Biology and natural history of human papillomavirus infection

José Veríssimo Fernandes¹
Josélio Maria Galvão de Araújo¹
Thales Allyrio Araújo de Medeiros Fernandes²

¹Laboratory of Molecular Biology for Infectious Diseases and Cancer, Federal University of Rio Grande do Norte, Natal, Brazil; ²Department of Biomedical Sciences, University of Rio Grande do Norte State, Mossoró, Brazil

Abstract: Human papillomavirus (HPV) is one of the most common causes of sexually transmitted diseases worldwide. It has been proposed that the great majority of women and men have been infected with HPV at least once during their lifetime. HPV infection is associated with a variety of clinical conditions, ranging from benign lesions to cervical cancer. In most cases, the infection is transient, where most of the individuals are healing, eliminating the virus without the presence of any clinical manifestation. Actually, more than 120 HPV types have been cataloged, of which approximately 40 can infect the mucosa of the anogenital tract and are collectively known as mucosal HPV, which are classified based on their oncogenic potential as either low- or high-risk HPV types. The low-risk HPV type causes benign hyperproliferative lesions or genital warts, with a very limited tendency for malignant progression, while the high-risk HPV type is strongly associated with premalignant and malignant cervical lesions. The HPV cycle initiates when the virus gains access to undifferentiated cells of the basement membrane of the squamous columnar junction epithelium of the ectocervix, after these regions are exposed to mechanical or chemical trauma. The basal cells in the transformation zone retain the ability to differentiate, a property required for virion production. Cervical infection with high-risk HPV typically lasts from 12 to 18 months and in most cases is cleared spontaneously. However, in some women the immune response is insufficient to eliminate the virus, resulting in a persistent, long-term infection that may progress to a malignant lesion. In this review, we discuss the biology and natural history of HPV infection and its association with cervical cancer.

Keywords: biology, HPV, cancer

Introduction

Human papillomavirus (HPV) is one of the most common causes of sexually transmitted diseases in both men and women around the world, with prevalence rates varying with the studied population and geographical localization.¹ Most HPV infections are transient, and some studies show that the majority of sexually active individuals are exposed to and acquire infection from this virus at some phase in their lives.²,³ Young women are more vulnerable to the viruses and often become infected by multiple strains of the virus. Thus, HPV infection is more prevalent in young adults at the beginning of their sexual activity, with a subsequent decline in the prevalence rate with increasing age, likely as a result of the development of an immune response against the virus and reduction of sexual activity.⁴-⁷

HPVs can roughly be divided into two tropism groups: those that infect the keratinized surface of the skin, causing common warts, and those that infect the mucosa of the mouth, throat, respiratory tract, and especially the anogenital tract.⁵,⁹ Some HPVs...
of mucosa are associated with anogenital warts, while others have been associated with premalignant squamous intraepithelial lesions of different degrees and malignant lesions, being the main risk factor for invasive cervical cancers.\textsuperscript{10–12} However, only a few HPV-infected individuals progress to invasive cervical cancer.\textsuperscript{8} Most infected individuals eliminate the virus without developing recognized clinical manifestation.\textsuperscript{13}

Today, more than 120 different HPV types have been cataloged, and about 40 can infect the epithelial lining of the anogenital tract and other mucosal areas of the human body.\textsuperscript{14} Those viruses are classified according to their involvement in the genesis of benign or malignant lesions. Considered as low-risk oncogenic HPV (LR-HPV) are those associated with anogenital warts, which are benign, hyperproliferative lesions with a very limited tendency for malignant progression, mostly caused by HPV 6 and 11. Considered as high-risk oncogenic HPV (HR-HPV) are those that are strongly associated with premalignant and malignant cervical lesions, especially caused by HPV 16 and 18.\textsuperscript{13–15}

HPV are also classified into five genera designated by letters of the Greek alphabet – alpha, beta, gamma, mu, and nu – according to the identity of the nucleotide sequences of their genomes, besides the phylogenetic and pathology features.\textsuperscript{14,15} The genus \textit{Alphapapillomavirus} is associated with infections of the anogenital and oral mucosa, and covers a group of HR-HPVs that cause virtually all cases of cervical cancer and a smaller proportion of cases of other cancers of the genital tract, such as vulval, penile, and some types of extragenital cancers, such as head and neck carcinomas, contributing to about 40% of oral cancers.\textsuperscript{16} The alpha genus covers three main branches (clades) where all HPV types with carcinogenic potential are gathered to form a single high-risk clade composed of five species groups (“species”): \textit{\textalpha}-5, \textit{\textalpha}-6, \textit{\textalpha}-7, \textit{\textalpha}-9, and \textit{\textalpha}-11. Although these five species belong to a high-risk clade, each has a different risk profile, with \textit{\textalpha}-9 being the most important species, consisting almost entirely of carcinogenic types, of which the main type is species HPV 16, covering also several other types of high-risk HPV 16-related species.\textsuperscript{17}

