynecol. Cancer* **9**, 1–11 (1999).

93. Klaes, R. *et al.* Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes. *Cancer Res.* **59**, 6132–6136 (1999).

94. Yoshinouchi, M. *et al.* Analysis by multiplex PCR of the physical status of human papillomavirus type 16 DNA in cervical cancers. *J. Clin. Microbiol.* **37**, 3514–3517 (1999).

95. Graham, D. A. & Herrington, C. S. HPV-16 E2 gene disruption and sequence variation in CIN 3 lesions and invasive squamous cell carcinomas of the cervix: relation to numerical chromosome abnormalities. *Mol. Pathol.* **53**, 201–206 (2000).

96. Szarka, K. *et al.* Integration status of virus DNA and p53 codon 72 polymorphism in human papillomavirus type 16 positive cervical cancers. *Anticancer Res.* **20**, 2161–2167 (2000).

97. Giannoudis, A., Duijn, M., Snijders, P. J. & Herrington, C. S. Variation in the E2-binding domain of HPV 16 is associated with high-grade squamous intraepithelial lesions of the cervix. *Br. J. Cancer* **84**, 1058–1063 (2001).

98. Tonon, S. A. *et al.* Physical status of the E2 human papilloma virus 16 viral gene in cervical preneoplastic and neoplastic lesions. *J. Clin. Virol.* **21**, 129–134 (2001).

99. Evans, M. F., Mount, S. L., Beatty, B. G. & Cooper, K. Biotinyl-tyramide-based *in situ* hybridization signal patterns distinguish human papillomavirus type and grade of cervical intraepithelial neoplasia. *Mod. Pathol.* **15**, 1339–1347 (2002).

100. Nagao, S. *et al.* Rapid and sensitive detection of physical status of human papillomavirus type 16 DNA by quantitative real-time PCR. *J. Clin. Microbiol.* **40**, 863–867 (2002).

101. Watts, K. J., Thompson, C. H., Cossart, Y. E. & Rose, B. R. Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia. *Int. J. Cancer* **97**, 868–874 (2002).

102. do Horto dos Santos Oliveira, L., Rodrigues Ede, V., de Salles Lopes, A. P., Fernandez Ade, P. & Cavalcanti, S. M. HPV 16 detection in cervical lesions, physical state of viral DNA and changes in p53 gene. *Sao Paulo Med. J.* **121**, 67–71 (2003).

103. Gallo, G. *et al.* Study of viral integration of HPV-16 in young patients with LSIL. *J. Clin. Pathol.* **56**, 532–536 (2003).

104. Lukaszuk, K., Liss, J., Wozniak, I., Emerich, J. & Wojcikowski, C. Human papillomavirus type 16 status in cervical carcinoma cell DNA assayed by multiplex PCR. *J. Clin. Microbiol.* **41**, 608–612 (2003).

105. Alazawi, W. *et al.* Genomic imbalances in 70 snap-frozen cervical squamous intraepithelial lesions: associations with lesion grade, state of the HPV16 E2 gene and clinical outcome. *Br. J. Cancer* **91**, 2063–2070 (2004).

This is an interesting paper that shows a high prevalence of integrated forms in women with cervical intraepithelial neoplasia.

106. Hudelist, G. *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.* **92**, 873–880 (2004).

107. Sathish, N. *et al.* Human papillomavirus 16 E6/E7 transcript and E2 gene status in patients with cervical neoplasia. *Mol. Diagn.* **8**, 57–64 (2004).

108. Chen, Q. Y., Bian, M. L., Chen, Z. H. & Liu, J. Detection of integration status of human papillomavirus 16 in cervical precancerous lesions. *Zhonghua Yi Xue Za Zhi* **85**, 400–404 (2005).

109. Swan, D. C. *et al.* Human papillomavirus type 16 E2 and E6/E7 variants. *Gynecol. Oncol.* **96**, 695–700 (2005).

110. Arias-Pulido, H., Peyton, C. L., Joste, N. E., Vargas, H. & Wheeler, C. M. Human papillomavirus type 16 integration in cervical carcinoma *in situ* and in invasive cervical cancer. *J. Clin. Microbiol.* **44**, 1755–1762 (2006).

