MOLECULAR BIOLOGY OF HUMAN PAPILLOMAVIRUS AND ITS ASSOCIATION WITH THE CANCER: A REVIEW

Pranjil Kulshrestha, Swati Kamal, Mayank Agarwal, Smriti Singh & A.K Dabral*
Department of Biotechnology, Madhav Institute of Technology and Science, Gwalior, M.P, India
* Associate Vice president (Bacterial Vaccine Division) at Biological E Limited, Hyderabad, India

Abstract
Human papillomavirus (HPV) is the most common sexually transmitted infection in the United States. The relationship of cervical cancer and sexual behaviour was suspected for more than 100 years and was established by epidemiologic studies in the 1960s. In the early 1980s, cervical cancer cells were demonstrated to contain HPV DNA. Epidemiologic studies showing a consistent association between HPV and cervical cancer were published in the 1990s. The first vaccine to prevent infection with four types of HPV was licensed in 2006. Many of the therapeutic compounds are still in the clinical trial phase. Sooner the permanent cure and prevention for the HPV will be available in the market.

Introduction
Human Papillomavirus (HPV) is one of the most common virus groups in the world today affecting the skin and mucosal areas of the body (Warts et al., 2006). Over 140 different strains of HPV have been identified (Warts et al., 2006). Human papillomaviruses (HPV) are small, double stranded DNA viruses that belong to family papillomaviridae (de Villiers EM et al., 2004). Papillomaviruses were first identified, cloned and sequenced from cervical tumor specimens and were subsequently established as important causative agents for development of cervical cancer, the discovery which was honoured by conferring Nobel Prize of Physiology and Medicine for the year 2008 to its inventor Harald zur Hausen (Hausen et al., 1974). The most visible forms of the virus produce warts (papillomas) on the hands, arms, legs and other areas of the skin (Kumar et al., 2014). Human papillomavirus (HPV) is the most common viral infection of the reproductive tract (Kumar et al., 2014). Most sexually active women and men will be infected at some point in their lives and some may be repeatedly infected (Kumar et al., 2014).

Cervical cancer is by far the most common HPV-related disease. Nearly all cases of cervical cancer can be attributable to HPV infection (Kumar et al., 2014). High-risk HPV encode two major oncoproteins termed as E6 and E7, and the respective genes are the only viral genes that are generally retained and expressed in cervical cancer tissues (Kumar et al., 2012). One key activity of E7 is to overcome the pRB tumour suppressor block (Dyson, Howley, Munger, 1989). Genitally transmitted HPV types are contained within supergroup A (also known as Alpha papillomaviruses) (de Villiers EM et al., 2004). Papillomaviruses were first identified, cloned and sequenced from cervical tumor specimens and were subsequently established as important causative agents for development of cervical cancer, the discovery which was honoured by conferring Nobel Prize of Physiology and Medicine for the year 2008 to its inventor Harald zur Hausen (Hausen et al., 1974). The most visible forms of the virus produce warts (papillomas) on the hands, arms, legs and other areas of the skin (Kumar et al., 2014). Human papillomavirus (HPV) is the most common viral infection of the reproductive tract (Kumar et al., 2014). Most sexually active women and men will be infected at some point in their lives and some may be repeatedly infected (Kumar et al., 2014).

The second major group of human papillomaviruses are contained within supergroup B. Viruses from the B1 subgroup such as HPV5 (also known as Beta papillomaviruses) (de Villiers et al., 2004; Myers et al., 1994) cause inapparent or latent infections in the general population, but can become a problem in immuno-suppressed individuals and in individuals who have an inherited defect (Ramoz et al., 2002) which renders them susceptible to infection by papillomaviruses from the B1 supergroup. Such patients can develop skin cancers at the site of HPV infection, and it is thought that B1 viruses may be involved in the development of non-melanoma skin cancer (NMSC) in the general population (Harwood et al., 2004). By contrast, viruses from the B2 subgroup such as HPV4 (also known as Gamma papillomaviruses; (de Villiers et al., 2004), cause cutaneous warts in the general population that can superficially resemble those caused by supergroup A papillomaviruses such as HPV2 (Myers et al., 1994).
The remaining group of HPVs are contained within supergroup E (Myers et al., 1994) (also classified as Mu and Nu-papillomaviruses (de Villiers et al., 2004)). Only three human members from this group are known, and all cause cutaneous papillomas in the general population (Castellsagué, 2008). HPV1 is the most well studied member of this group, and like HPV2 in supergroup A, causes verrucas and palmar warts (Castellsagué, 2008). The neoplastic cervical (CIN), vulvar (VIN), vaginal (VaIN), penile (PIN) and anal (AIN) lesions are associated occasionally with “benign” or “low risk” HPVs such as HPV 6 and HPV 11, but more frequently with the typically “high risk” carcinogenic or oncogenic HPVs like HPV 16, 18, 45 and 31. More than 35 types of HPV have been isolated in neoplastic lesions of the anogenital tract (Castellsagué, 2008).