Persistent infection with HR-HPV is unequivocally established as a necessary cause of cervical cancer.\textsuperscript{7} The molecules key to the initiation and progression to cancer are oncoproteins E5, E6, and E7, which act primarily by removing the negative regulation of growth of the host cell, generating genomic instability, a hallmark of HPV-associated cancers.\textsuperscript{18,19} Once HPV transmission to the genital tract occurs through sexual contact, the risk factors for the infection and cervical lesions, including cervical cancer, are the same classic risk factors as for other sexually transmitted diseases. The number of sexual partners is the risk factor more consistently associated with genital HPV infection and therefore with cervical cancer. In addition, other indicators of sexual behavior, reproductive activity, heredity, immune and nutritional status, and smoking habit also contribute in some way to the development of cervical cancer.\textsuperscript{12,20,21}

In this review, we discuss in detail aspects of the biology, natural history, and pathogenesis of HPVs, analyzing some specific aspects of their interactions with the infected host and specific host cell components. We also briefly review the current prevention and treatment options for HPV-related diseases. This knowledge is essential for inventing new therapies and effective treatments.

The literature search was conducted in the electronic databases PubMed (National Institutes of Health), Scopus (Elsevier), and Web of Knowledge (Thomson Reuters), using the following keywords: HPV, biology, natural history, immune response, prevention, and treatment. The databases retrieved thousands of articles, and we selected the manuscripts that we thought to be most relevant to our purpose.

**Characteristics of HPV**

HPV is a relatively small nonenveloped virus that contains a double-stranded, closed circular DNA genome, associated with histone-like proteins and protected by a capsid with icosahedron symmetry, formed by two protein types. Each capsid is composed of 72 capsomeres, each of which is composed of five monomeric 55 KDa units that join to form a pentamer corresponding to the major protein capsid – L1. The L1 pentamers are distributed forming a network of intra- and interpentameric disulfide interactions, which serve to stabilize the capsid.\textsuperscript{22} L2 protein is the secondary component of viral capsid, with about 75 KDa located within the virion. To assemble the viral capsids, the pentamers join to copies of L2, which occludes the center of each pentavalent capsomere.\textsuperscript{23,24} Thus, each virion contains 72 copies of L1, the major component of the capsid, and a variable number of copies of L2, the secondary component of the viral capsid, to form a particle approximately 50–60 nm in diameter.\textsuperscript{25,26}

The genome of HPV contains approximately 8000 base pairs and harbors an average of eight open reading frames, divided into three regions. The first is a long control region (LCR), which has the regulatory function of the transcription of the \textit{E6} and \textit{E7} viral genes, and the second is an early region (E), consisting of six open reading frames (\textit{E1}, \textit{E2}, \textit{E3}, \textit{L1}, \textit{L2}, \textit{L3}).
of E4, E5, E6, and E7), which encodes no structural proteins involved in viral replication and oncogenesis. The third is a late region (L), which encodes the L1 and L2 structural proteins. The LCR region of the anogenital HPVs varies in size between 800 and 900 pb, representing about 10% of the genome, and varies substantially in nucleotide composition among individual HPV types.\textsuperscript{27,28} Table 1 shows the main functions of the viral proteins.

Only one strand of the double-stranded DNA serves as the template for viral gene expression, coding for a number of polycistronic mRNA transcripts.\textsuperscript{29} The regulation of viral gene expression is complex and controlled by cellular and viral transcription factors. Most of these regulations occur within the LCR region, which contains cis-acting element-transcription regulators. These sequences are bound by a number of cellular factors as well as the viral E2 product.\textsuperscript{30} A large number of cellular transcription factors have been identified, and the dysfunction of some of them appears to play a significant role in HPV-linked carcinogenesis.\textsuperscript{31}

Table I: The human papillomavirus (HPV) proteins and functions

<table>
<thead>
<tr>
<th>Protein</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>Has DNA-binding functions and a binding site in the origin of replication localized in the long control region, a prerequisite for viral DNA replication</td>
</tr>
<tr>
<td>E2</td>
<td>Controls viral transcription, DNA replication, and segregation of viral genomes</td>
</tr>
<tr>
<td>E4</td>
<td>Favors and supports the HPV genome amplification, besides regulating the expression of late genes, controlling the virus maturation, and facilitating the release of virions</td>
</tr>
<tr>
<td>E5</td>
<td>Enhances the transforming activity of E6 and E7; promotes fusion between cells; contributes to immune response evasion</td>
</tr>
<tr>
<td>E6</td>
<td>Binds and degrades the tumor-suppressor protein p53, inhibiting apoptosis; interacts with proteins of the innate immune response; activates the expression of telomerase</td>
</tr>
<tr>
<td>E7</td>
<td>Binds and degrades the tumor-suppressor protein pRB; increases cdk activity; affects the expression of S-phase genes by directly interacting with EZF factors and with histone deacetylases; induces a peripheral tolerance in cytotoxic T lymphocytes; downregulates the expression of TLR9, contributing to immune response evasion</td>
</tr>
<tr>
<td>L1</td>
<td>Major capsid protein; contains the major determinant required for attachment to cell-surface receptors; it is highly immunogenic and has conformational epitopes that induce the production of neutralizing type-specific antibodies against the virus</td>
</tr>
<tr>
<td>L2</td>
<td>Minor capsid protein; L2 contributes to the binding of virion in the cell receptor, favoring its uptake, transport to the nucleus, and delivery of viral DNA to replication centers; also, E2 helps the packaging of viral DNA into capsids</td>
</tr>
</tbody>
</table>