This paper is interesting in that it shows that real-time PCR assays can only distinguish integrated forms in the presence of episomal forms when they are at least 100 times as common.

111. Bhattacharjee, B. & Sengupta, S. HPV16 E2 gene disruption and polymorphisms of E2 and LCR: some significant associations with cervical cancer in Indian women. *Gynecol. Oncol.* **100**, 372–378 (2006).

112. Bryan, J. T. *et al.* Detection of specific human papillomavirus types in paraffin-embedded sections of cervical carcinomas. *J. Med. Virol.* **78**, 117–124 (2006).

113. Ho, C. M., Chien, T. Y., Huang, S. H., Lee, B. H. & Chang, S. F. Integrated human papillomavirus types 52 and 58 are infrequently found in cervical cancer, and high viral loads predict risk of cervical cancer. *Gynecol. Oncol.* **102**, 54–60 (2006).

114. Kulmala, S. M. *et al.* Early integration of high copy HPV16 detectable in women with normal and low grade cervical cytology and histology. *J. Clin. Pathol.* **59**, 513–517 (2006).

115. Zheng, Y., Peng, Z. L., Lou, J. Y. & Wang, H. Detection of physical state of human papillomavirus 16 in cervical cancer tissue and SiHa cell line by multiplex real-time polymerase chain reaction. *Ai Zheng* **25**, 373–377 (2006).

116. Vinokurova, S. *et al.* Clonal history of papillomavirus-induced dysplasia in the female lower genital tract. *J. Natl Cancer Inst.* **97**, 1816–1821 (2005).

117. Wilke, C. M. *et al.* FRA3B extends over a broad region and contains a spontaneous HPV16 integration site: direct evidence for the coincidence of viral integration sites and fragile sites. *Hum. Mol. Genet.* **5**, 187–195 (1996).

118. Thorland, E. C. *et al.* Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites. *Cancer Res.* **60**, 5916–5921 (2000).

119. Ferber, M. J. *et al.* Integrations of the hepatitis B virus (HBV) and human papillomavirus (HPV) into the human telomerase reverse transcriptase (hTERT) gene in liver and cervical cancers. *Oncogene* **22**, 3813–3820 (2003).

120. Ferber, M. J. *et al.* Preferential integration of human papillomavirus type 18 near the c-MYC locus in cervical carcinoma. *Oncogene* **22**, 7233–7242 (2003).

121. Thorland, E. C., Myers, S. L., Gostout, B. S. & Smith, D. I. Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. *Oncogene* **22**, 1225–1237 (2003).

122. Ferber, M. J. *et al.* Positioning of cervical carcinoma and Burkitt lymphoma translocation breakpoints with respect to the human papillomavirus integration cluster in FRA3C at 8q24.13. *Cancer Genet. Cytogenet.* **154**, 1–9 (2004).

123. Yu, T. *et al.* The role of viral integration in the development of cervical cancer. *Cancer Genet. Cytogenet.* **158**, 27–34 (2005).

124. Wentzensen, N. *et al.* Characterization of viral-cellular fusion transcripts in a large series of HPV16 and 18 positive anogenital lesions. *Oncogene* **21**, 419–426 (2002).

125. Ziegert, C. *et al.* A comprehensive analysis of HPV integration loci in anogenital lesions combining transcript and genome-based amplification techniques. *Oncogene* **22**, 3977–3984 (2003).