Clinico-epidemiological and molecular studies have established the casual link between Human Papillomavirus (HPV) infection and cervical cancer as also association of HPV infection with several other cancers. In India, cervical cancer is a leading cancer among women and almost all cases of cervical cancer show prevalence of High Risk (HR)-HPV infection. HPV has been also detected in a significant proportion of oral, esophageal, anal, vaginal, vulvar, and penile cancer and in a small percentage of lung, laryngeal, and stomach cancer in India. Due to lack of organized HPV screening program, insufficient infrastructure and trained manpower and inadequacy in cancer registries, there are not much data available on the countrywide HPV prevalence and its type distribution in different cancers in India. Forthcoming introduction of recently developed HPV vaccines in India given a new urgency to know the prevalence and distribution of various HPV types in different organ sites for the management and monitoring of vaccination program and its impact on prevalence of other cancers.

Cervical cancer is estimated to affect approximately 500,000 women each year, of whom 80% live in developing countries. Virtually all cervical cancer cases result from genital infection with human papillomavirus (HPV). Well-organized programmes of regular gynaecological screening and treatment of precancerous lesions have been very effective in preventing squamous cervical cancer (the most common kind) but have had less impact on adenocarcinomai and are difficult to implement in low-resource settings. In 2006, a quadrivalent vaccine was licensed in several countries, and a bivalent vaccine has recently been licensed in Australia.

Life cycle of HPV
Papillomaviruses are non-lytic, and are not released until the infected cells reach the epithelial surface. Papillomaviruses are resistant to desiccation (Roden et al., 1997) and their extra-cellular survival may be enhanced if they are shed from the epithelial surface within a cornified squame (Bryan and Brown, 2001). A similar pattern has been reported in cattle (Knowles et al., 1996) and may also occur in humans under some circumstances (Coleman et al., 1994). The importance of the immune system in controlling the spread of HPV-associated disease is well established, and patients with immune defects are particularly susceptible to infection, and can develop widespread lesions that are refractory to treatment (Middleton et al., 2003). In the former group, which includes canine oral papillomavirus (COPV), genome amplification begins as soon as cells leave the basal layer, without the intervening proliferative-phase characteristic of viruses such as HPV2 or HPV11 (Middleton et al., 2003).

![Fig.1: The HPV life cycle. Shown is the coordinate expression of the different viral proteins during the course of a productive infection.](image-url)
dysplasia) lesions show relatively low levels of E6 and E7 expression in which the viral genomes replicate episomally, whereas CIN III (severe dysplasia, carcinoma in situ) and invasive cancer lesions often display high-level expression of E6 and E7, in most cases with the integration of viral DNA into the host cell genome (Thomas et al, 1999). Since HPV's use cellular enzymes to replicate their genomes, they need to induce the cellular replication machinery while simultaneously maintaining differentiation, which as we have seen, is achieved by the combined activity of the viral E6 and E7 oncoproteins (Flores et al, 2000).

