Human papillomaviruses: diversity, infection and host interactions

Alison A. McBride

Abstract | Human papillomaviruses (HPVs) are an ancient and highly successful group of viruses that have co-evolved with their host to replicate in specific anatomical niches of the stratified epithelia. They replicate persistently in dividing cells, hijack key host cellular processes to manipulate the cellular environment and escape immune detection, and produce virions in terminally differentiated cells that are shed from the host. Some HPVs cause benign, proliferative lesions on the skin and mucosa, and others are associated with the development of cancer. However, most HPVs cause infections that are asymptomatic and inapparent unless the immune system becomes compromised. To date, the genomes of almost 450 distinct HPV types have been isolated and sequenced. In this Review, I explore the diversity, evolution, infectious cycle, host interactions and disease association of HPVs.

Cutaneous epithelium
Layers of stratified keratinocytes that form the outer layer of the skin.

Mucosal epithelium
The moist mucous epithelium that is present at the entrance to body cavities.

Oncogenic HPV types
'High-risk' human papillomavirus (HPV) types that are associated with the development of several human cancers.

Laboratory of Viral Diseases, National Institutes of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA.

e-mail: AMCBRIDE@nih.gov

https://doi.org/10.1038/s41579-021-00617-5

Papillomaviruses are a ubiquitous and ancient group of small DNA viruses that infect the skin and mucosa of vertebrates ranging from fish to mammals. Papillomaviruses are highly species specific and have co-evolved with their vertebrate hosts for millions of years; ancestral forms of these viruses are likely to have infected archaic human populations before the emergence of *Homo sapiens*. The *Papillomaviridae* family currently contains ~450 individual human papillomavirus (HPV) types (Papillomavirus Episteme), which are organized into five phylogenetic genera: Alphapapillomaviruses, Betapapillomaviruses, Gammapapillomaviruses, Mupapillomaviruses and Nupapillomaviruses (FIG. 1a). All five genera of HPVs contain virus types that infect specific regions of the cutaneous epithelium, but the Alphapapillomavirus genus also contains HPV types that are tropic for oral and genital mucosal epithelium. Many HPVs cause asymptomatic infections and can be considered part of the normal microbial skin flora. Others cause benign lesions (warts or papillomas), and some HPV types cause lesions that can progress to cancer after long-term infection. The current taxonomy uses the nucleotide sequence of the L1 gene (encoding the major capsid protein) to define and classify HPV types, species and genera (BOX 1). Each HPV type has <90% nucleotide sequence similarity in the L1 gene compared with any previously identified HPV type[1].