126. 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.* **64**, 3878–3884 (2004). **This is an important review of the sites of HPV integration into the human genome.**
127. Peter, M. *et al.* MYC activation associated with the integration of HPV DNA at the MYC locus in genital tumors. *Oncogene* **25**, 5985–5993 (2006). **This paper describes transcriptional and protein overexpression of MYC associated with HPV16 or HPV18 integration at 8q24, which includes the MYC locus.**
128. Melsheimer, P., Vinokurova, S., Wentzensen, N., Bastert, G. & von Knebel Doeberitz, M. DNA aneuploidy and integration of human papillomavirus type 16 E6/E7 oncogenes in intraepithelial neoplasia and invasive squamous cell carcinoma of the cervix uteri. *Clin. Cancer Res.* **10**, 3059–3063 (2004).
129. Pett, M. R. *et al.* Selection of cervical keratinocytes containing integrated HPV16 associates with episome loss and an endogenous antiviral response. *Proc. Natl Acad. Sci. USA* **103**, 3822–3827 (2006). **This paper offers an interesting hypothesis: it proposes that it is the loss of episomal E2 in cells in which integration has already occurred that confers the growth advantage on selected cells.**
130. Bechtold, V., Beard, P. & Raj, K. Human papillomavirus type 16 E2 protein has no effect on transcription from from episomal viral DNA. *J. Virol.* **77**, 2021–2028 (2003).
131. Herdman, M. T. *et al.* Interferon- β treatment of cervical keratinocytes naturally infected with human papillomavirus 16 episomes promotes rapid reduction in episome numbers and emergence of latent integrants. *Carcinogenesis* **27**, 2341–2353 (2006).
132. Kjaer, S. K. *et al.* Acquisition and persistence of human papillomavirus infection in younger men: a prospective follow-up study among Danish soldiers. *Cancer Epidemiol. Biomarkers Prev.* **14**, 1528–1533 (2005).
133. Lajous, M. *et al.* Determinants of prevalence, acquisition, and persistence of human papillomavirus in healthy mexican military men. *Cancer Epidemiol. Biomarkers Prev.* **14**, 1710–1716 (2005).
134. Giuliano, A. R., Papenfuss, M., Abrahamson, M. & Insera, P. Differences in factors associated with oncogenic and nononcogenic human papillomavirus infection at the United States-Mexico border. *Cancer Epidemiol. Biomarkers Prev.* **11**, 930–934 (2002).
135. Rousseau, M. C. *et al.* Predictors of cervical coinfection with multiple human papillomavirus types. *Cancer Epidemiol. Biomarkers Prev.* **12**, 1029–1037 (2003).
136. Weissenborn, S. J. *et al.* Oncogenic human papillomavirus DNA loads in human immunodeficiency virus-positive women with high-grade cervical lesions are strongly elevated. *J. Clin. Microbiol.* **41**, 2763–2767 (2003).
137. Badaracco, G. & Venuti, A. Physical status of HPV types 16 and 18 in topographically different areas of genital tumours and in paired tumour-free mucosa. *Int. J. Oncol.* **27**, 161–167 (2005). **This paper indicates that the integration status of HPV16 might vary with the presence or absence of HPV18.**
138. Kim, K., Garner-Hamrick, P. A., Fisher, C., Lee, D. & Lambert, P. F. Methylation patterns of papillomavirus DNA, its influence on E2 function, and implications in viral infection. *J. Virol.* **77**, 12450–12459 (2003).
139. Badal, S. *et al.* The human papillomavirus-18 genome is efficiently targeted by cellular DNA methylation. *Virology* **324**, 483–492 (2004).
140. Kalantari, M. *et al.* Conserved methylation patterns of human papillomavirus type 16 DNA in asymptomatic infection and cervical neoplasia. *J. Virol.* **78**, 12762–12772 (2004).
141. Turan, T. *et al.* Methylation of the human papillomavirus-18 L1 gene: a biomarker of neoplastic progression? *Virology* **349**, 175–183 (2006).
142. Remus, R. *et al.* Insertion of foreign DNA into an established mammalian genome can alter the methylation of cellular DNA sequences. *J. Virol.* **73**, 1010–1022 (1999).
143. Badal, V. *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.* **77**, 6227–6234 (2003).