The replication origin and many transcriptional regulatory elements are found in the upstream LCR region. The virus early promoter, the differentiation-dependent late promoter, and two polyadenylation signals define three general groups of viral genes that are coordinately regulated during host-cell differentiation. The E6 and E7 genes maintain replication competence, and E1, E2, E4, E5, and E8 genes are involved in viral DNA replication and transcriptional control, in addition to other late functions, while the late gene products, L1 and L2, are responsible for the assembly of viral particles.\textsuperscript{32}

The regulation of expression of the late genes in genital HPVs is not well understood. However, it has been shown that the second promoter, or late promoter, is initiated in a differentiation-dependent manner, and thus is activated only when cells are grown in the host’s stratifying/differentiating tissue. Once activated, the late promoter directs transcription from a heterogeneous set of start sites and will serve to produce a set of transcripts that facilitate the translation of L1 and L2 proteins.\textsuperscript{24,33} Activation of the late promoter is accompanied by acceleration of viral DNA replication and by high levels of viral protein expression. As a result, the virus copy number amplifies from 50 copies to several thousands of copies per cell. So when a late promoter is activated, the expression of genes will occur, encoding the structural proteins L1 and L2, which join to assemble the capsids and to form virions.\textsuperscript{29}

**Natural history of HPV infection**

HPV infection is diagnosed by detection of viral DNA, but not by isolation of virus. The genital infection by HPVs is predominantly but not exclusively a sexually transmitted infection. Penetrative vaginal or anal intercourse is not a necessary prerequisite for acquiring the infection by this virus, because it can be transmitted by direct contact with skin or mucosa, during intimate contacts of the genitalia or other mucosal surfaces infected.\textsuperscript{34} HPV infection is very common among human populations worldwide, with prevalence rates varying between different geographical regions according to their level of development and characteristics of the study population. It is estimated that that 50%–80% of sexually active men and women will acquire a genital HPV infection throughout their lives. The peak incidence of infection occurs in the period immediately after the start of sexual activity,\textsuperscript{35} and the risk of infection increases with the number of sexual partners.\textsuperscript{36} HPV is highly infectious, with incubation periods ranging from 3–4 weeks to months or years; the duration of this phase of latency probably relates to the dose of virus received.
Eventually, for reasons not yet very well understood, the cell becomes permissive and viral growth commences, so that viral DNA can be detected and infectious virus is produced and released. This phase of active replication also remains for a variable length of time, but eventually the vast majority of infected individuals develop an effective immune response, becoming viral DNA-negative with subsequent sustained clinical remission from disease.\textsuperscript{37} Consistent immunity depends on cell-mediated effective response to the early proteins, principally E2 and E6. This kind of immune response promotes the regression of the lesion, accompanied or followed by seroconversion with neutralizing antibody production for the major capsid protein L1.\textsuperscript{38}

Only a minority -- estimated to be between 10\% and 20\% -- of infected individuals do not effectively clear the virus. They remain DNA-positive with a persistent active viral infection, and it is these individuals who are at risk for progression to high-grade precancers in the cervix: cervical intraepithelial neoplasia (CIN2/3) and invasive cancer.\textsuperscript{14} The presence of infectious HPV is exclusively intraepithelial, where the virus is maintained in the basal cells of epithelium, probably with low copy number, and then amplifies in a first round of DNA replication to 100 or so nuclear episomes per cell. The virus remains in an episome-maintenance phase, with minimal viral gene expression in the proliferating compartment of the epithelium. The viral genes’ expression becomes maximal in the differentiating noncycling cells of the upper one- to two-thirds of the epithelium with viral capsid synthesis proteins, assembly and shredding confined to the cells layer superficial of squamous epithium in the differentiating.\textsuperscript{39} This replication strategy involving tight control over early gene expression ensures that in dividing cells the powerful oncogenic properties of the E6/E7 proteins of the HR-HPV are not allowed to flourish. However, if viral infection becomes persistent, then the probability of molecular accidents deregulating control of E6/E7 oncogene expression in mitotically active cells will increase and neoplastic progression starts.\textsuperscript{40}