Numerous approaches have been directed against the polycistronic E6/E7 mRNA to block the expression of both E6 and E7 in HPV-positive cervical cancer cells. This includes selectively inhibiting viral transcription (Goodwin, et al, 2000) or by using antisense constructs (Hamada et al, 1996), ribozymes (Hall, Alexsander, 2003), or short-interfering RNA (siRNA), which lead to the degradation of the E6/E7 mRNAs. As E6 and E7 are expressed together from a bicistronic mRNA, such approaches generally lead to the reactivation of both pRb and p53, and to the attainment of a growth-arrested state that resembles the replicative senescence that is achieved by primary cells after they have reached their normal life span in culture (Grm et al, 2009). A superior strategy appears to be to inhibit the function of E6 alone, which exposes the cell to the pro-apoptotic activity of the E7 protein. This has been achieved using E6-binding peptides and an intrabody-based approach. Indeed, administration of peptide aptamers and intrabodys to E6 was found to induce apoptosis of HPV-positive cancer cells (Butz et al, 2000) (Griffin et al, 2006).

**Papillomavirus Genome and molecular events during cancer progression**

Papillomavirus uncoating may be facilitated by the disruption of intracapsomeric disulphide bonds in the reducing environment of the cell (Li et al., 1998) allowing viral DNA to be transported into the nucleus. The pattern of viral gene expression in these cells is not well defined, but it is generally thought that the viral E1 and E2 proteins are expressed in order to maintain the viral DNA as an episome (Wilson et al., 2002) and to facilitate the correct segregation of genomes during cell division (You et al., 2004). Whether the viral ‘transforming’ proteins (E6 and E7) are also expressed in cells of the basal layer is not certain (Crum et al., 1988), although it appears that initial infection is followed by a proliferative-phase that results in an increase in the number of basal cells harbouring viral episomes. It has been suggested that the viral genome is maintained in the basal layer at around 10–200 copies per cell, and that viral early proteins (E6, E7, E1 and E2) are expressed at low level (Geest et al., 1993; Stanley et al., 1989). The contribution of E6 and E7 to basal cell proliferation during in vivo infection is currently

---

Fig. 2. Cellular protein targets of the E6 and E7 oncoproteins. Some of the important cellular targets are highlighted together with their respective positions in the progression to malignancy.
uncertain, and it has been suggested that expression of E1 (and possibly also E2) may be sufficient for the basal maintenance of viral episomes (Zhang et al., 1999). During papillomavirus infection, E7 (and presumably also E6) is expressed in these cells, the restraint on cell cycle progression is abolished and normal terminal differentiation is retarded (Sherman et al., 1997). E6 and E7 are thought to work together to achieve these effects, and in lesions caused by high-risk HPV types (such as HPV16), the two proteins are expressed from a bicistronic mRNA (Stacey et al., 2000) expressed from the viral early promoter. Papillomaviruses encode two structural proteins that are expressed in the upper layers of infected tissue once viral genome amplification has been completed (Ozbun and Meyers, 1998b). L2 is a minor coat protein that like L1 is produced in a subset of the cells that express E4 (Doorbar et al., 1997). The major capsid protein (L1) is expressed after L2 allowing the assembly of infectious particles in the upper layers of the epithelium (Florin et al., 2002). To be successful, the virus must eventually escape from the infected skin cell and survive extra-cellularly prior to re-infection.

**Cause of infection**

Infection by HPV is basically a sexually transmitted disease. As such, both men and women are involved in the epidemiological chain of infection and are able at the same time to be asymptomatic carriers, transmitters and also victims of the infection by HPV (Castellsague et al, 2003). The most important are: early age at the start of the first sexual relationships, high number of sexual partners throughout life, sexual contacts with high risk individuals (in men, frequent contact with women that practice prostitution and in women, frequent contacts with men with multiple sexual partners). Male circumcision and the strict and systematic use of condoms are factors that can reduce, although without totally preventing, the risk of transmission of HPV between sexual partners (Castellsague et al, 2002). The half-life of infections by HPV has been estimated at 8–10 months for the high risk types and of approximately half that for low risk viral types. The infections by HPV 16 are those that present the most prolonged longevities with average persistence values of 16 months in some studies (Woodman CB, 2001). Initial infection requires access of infectious particles to cells in the basal layer, which for some HPV types is thought to require a break in the stratified epithelium. Such breaks may not be readily apparent, and may occur under conditions where the skin is exposed to water or is abraded (e.g., swimming pool surfaces), or is subjected to other environments where micro traumas may develop. It has been suggested that for a lesion to be maintained, the virus must infect an epithelial stem cell (Egawa, 2003; Schmitt et al., 1996).