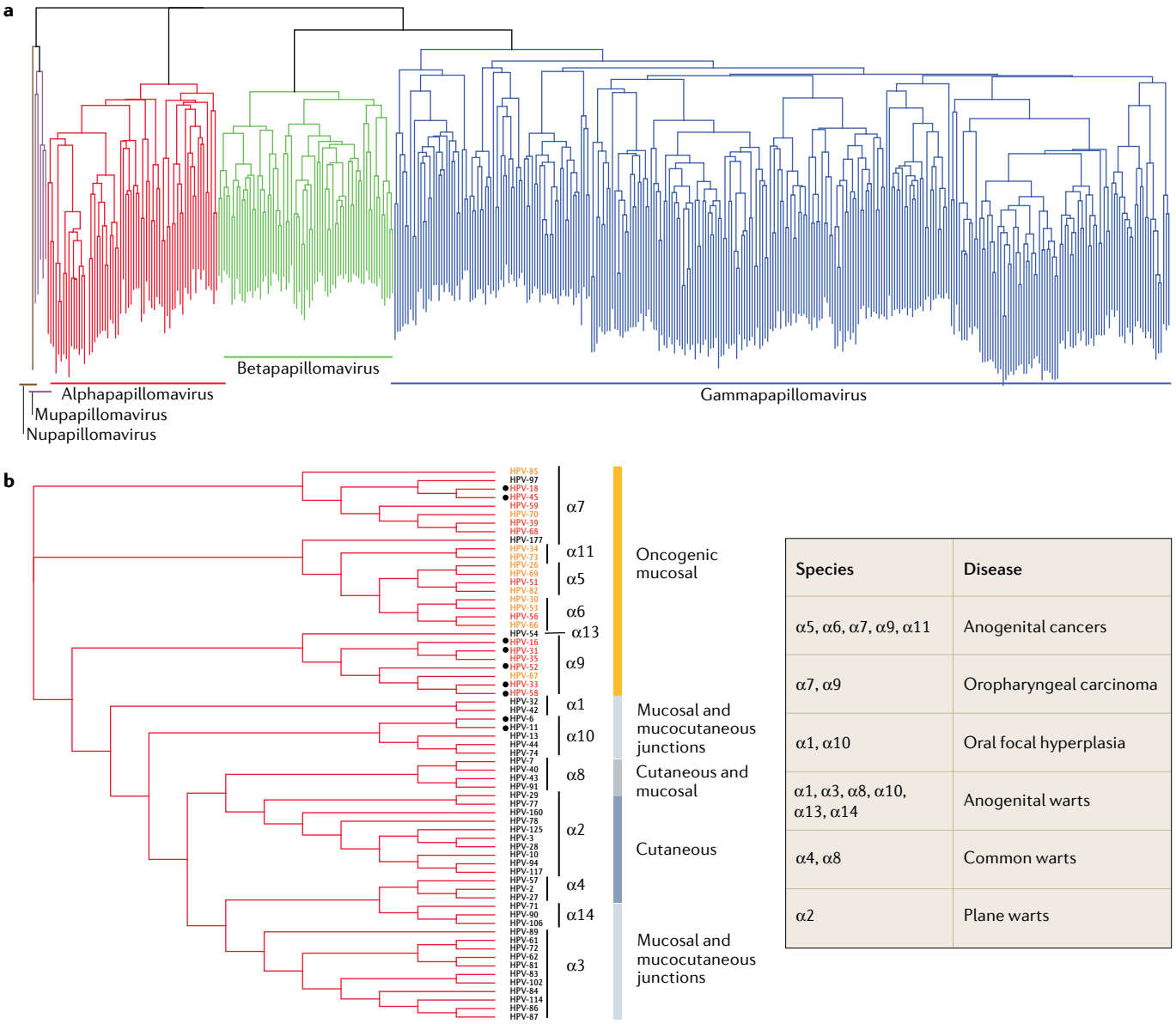
Warts have been described since the time of Hippocrates[2], and in 1907 Giuseppe Ciuffo demonstrated their viral aetiology by transmitting warts from wart extracts passed through a Berkefeld filter[2]. The earliest defined HPV types were isolated from clinically apparent lesions, and researchers noted that diverse skin lesions had common histological features, which suggests that they were caused by related viruses[2]. In 1976, Harald zur Hausen postulated that cervical cancer was associated with HPV infection[3], and this led to decades of research on genital HPVs (Alphapapillomaviruses) and the discovery of oncogenic HPV types. With the advent of sensitive PCR technologies, it became possible to detect low copy numbers of many cutaneous HPV types in healthy human skin[4]. In 2000, Antonsson et al. showed that about 50 different HPV types were common in healthy individuals, and prevalence increased in recipients of organ transplants who were immunosuppressed[4]. These cutaneous HPV types belong to the Betapapillomavirus and Gammapapillomavirus genera and are often acquired in infancy[5]. In recent years, the realization that infection by most HPVs is subclinical or asymptomatic, in combination with the great sampling depth from next-generation sequencing techniques, has led to a great expansion in the number of Betapapillomavirus and Gammapapillomavirus HPV types[6,7] (FIG. 1a).

The Alphapapillomavirus genus is the best studied because of the strong association with clinical disease (FIG. 1b). A subset of the mucosal Alphapapillomaviruses are oncogenic or 'high-risk' HPVs and responsible for 4.5% of cancers worldwide[8]. When Alphapapillomaviruses are organized according to the E7 protein sequence, the oncogenic viruses cluster together (FIG. 1b). Since 2018, there has been a huge expansion in the discovery of Betapapillomaviruses and Gammapapillomaviruses, isolated primarily from individuals with immunodeficiencies[6,7]. For many years, a few Betapapillomaviruses have been implicated in the development of squamous cell skin cancer but, recently, increasing data suggest that these normally commensal viruses can sometimes predispose to the development of UV-mediated skin cancer[9].

Viruses from each genus have evolved to take advantage of different anatomical and biological niches in the cutaneous and mucosal epithelia. Remarkably, different HPVs have evolved the ability to persist and replicate in diverse types of stratified epithelia by exploiting key host cell pathways related to proliferation, differentiation and immune defence. In this Review, I provide a broad overview of the evolution, virology, host interaction and clinical association of the diverse HPVs discovered to date. This is intended to inspire the reader to dive deeper into the remarkable biology of this highly successful and ubiquitous group of human viruses.