HR-HPVs have the ability to infect several types of epithelial cells, but they can cause cancer more frequently in the uterine cervix.\textsuperscript{41} Cervical cancer arises preferentially in the cervical transformation zone (TZ), located in the boundary between the squamous epithelium of the ectocervix and the columnar epithelium of the endocervix. Basal cells in the TZ retain the ability to differentiate, a property required for virion production.\textsuperscript{42} The basal cells in the TZ are more susceptible to HPV infection in that there are fewer overlying layers than in other locations. In addition, the presence of hormones, such as estrogen and progesterone, which orchestrate cervical changes during menstruation and pregnancy, can help both HPV infection and cancer development.\textsuperscript{41,43,44}

It has been reported that two types of cells are present in the basal layer of the cervix. The first type comprises the transit-amplifying (TA) cells, which are proliferating cells that are able to undergo terminal differentiation. TA cells divide and differentiate, representing the majority of cells in the suprabasal layers. The second class of basal cells is the stem cells, which have unlimited proliferation potential but divide only rarely in order to replenish the TA pool, serving as reserve cells to enable long-term maintenance of the tissue. Only one daughter cell of a stem cell division goes on to become a TA cell, while the other remains a stem cell. It is unclear which cells in the basal layer are the target of HPV infection, and perhaps both cell classes can be infected. If this is true, infection of stem cells could lead to one long-term persistent infection, whereas infection of TA cells could lead to short-term infections, followed by a cure.\textsuperscript{45}

Both in vitro and in vivo studies revealed that the L1 major capsid protein is most probably responsible for the first interaction with the cell surface, because it contains the major determinant required for the initial attachment of the viral particles to the cell-surface receptor – the heparan sulfate proteoglycans. Lammin-5 can also contribute to the binding of viral capsids to the extracellular matrix in the epithelial cell lines.\textsuperscript{9,46,47}

In vivo, the viral particles bind efficiently to regions of the basement membrane (BM) only after these regions have been exposed to mechanical or chemical trauma of the epithelium. The L1 capsid protein binds to heparan sulfate proteoglycans in segments of the BM exposed after epithelial trauma. After this, L1 undergoes a conformational change that exposes the N-terminus of the L2 minor capsid protein, which is cleaved by furin or the closely related proprotein convertase 5 and 6.\textsuperscript{48} L2 proteolysis exposes a previously occluded surface of L1 that binds to a surface receptor of keratinocytes that have migrated over the BM to close the wound. This receptor is still unknown, but in vitro studies indicate the \(\alpha\)-6 integrin as a possible candidate.\textsuperscript{49} The cleavage of L2 may be necessary due to the fact that the intact surface of the epithelia apparently contains sulfation patterns that do not bind capsids. Binding to the BM may promote the preferential interaction with basal keratinocytes that are migrating over the exposed BM to close the wound. Thus, papillomaviruses are the only viruses that initiate the infectious process at an extracellular site.\textsuperscript{50}
The viral particles are internalized via the keratinocyte-surface receptor and subsequently surf toward the cell body. The first phase in infection is internalization, which usually occurs 2–4 hours after cell-surface binding. Spliced viral mRNA was first detected at 12 hours postinfection, and in most assay systems infection is not robustly detected until at least 24 hours after capsid binding. On the other hand, it has been proposed that HPV 16 initially enters via clathrin-coated pits but the traffic occurs through caveosomes to eventually reach the endoplasmic reticulum. Moreover, it has been suggested that the capsids might be internalized via a novel pathway involving tetraspanin-enriched microdomains.

The uncoating is not observed until 8–12 hours after cell-surface binding, and it seems that L2 plays a critical role in the endosome escape. The cytoplasm transport along microtubules is mediated by protein complex, and L2 has been found to interact with the microtubule network via the motor protein dynein during infectious entry. After the entry of the viral genome into the nucleus, the complexes predominantly localize in distinct punctate nuclear domains designated as ND10 bodies or promyelocytic leukemia oncogenic domains. There is evidence that cell division is required for establishment and expression of the viral genome in the nucleus.

**Productive viral life cycle**

The HPV life cycle begins with the infection of stem cells in the basal layer of the epithelium. After the entry in the cells, the virus requires the expression of E1 and E2 genes to maintain a low number of copies of the genome. These proteins bind to the viral origin of replication and recruit cellular DNA polymerases and other proteins necessary for DNA replication. In the suprabasal layer, the expression of genes E1, E2, E5, E6, and E7 contributes to the maintenance of the viral genome and induces cell proliferation, increasing the number of HPV-infected cells in the epithelium, resulting in a higher number of cells that will eventually produce infectious virions. Occurring in the more differentiated cells of this same layer of the epithelium is the activation of the differentiation-dependent promoter and maintenance of the gene expression of E1, E2, E6, and E7. In addition, there is activation of the expression of the E4 gene, whose product will induce amplification of the viral genome replication, greatly increasing the number of virus copies per cell, while at the same time the expression of genes L1 and L2 occurs. In the granular layer, the products of late genes, the major and minor proteins of the viral capsid, L1 and L2 respectively, gather to assemble the viral capsid and formation of new virions, which reach the cornified layer of the epithelium and are released. For a better understanding of the HPV life cycle, Bodily and Laimins divided the process of infection into two stages: a maintenance phase and a differentiation-dependent phase.