**Host Cell Entry**

Host cell entry of HPV is initiated by binding of the virus particle to cell surface receptors (Fig. 3). It has been suggested that virions bind initially to the basement membrane prior to transfer to the basal keratinocyte cell surface (Roberts et al, 2007). It is important to note that the entry of HPV in vitro is initiated by binding to a cell surface receptor in contrast to the in vivo situation where the basement membrane has recently been identified as the primary site of virus binding (Muñoz et al, 2006). Early work investigating the cell surface receptors found that HPVs bind to a widely expressed and evolutionary conserved cell surface receptor and that the interaction depends primarily on L1 (Roden et al, 1994) Glycosaminoglycans (GAGs), especially heparan sulfate, were suggested as initial attachment receptors for HPV VLPs (Roden et al, 1994) (Drobni et al, 2003). Heparan sulfate proteoglycans (HSPG) are frequently found in the extracellular matrix (ECM) and on the surface of most cells. They are involved in several biological functions and because of their location they are appropriate molecules for viral infection (Müller et al, 1995) (Volpers et al, 1995) (Sapp et al, 2009).

![Fig 3. Putative model of interaction of HPV capsids with the ECM and cell surface.](image-url)
Heparan sulfate is often found on two membrane-bound proteoglycans, syndecans and glypicans (Bernfield et al., 1999). Glypicans are predominantly expressed in the central nervous system, whereas syndecans are the predominant HSPG in epithelial cells, the target cells of HPV. Especially syndecan-1 may serve as the primary attachment receptor in vivo due to its high expression level in the appropriate target cells and upregulation during wound healing (Sapp, et al, 2009, Keramat S et al, 2003). Furthermore, other candidate receptors for HPV have been suggested, such as laminin-5. Several in vitro studies have shown that HPV can also bind to a receptor in the ECM, identified as laminin-5 which is able to mediate binding to the ECM (Culp, Budgeon, Christensen, 2006) (Culp, Budgeon et al, 2006). However, laminin-5 interaction seems to be of lesser importance for a productive infection and even though the affinity to laminin-5 is higher than to heparan sulfate, infectious transfer from the ECM seems to require heparan sulphate binding (Sapp, et al, 2009 ; Culp et al, 2006 ; Selinka, 2007). The cell adhesion receptor a6-integrin, which is involved in cell to cell interactions, has been suggested as secondary receptor even though its involvement in HPV infection is rather controversial (Giroglou et al, 2001; Keramat et al., 2003; Culp et al, 2006; Evander et al.; 1997; Yoon et al, 2001). Given the close association of proteoglycans and integrins as matrix components, it is possible that the experimental association of a6-integrin with HPV binding and entry is a secondary effect due to its interaction with HSPGs (Day, Schiller, 2006).

Several studies suggest a role for L2 in facilitating infection via interaction with a secondary receptor(s) (Rodan et al, 2001, Yang et al, 2003). Although cell surface interactions predominantly depend on the major capsid protein L1, it seems likely that the secondary cell surface receptor is L1-specific, although it is possible that L2 may contribute to surface interactions (Sapp et al, 2009).

**Mode of Action**
The high-risk HPV E6 and E7 oncoproteins target critical negative growth regulatory signaling circuits to allow viral genome replication in these post-mitotic cells (Fig. 4) (Noya et al. 2001).