144. Van Tine, B. A. *et al.* Clonal selection for transcriptionally active viral oncogenes during progression to cancer. *J. Virol.* **78**, 11172–11186 (2004).
145. Esteller, M. Epigenetics provides a new generation of oncogenes and tumour-suppressor genes. *Br. J. Cancer* **94**, 179–183 (2006).
146. Kang, G. H. *et al.* Epstein-Barr virus-positive gastric carcinoma demonstrates frequent aberrant methylation of multiple genes and constitutes CpG island methylator phenotype-positive gastric carcinoma. *Am. J. Pathol.* **160**, 787–794 (2002).
147. Vo, Q. N. *et al.* Epstein-Barr virus in gastric adenocarcinomas: association with ethnicity and CDKN2A promoter methylation. *J. Clin. Pathol.* **55**, 669–675 (2002).
148. Yang, B., Guo, M., Herman, J. G. & Clark, D. P. Aberrant promoter methylation profiles of tumor suppressor genes in hepatocellular carcinoma. *Am. J. Pathol.* **163**, 1101–1107 (2005).
149. Li, X. *et al.* p16INK4A hypermethylation is associated with hepatitis virus infection, age, and gender in hepatocellular carcinoma. *Clin. Cancer Res.* **10**, 7484–7489 (2004).
150. Sakuma, K. *et al.* High-density methylation of p14ARF and p16INK4A in Epstein-Barr virus-associated gastric carcinoma. *Int. J. Cancer* **112**, 273–278 (2004).
151. Sudo, M. *et al.* Promoter hypermethylation of E-cadherin and its abnormal expression in Epstein-Barr virus-associated gastric carcinoma. *Int. J. Cancer* **109**, 194–199 (2004).
152. Kim, N. R., Lin, Z., Kim, K. R., Cho, H. Y. & Kim, I. Epstein-Barr virus and p16INK4A methylation in squamous cell carcinoma and precancerous lesions of the cervix uteri. *J. Korean Med. Sci.* **20**, 636–642 (2005).
153. Suzuki, M. *et al.* Aberrant methylation profile of human malignant mesotheliomas and its relationship to SV40 infection. *Oncogene* **24**, 1302–1308 (2005).
154. Chang, M. S. *et al.* CpG island methylation status in gastric carcinoma with and without infection of Epstein-Barr virus. *Clin. Cancer Res.* **12**, 2995–3002 (2006).
155. Jicai, Z. *et al.* Persistent infection of hepatitis B virus is involved in high rate of p16 methylation in hepatocellular carcinoma. *Mol. Carcinog.* **45**, 530–536 (2006).
156. Katoh, H. *et al.* Epigenetic instability and chromosomal instability in hepatocellular carcinoma. *Am. J. Pathol.* **168**, 1375–1384 (2006).
157. Kusano, M. *et al.* Genetic, epigenetic, and clinicopathologic features of gastric carcinomas with the CpG island methylator phenotype and an association with Epstein-Barr virus. *Cancer* **106**, 1467–1479 (2006).
158. O’Nions, J. *et al.* p73 is over-expressed in vulval cancer principally as the δ 2 isoform. *Br. J. Cancer* **85**, 1551–1556 (2001).
159. Gasco, M. *et al.* Epigenetic inactivation of 14-3-3 σ in oral carcinoma: association with p16(INK4a) silencing and human papillomavirus negativity. *Cancer Res.* **62**, 2072–2076 (2002).
160. Dong, S. M. *et al.* Epigenetic inactivation of RASSF1A in head and neck cancer. *Clin. Cancer Res.* **9**, 3635–3640 (2003).
161. Marsit, C. J., McClean, M. D., Furniss, C. S. & Kelsey, K. T. Epigenetic inactivation of the SFRP genes is associated with drinking, smoking and HPV in head and neck squamous cell carcinoma. *Int. J. Cancer* **119**, 1761–1766 (2006). **This paper draws attention to an interaction between smoking and HPV16 infection, and promoter methylation at another HPV-associated site of cancer.**
162. Santin, A. D. *et al.* Gene expression profiles of primary HPV16- and HPV18-infected early stage cervical cancers and normal cervical epithelium: identification of novel candidate molecular markers for cervical cancer diagnosis and therapy. *Virology* **331**, 269–291 (2005).
163. Wilting, S. M. *et al.* Increased gene copy numbers at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. *J. Pathol.* **209**, 220–230 (2006).
164. Tsai, C. N., Tsai, C. L., Tse, K. P., Chang, H. Y. & Chang, Y. S. The Epstein-Barr virus oncogene product, latent membrane protein 1, induces the downregulation of E-cadherin gene expression via activation of DNA methyltransferases. *Proc. Natl Acad. Sci. USA* **99**, 10084–10089 (2002).