**Maintenance phase**

HPV virions infect cells in the basal epithelial layer that become exposed through microlesions. The viral capsid binds initially to the basal cell layer, and infection occurs when activated keratinocytes move into the wound to the upper layers of the epithelium. HPV genomes replicate in the nucleus of the basal cell layer, where the viral replication is considered nonproductive and the virus establishes itself as a low-copy-number episome by using the host DNA replication machinery. In this way, viral proteins are expressed at very low levels in undifferentiated cells, and this contributes to immune evasion and viral persistence.

The maintenance of the viral genome in episomal form in the basal cell layer is a prerequisite for the initiation or maintenance of the viral cycle. The expression of E6, E7, E1, and E2 is necessary for continued episomal maintenance. E1 and E2 cooperate to initiate viral DNA replication, while E6 and E7 modulate cell-cycle regulators to maintain long-term replication competence. The E2 protein is probably a major regulator of this process, because it is able to make both positive and negative control of the early viral promoter that regulates expression of E6, E7, and E1, as well as E2 itself. Following this establishment phase, viral DNA is replicated coordinately with host-cell chromosomes, and virus genomes are distributed to the daughter cells. However, in the differentiated keratinocytes of the suprabasal layers of the epithelium, the virus switches to a rolling-circle mode of DNA replication, amplifying its DNA to a high copy number, synthesizing capsid proteins, and assembling the viral particle.

Thus, HPV replication begins when the host-cell factors interact with the LCR region of the viral genome and begin the transcription of the early genes, mainly E6 and E7. The products of these viral genes promote dysregulation of the cell cycle, subverting the cell growth-regulatory pathways and modifying the cellular environment in order to facilitate
viral replication in a cell that is terminally differentiated and has exited the cell cycle.

**Differentiation-dependent phase**

During the maintenance phase in undifferentiated cells, viral proteins are expressed in extremely low levels. However, when HPV-infected cells leave the basal layer, they undergo differentiation, and high levels of viral proteins synthesis are induced. This restriction of viral protein synthesis to highly differentiated cells delays the expression of viral antigens to locations less susceptible to the host immune response.62

This compartmentalization of gene expression by HPVs constitutes an important strategy to sustain long-term infection, but it creates some problems for the virus. To solve this, the virus forces the cell to remain active in the cell cycle, enabling productive replication in differentiating cells. The viral protein E7 is responsible for maintaining the replication competence in differentiated cells, and this is accomplished in part by inactivation of retinoblastoma (pRB) family members.18 The activation of the late viral promoter in response to host-cell differentiation occurs in the vicinity of the spinous epithelial layer and is responsible for high levels of viral protein expression. As a result, the virus copy number amplifies from 50–200 copies to several thousands of copies per cell.63 This replication strategy adopted by HPV reduces expression of proteins in undifferentiated cells, particularly those that comprise the viral capsid; in the more differentiated cells, the replication strategy functions as an escape mechanism of the immune response.

The viral proteins E1, E4, and E5 contribute to the activation of late viral functions upon differentiation.64,65 The E2-mediated downregulation of E6 and E7 transcription results in the release of the p53 and pRB cellular proteins, and allows the normal differentiation process of the host cell. Then, a putative late promoter activates the capsid genes L1 and L2. Finally, the viral particles are assembled in the nucleus, and the complete virions are released when the cornified layers of the epithelium are shed. The virions are shed in an environment with desquamated cells in the absence of lysis or necrosis, thereby further contributing to virus persistence, because this avoids inflammation.66

Thus, the infection occurs without the proper production of proinflammatory cytokines, which hinders the mobilization and activation of Langerhans cells (LCs) to the site of the infection, which prevents the total elimination of infected cells, favoring their persistence.

It has been observed that most women infected with a specific HPV type will not show evidence of clearance of this same viral type after 6–12 months. It is not known whether HR-HPV can be detected for periods similar to those for LR-HPV. Some studies show similar duration, but others reveal longer durations of infection for HR-HPV types.67–69

It appears that HR-HPV, particularly HPV 16, has a longer clearance time and is more likely to develop persistent infection.67 This is probably due to a greater affinity of their E6 and E7 proteins to form complexes with several cellular proteins responsible for maintaining control over the functions of the cell cycle, resulting in genomic instability that facilitates viral persistence.

**Viral persistence and clearance**

Infection with HR-HPV typically lasts 12–18 months and is eventually cleared by the immune system.67 However, approximately 10% of women fail to clear HPV infections, resulting in a persistent infection. The main consequence of persistent infection with HR-HPV is the development of lesions that may progress to malignancy, and this constitutes the most important risk factor for cervical cancer.19,32,66

Details about the immune response that results in clearance of HPV infection are still unknown. HPV clearance seems to result in long-term humoral and/or cellular protection against reinfection by the same HPV type; whether the protection is lifelong is not known.69 Although the term “clearance” is used when an HPV infection can no longer be detected using sensitive test methods, HPV presence might not be completely eliminated, because the latent state of HPV is still poorly understood. Reappearance of HPV from latency even in the absence of definite immunosuppression is common, but most cases are probably benign.70 In contrast to HPV infections that clear, the risk of cancer increases dramatically when infection becomes persistent.17 This occurs due to the characteristics of the infection and the strategy adopted by the viruses to evade the defense mechanisms of the host, which becomes indifferent to the presence of the virus.