![Fig 4. HPV-16 E6 and E7 oncoproteins abrogate negative growth regulatory signalling pathways of the host cell.](image)

During natural infection, however, the ability of E7 to drive cell proliferation is inhibited in some cells, depending on the levels of the p21 and p27 cyclin-dependent kinase inhibitors (Figure 5) (Noya et al. 2001). High levels of p21 and p27 in differentiating keratinocytes can lead to the formation of inactive complexes with E7 and cyclinE within the cell. It appears that the ability of E7 to drive cells through mitosis in differentiating epithelium may be limited to those cells which express p21 and p27 at a low level or which express sufficient E7 to overcome the block to cell-cycle progression (Noya et al. 2001). This is important given that deregulated expression of the viral oncoproteins is a predisposing factor in the development of HPV-associated cancers (Figure 5) (Noya et al. 2001). The function of the viral E6 protein complements that of E7 and, in the high-risk HPV types, the two proteins are expressed together from a single polycistronic mRNA species (Stacey et al, 2000). A primary role of E6 is its
association with p53 which, in the case of the high-risk HPV types, mediates p53 ubiquitination and degradation (Noya et al, 2001). This is thought to prevent growth arrest or apoptosis in response to E7-mediated cell-cycle entry in the upper epithelial layers, which might otherwise occur through activation of the ARF (ADPribosylation factor) pathway (see Figure 5) (Noya et al. 2001). The general role of E6 as an anti-apoptotic protein is emphasized further by the finding that it also associates with Bak (Thomas and Banks, 1998) and Bax (Li and Dou, 2000). This role of E6 is of key significance in the development of cervical cancers, as it compromises the effectiveness of the cellular DNA damage response and allows the accumulation of secondary mutations to go unchecked (Li and Dou, 2000).

Statistics of HPV infection
Most HPV infections of the cervix are asymptomatic and more than 90% of detected infections are cleared within 2 years (Moscicki ; Schiffman ; Kjaer ; Villa, 2006). The degree of protection and duration of immunity after natural infection are not known. Only 50–60% of women develop serum antibodies to HPV after natural infection (Carter et al, 2004).

<table>
<thead>
<tr>
<th>Site</th>
<th>Attributable to HPV (%)</th>
<th>Developed countries</th>
<th>Developing countries</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total cancers</td>
<td>Attributable to HPV</td>
<td>Total cancers</td>
</tr>
<tr>
<td>Cervix</td>
<td>100</td>
<td>83400</td>
<td>83400</td>
</tr>
<tr>
<td>Penis</td>
<td>40</td>
<td>5200</td>
<td>2100</td>
</tr>
<tr>
<td>Vulva, vagina</td>
<td>40</td>
<td>18300</td>
<td>7300</td>
</tr>
<tr>
<td>Anus</td>
<td>90</td>
<td>14500</td>
<td>13100</td>
</tr>
<tr>
<td>Mouth</td>
<td>&gt; = 3</td>
<td>91200</td>
<td>2700</td>
</tr>
<tr>
<td>Oro-pharynx</td>
<td>&gt; = 12</td>
<td>24400</td>
<td>2900</td>
</tr>
<tr>
<td>All cancers</td>
<td>5</td>
<td>5016100</td>
<td>111500</td>
</tr>
</tbody>
</table>

Table 1. HPV-infection attributable cancer in 2002: developed and developing countries
Therapy
Given the worldwide burden of HPV infection (anogenital warts and neoplasia of several sites), prevention of infection could provide relief from an important public health threat. With the introduction of cervical screening in developed countries, the number of deaths from cervical cancer has declined dramatically (Doorbar et al, 2005) but in developing countries it still remains the number one female cancer, with approximately a quarter of a million deaths occurring each year (Lin et al, 2007). It is thus a major goal to develop safe and effective therapeutics to prevent and to treat HPV infections and their associated diseases. Because of the major role played by host defence mechanisms against HPV, major efforts have been made in the development of candidate prophylactic and therapeutic vaccines against cervical cancer and HPV-related infections in the last few years (Lin et al, 2007). These efforts have led to the approval of the HPV vaccine GARDASIL in several countries in 2006.

Importantly, however, this vaccine does not appear to affect a current HPV infection, cervical cancer precursor lesions or genital warts pre-existing at the time of vaccination (Lin et al, 2007). Furthermore, because HPV infection is pandemic in humans and there is a long latency from HPV infection to the development of invasive cervical cancer in women, several decades will pass before cancer incidences in developing countries begin to decline, even if widespread vaccination is introduced immediately (Wright et al, 2006). Clearly there is a gap in the current treatment arsenal and, even if we currently had full vaccine coverage, we would still require alternative therapies for existing patients.

References