165. Etoh, T. *et al.* Increased DNA methyltransferase 1 (DNMT1) protein expression correlates significantly with poorer tumor differentiation and frequent DNA hypermethylation of multiple CpG islands in gastric cancers. *Am. J. Pathol.* **164**, 689–699 (2004).
166. Lee, J. O. *et al.* Hepatitis B virus X protein represses E-cadherin expression via activation of DNA methyltransferase 1. *Oncogene* **24**, 6617–6625 (2005).
167. Li, H. P., Leu, Y. W. & Chang, Y. S. Epigenetic changes in virus-associated human cancers. *Cell Res.* **15**, 262–271 (2005).
168. Liu, J. *et al.* Downregulation of E-cadherin by hepatitis B virus X antigen in hepatocellular carcinoma. *Oncogene* **25**, 1008–1017 (2006).
169. Burgers, W. A. *et al.* Viral oncoproteins target the DNA methyltransferases. *Oncogene* 18 Sep 2006 (doi: 10.1038/sj.onc.1209950). **This is an important paper that shows that HPV oncogenes can activate DNA methyltransferases.**
170. Robertson, K. D. *et al.* DNMT1 forms a complex with RB, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nature Genet.* **25**, 338–342 (2000).
171. McCabe, M. T., Davis, J. N. & Day, M. L. Regulation of DNA methyltransferase 1 by the pRb/E2F1 pathway. *Cancer Res.* **65**, 3624–3632 (2005).
172. McCabe, M. T., Low, J. A., Imperiale, M. J. & Day, M. L. Human polyomavirus BKV transcriptionally activates DNA methyltransferase 1 through the pRb/E2F pathway. *Oncogene* **25**, 2727–2735 (2006).
173. Uren, A. *et al.* Activation of the canonical Wnt pathway during genital keratinocyte transformation: a model for cervical cancer progression. *Cancer Res.* **65**, 6199–6206 (2005).
174. Yukawa, K., Butz, K., Yasui, T., Kikutani, H. & Hoppe-Seyler, F. Regulation of human papillomavirus transcription by the differentiation-dependent epithelial factor Epc-1/skn-1a. *J. Virol.* **70**, 10–16 (1996).
175. Sun, S. & Steinberg, B. M. PTEN is a negative regulator of STAT3 activation in human papillomavirus-infected cells. *J. Gen. Virol.* **83**, 1651–1658 (2002).
176. Kamio, M. *et al.* SOCS1 inhibits HPV-E7-mediated transformation by inducing degradation of E7 protein. *Oncogene* **23**, 3107–3115 (2004).
177. Liu, L. *et al.* A methylation profile of *in vitro* immortalized human cell lines. *Int. J. Oncol.* **26**, 275–285 (2005).
178. Gustafson, K. S., Furth, E. E., Heitjan, D. F., Fansler, Z. B. & Clark, D. P. DNA methylation profiling of cervical squamous intraepithelial lesions using liquid-based cytology specimens: an approach that utilizes receiver-operating characteristic analysis. *Cancer* **102**, 259–268 (2004). **This study shows how the detection of methylated forms in cervical material might assist in the triage of women with cytological abnormality.**
179. Maliukova, A. V. *et al.* Methylation of the putative tumor suppressor gene, RASSF1A, in primary cervical tumors. *Mol. Biol. (Mosk.)* **38**, 1005–1013 (2004).
180. Chan, A. O. *et al.* Eradication of *Helicobacter pylori* infection reverses E-cadherin promoter hypermethylation. *Gut* **55**, 463–468 (2006). **This study reports how the effective treatment of an infectious agent can reverse the hypermethylation of a TSC.**

Acknowledgements

We gratefully acknowledge the support of Cancer Research UK.

Competing interests statement

The authors declare competing financial interests: see Web version for details.

DATABASES

The following terms in this article are linked online to Entrez Gene: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene> CDH1 | DNMT1 | DNMT3A | DNMT3B | IFNB1 | LMP1 | MYC | POU2F3 | PTEN | SFRP1 | SFRP4 | SOCS1 | STAT3 National Cancer Institute: <http://www.cancer.gov> Cervical cancer

FURTHER INFORMATION

Division of Cancer Studies, University of Birmingham: <http://www.cancerstudies.bham.ac.uk>

SUPPLEMENTARY INFORMATION

See online article: S1 (table) Access to this links box is available online.