It is important to remember that it is not easy to characterize a persistent HPV infection or differentiate persistent infection from healing followed by reinfection, although reinfection with the same HPV type appears to be uncommon. Many studies classify HPV infection as persistent if the HPV was detected in two consecutive follow-up visits 4–6 months apart. However, because the interval between follow-up visits varies among studies and there are many unknown questions regarding the natural history of HPV, it is complicated to distinguish persistent and transient infections.2 Furthermore,
an undetectable HPV infection could be a period of viral latency, in which the HPV levels are below the detectable threshold of current HPV DNA assays, instead of representing a cleared host.

The persistent nature of HPV infection and DNA viral integration into the genome of the cell contributes to increasing the risk of high-grade and malignant lesions because of the genomic instability generated. E5, E6, and E7 can induce cellular abnormalities, including fusion between cells, generating aneuploidy and chromosomal instability. Abnormal centrosome reduplication can also occur, leading to abnormal numbers of centrosomes. Furthermore, abrogation of cell-cycle checkpoints through the targeting of p53 and pRB family members allows retention of cells with chromosomal abnormalities. These conditions lead to a genomic instability in the HPV-infected cell, increasing the risk of occurrence and of accumulation of genetic changes that extend over a long period of time. The combination of genetic changes and the deficiency of the immune response to detect and eliminate these altered cells represent favorable conditions that allow cancer development.

In benign and malignant HPV lesions, the cellular proliferation increases the demand for nutrients, generating a competition for nutrients and oxygen. To overcome this constraint, both HR-HPV and LR-HPV E7 proteins enhance the levels of the transcription hypoxia-inducible factor-1 (HIF-1), as well as induce the increased expression of HIF-1 target genes under hypoxia conditions. The enhancement of HIF-1 activity results in an increased transcription of a subset of genes that favor angiogenesis, and this induction of angiogenesis plays a crucial role in both the persistence and growth of HPV-induced lesions, because it allows for the availability of nutrients and oxygen in sufficient quantities to meet the needs of cells that are characterized by active proliferation, in both benign and malignant induced HPV lesions.

Mucosal junctions and HPV infection

The cervical and anal TZs are dynamic areas of a few millimeters in size in which a columnar glandular epithelium coexists with a squamous epithelium, and this results from an adaptive process called metaplasia. The constant changes of the cervical epithelium of the mucosa increase vulnerability to disruption of epithelial barrier integrity, facilitating invasion by pathogens, including HPV. These metaplastic conversions are influenced by both the acidification of vaginal pH by traumas such as those resulting from receptive anal intercourse, and can be considered as a stepwise progression of changes. Although these adaptive responses frequently occur at the cervical and anal squamocolumnar junctions, the molecular mechanism underlying the development and the maintenance of the metaplastic epithelium is still not completely understood.

It is believed that this phenomenon could result from the reprogramming of adult stem cells and that the metaplastic epithelium is associated with a deregulated production of receptors, adhesion molecules, and soluble mediators of the inflammatory response, such as cytokines, chemokines, prostaglandins, and growth factors. These molecules might not only exercise influence on epithelial differentiation but also alter the local antiviral immune response, favoring HPV infections. Importantly, a substantial majority of cervical and anal preneoplastic lesions develop within the metaplastic microenvironment of TZs.

This implies that exogenous or endogenous factors specific to the anatomical milieu of squamocolumnar junctions, associated with overexpression of prostaglandin and cytokines such as interleukin 10, create a condition whereby several mechanisms may contribute to the loss of function or impaired chemotaxis of LCs to the site of infection, thus favoring the virus evasion of the immune response and consequently the development of persistent HPV infection at this anatomical site.

In contrast to normal squamous epithelia, metaplastic epithelia have an altered maturation characterized by a weak expression of several keratin intermediate filaments and cell-envelope components such as involucrin and loricrin. The primary function of keratins and other cytoskeletal proteins is to provide resistance to mechanical and nonmechanical stresses that can cause cell rupture and death. Because of their immature state, keratinocytes of the squamocolumnar junctions could be more vulnerable to the trauma required for HPV infection. Therefore, the increased sensitivity of metaplastic epithelia of anal and cervical TZs to preneoplastic lesions can be attributed to the fact that both basement membrane and basal cells could be more accessible to the virus in metaplastic areas, where monostratified glandular and pluristratified squamous forms coexist.

Among the cofactors involved in the malignant transformation of cells infected by HPV, sex hormones may contribute to the process of carcinogenesis by different mechanisms, including the induction of squamous metaplasia in the TZ. In the interactions between steroid hormones and HPV gene expression, alterations of the local immune
microenvironment may be involved. It has been proposed that the cervical TZs with squamous metaplasia have a higher density of estrogen and progesterone receptor-positive cells compared with normal squamous epithelia.\textsuperscript{76,77} Besides, the cervical TZs are more sensitive to the induction of squamous cell carcinogenesis by estrogen.\textsuperscript{80} Among the possible mechanisms by which sex hormones could facilitate HPV-induced carcinogenesis would be the stimulation of expression of the viral genes \textit{E6} and \textit{E7}, directly and/or indirectly through steroid response elements in the viral genome or stimulating cellular proliferation.\textsuperscript{81,82}

Possibly, the hormones can also increase the sensibility of the TZs to persistent HPV infection by altering the local immune microenvironment. It has been observed that 17-\(\beta\) estradiol can both reduce the migration and/or functional capacity of antigen-presenting cells and promote the initiation of T-helper 2 immune response, which is generally associated with the progression of the disease.\textsuperscript{83} Together with these observations, the topography of the cervical TZs is affected by the hormonal status of women, suggesting that sex hormones might not only be involved in the development and maintenance of the metaplastic epithelium but might also be implicated in the high sensitivity of TZs to HPV infection and cancer progression.

It has been demonstrated that there is a reduced secretion of soluble factors of the innate immune response involved in antiviral defense in anal and cervical squamocolumnar epithelial junctions. The \(\beta\)-defensin 2 is weakly expressed in cervical TZs with preneoplastic lesions compared with normal ectocervix.\textsuperscript{75} Furthermore, the density of langerin-positive LCs and their function are significantly altered in the anal and cervical TZs compared with the normal squamous epithelia, suggesting that keratinocyte–LC interaction could play an important role in the establishment of HPV infection in these regions.\textsuperscript{76,77} It was also suggested that a Th2 immunodereivation of immune response could participate in the immunological escape of virus-infected cells. Considering that this immune-response profile is less efficient in the recognition and elimination of virus-infected cells, it is quite possible that it favors evasion of HPV of the immune response, thus facilitating the development of persistent infection and thereby increasing the risk of malignant progression.

Once that immaturity of the metaplastic epithelium is characterized by a weak expression of several keratin intermediate filaments and cell-envelope components, such as involucrin and loricin, it has a lower resistance to mechanical and nonmechanical stresses that cause cell rupture and death.\textsuperscript{76,77,84} In addition, the metaplastic epithelium exhibits a significantly higher density of estrogen and progesterone receptor-positive cells. This appears to increase the sensitivity of TZs to HPV infection, associated with the changes in the local immune response, and it might not only promote viral infections but might also be involved in the HPV-induced development of cervical and anal carcinoma in the TZs. This could be explained, at least in part, by the fact that HPV-associated lesions located elsewhere in the anogenital tract outside the TZs, such as the vagina and vulva, are less likely to progress to cancer than those that develop within the TZs. Thus, it is believed that the anatomical, histological, physiological, and immunological features of the TZs might not only facilitate the mucosal entry of HPV but also may be involved in the HPV-induced development of cervical and anal carcinoma.

### Prevention and treatment

Vaccines to prevent HPV infection were the first vaccines developed specifically to prevent a human cancer. They are composed of noninfectious virus-like particles that resemble the native virus immunologically, and administered with or without adjuvants can induce neutralizing antibody to the relevant virus genotype.\textsuperscript{85}

Two prophylactic HPV vaccines are currently licensed and are available to teenage girls as part of the routine immunization schedules of some countries. The bivalent vaccine incorporates the two HPV types responsible for the majority of cervical cancer (HPV 16 and HPV 18), and a quadrivalent vaccine additionally incorporates HPV 6 and HPV 11, which together are responsible for about 90% of genital warts.\textsuperscript{85} Both vaccines are administered intramuscularly on three occasions, and provide durable protection that has been observed for up to 6.4 years for the bivalent vaccine\textsuperscript{86} and up to 8.5 years for the quadrivalent vaccine.\textsuperscript{87,88} Future vaccines may contain more HPV genotypes to increase the potential for prevention of cervical cancer from 70% to nearer 100%, and nine valent vaccines are currently in clinical trials.\textsuperscript{85}

The protective efficacy of immunization with these vaccines has generally been defined as absence of evidence of persisting infection and/or associated disease attributable to the virus genotype in the vaccine.\textsuperscript{85} Recent data from Australia show a marked decline in presentations with HPV-associated genital warts among women under the age of 28 years, providing the first evidence of field efficacy of HPV vaccination in the general population.\textsuperscript{89}

However, some studies suggest that the currently available vaccines are neither therapeutic for existing
HPV infection nor effective in preventing progression of established HPV infection to disease.\textsuperscript{90} Besides, other HPV types may also cause cervical diseases, although some cross-protection against other related HPV types has been observed in preclinical studies and clinical trials.\textsuperscript{85}

Thus, the primary prevention should also include some lifestyle modifications, such as: delaying first intercourse until age 21, allowing maturation of the TZ, making it less vulnerable to HPV effect; limiting the number of sexual partners; avoidance of tobacco use; use of latex or vinyl condoms with each and every sexual encounter, for patients who are not in stable mutually monogamous relationships.\textsuperscript{91} Circumcision of neonatal boys and adult males may also contribute directly to HPV control, as well as to the control of other sexually transmitted diseases acting as cofactors for HPV transmission.\textsuperscript{92}

The secondary prevention of HPV-related diseases includes the screening of cellular alterations with cytology (Pap smear), and thorough physical and clinical examination of the patients by health-care providers on a regular basis. Patients may be taught self-examination and should be encouraged to report any new findings to their physician. These practical measures must not be abandoned, even in vaccinated populations, given that other high-risk HPV strains also may cause disease, and the success of vaccination campaigns is dependent on a host of factors that includes accessibility and affordability of vaccines, completion of the entire vaccination protocol, and patient education and acceptance of the vaccine.\textsuperscript{91}

Recently, randomized controlled trials have provided evidence that HPV-based screening is more effective than cytology-based cervical screening. So cervical cancer-screening effectiveness may be improved by replacing frequent cytology with HPV screening of women aged 30–35 years or older every 5–8 years, using validated assays.\textsuperscript{92}

The treatment of HPV-related disease aims to eliminate the HPV-infected cells. It generally involves ablative and/or excisional procedures. Ablative procedures include chemodestructive agents, like trichloroacetic acid and dichloroacetic acid, infrared coagulation, cryoablation (cryosurgery), and carbon dioxide laser ablation. Some topical treatments (5-fluorouracil, podophyllin, and podophyllotoxin), or immune-modulating therapy (topical use of imiquimod or intralesion interferon injections) can also be used. The treatment of advanced disease depends on the stage of the malignancy and other factors such as patient comorbidities. It is generally surgical and can also include radiotherapy.\textsuperscript{91}

**Conclusion**

In this review, we tried to summarize the current knowledge about the biology, natural history, pathogenesis, host–pathogen interactions, prevention, and treatment of HPV infection in humans.

The life cycle of HPV begins with infection of stem cells in the basal layer, exposed due to some trauma of the epithelium. After entering these cells, the virus requires the expression of the genes $E1$ and $E2$ to maintain a low number of genome copies. The expression of the genes $E6$, $E7$, $E1$, and $E2$ is required for the continued maintenance of the viral genome in the episomal form. The $E1$ and $E2$ genes act together to initiate replication of viral DNA, while $E6$ and $E7$ modulate cell-cycle regulators to maintain replication competency in the long term. The activation of differentiation-dependent promoters leads to an increased production of proteins $E1$, $E4$, and $E5$. The viral proteins $E6$ and $E7$ act to increase the cellular proliferation and consequently the number of HPV-infected cells, as well as the number of viral genome copies per cell, resulting in a greater number of cells producing infectious virions. In the more differentiated cell layers of the epithelium, expression of the $L1$ and $L2$ genes occurs, whose products represent the major and minor proteins in the viral capsid, which meet to assemble viral particles forming virions, which are released along with desquamated epithelial cells.

HPV has developed an ingenious strategy in which DNA replication and assembly of viral particles occur only in terminally differentiated cells that are scheduled to die of natural causes. This allows the virus to remain in the host for many months, even years, invisible to the host defense mechanisms that become apparently ignorant to the presence of the pathogen over a long period of time.

Another explanation to immune-response evasion is the fact that the viral infectious cycle is confined to the epithelial compartment in the absence of viremia and of spread of the virus by the blood. In addition, the viral particles are eliminated from mucosal surfaces, far from vascular and lymphatic channels. As a result, there is little access of virus and viral proteins to lymph nodes where adaptive immune responses are initiated. Besides, there is no necrosis or virus-induced cell death, and the proinflammatory cytokines that activate the antigen-presenting cells in the epithelium are absent. On the other hand, HPV downregulates the interferon-induced responses, and inhibits the migration and activation of epithelial LCs. This allows for long, uninterrupted periods of viral replication in the epithelium in the absence of host defense mechanisms.
The replication strategy implemented by the virus during its life cycle, slowing down the process of infected cell differentiation and assembly of viral particles only in terminally differentiated cells, hampers their recognition by the host defense mechanisms. This represents a high risk for the host when the infection is by HPV with high oncogenic potential, because it increases the risk of the host immune system becoming tolerant or unresponsive to viral proteins. The combination of virus lifestyle with these factors contributes to evasion of immune response, facilitating the development of persistent infection.

The breakthrough of the HPV vaccine represented the first time in history that girls and women were offered primary prevention against cervical cancer. In the future, it will reduce the burden of cervical precancer and probably also of invasive cervical and other HPV-related diseases.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**