Papillomavirus evolution

Papillomaviruses are an ancient group of viruses that have, for the most part, co-evolved with their hosts[10]. This long relationship has resulted in the species-specific tropism and the asymptomatic or benign nature of most HPV infections[11]. Ancestral papillomaviruses infected amniotes in the Palaeozoic era, and when mammals evolved fur or hair and glands in their epithelia ~250–150 million years ago (Ma), papillomaviruses took advantage of these new epithelial ecological niches and expanded[12]. Indeed, primates were infected with viruses closely related to those in the five



Species	Disease
α5, α6, α7, α9, α11	Anogenital cancers
α7, α9	Oropharyngeal carcinoma
α1, α10	Oral focal hyperplasia
α1, α3, α8, α10, α13, α14	Anogenital warts
α4, α8	Common warts
α2	Plane warts

Fig. 1 | **Phylogenetic tree of human papillomaviruses.** Papillomaviruses are classified into genera, species and types (BOX 1). **a** | Human papillomaviruses (HPVs) are classified into five phylogenetic genera: Alphapapillomaviruses, Betapapillomaviruses, Gammapapillomaviruses, Mupapillomaviruses and Nupapillomaviruses. Distribution of viruses across the five genera is indicated. They are classified using a neighbour-joining phylogenetic tree derived from the L1 gene nucleotide sequences of 440 HPV types (reference and non-reference) from the Papillomavirus Episteme. **b** | E7 protein sequences from 65 Alphapapillomavirus HPV reference types (Papillomavirus Episteme) used to generate a neighbour-joining phylogenetic tree. In this E7 tree, oncogenic HPVs (shown in red) and probably or possibly oncogenic viruses (shown in orange) cluster together. The 14 Alphapapillomavirus species are indicated (α1–α14). The virus types currently in the nonavalent HPV vaccine are indicated with a black dot (●) and the anatomical sites most commonly infected with each species are indicated. Table shows common disease association of each Alphapapillomavirus species[132].

Box 1 | Taxonomy of human papillomaviruses

Papillomavirus genomes were first isolated from warts in the 1970s and different isolates were distinguished by restriction enzyme digest and liquid hybridization[134,135]. With the advent of DNA sequencing, a classification system based on the DNA sequence of the L1 gene was ratified by the International Committee on Taxonomy of Viruses[1]. Human papillomaviruses (HPVs) are divided into five genera (named using the Greek alphabet, for example, Alphapapillomaviruses) with each member sharing >60% L1 nucleotide sequence identity. Each genus is further divided into species (with ~71–89% shared sequence identity) and types (>10% different from the closest known type). Variants are isolates of the same type with <10% sequence difference in L1 (REF.[15]). The papillomavirus community has developed strict requirements for naming novel types: the L1 gene sequence must have <90% nucleotide sequence identity with other types and the viral genome must be cloned and submitted to the International HPV Reference Center. These criteria and requirements are ardently debated in the papillomavirus community. To date, there are 220 HPV reference types (submitted and named by the HPV Reference Center) and 221 non-reference types. Sequences of both types can be found at the Papillomavirus Episteme[116].

extant genera (Alphapapillomaviruses, Betapapillomaviruses, Gammapapillomaviruses, Mupapillomaviruses and Nupapillomaviruses) before the emergence of *H. sapiens*[13].

The Alphapapillomaviruses have been well studied at the evolutionary level. A key event in the emergence of the Alphapapillomaviruses was a DNA insertion that occurred between the E2 and L2 regions of the viral genome, which encoded hydrophobic peptides that evolved into the extant E5 proteins. Indeed, the oncogenic Alphapapillomaviruses have been co-evolving with their primate hosts for ~40 million years[13]. Additional key events in their evolutionary history are the acquisition of the ability of the E6 protein to degrade p53, and a small DNA insertion at the 3′ end of the E6 gene that encodes the PDZ-binding domain[12,14]. The Alphapapillomavirus HPV-16 is the most prevalent oncogenic HPV, and can be further classified into four major variant lineages[15]. These lineages originated geographically in different continents and mark the migration of human populations, as well as transmission of HPVs among archaic human populations[13,16].

HPV evolution is extremely slow, with a mutation rate that is only five to ten times that of the host[10]. This rate could be the result of the dependency on high-fidelity host DNA polymerases to replicate viral DNA, balanced against the increased replicative cycles of the virus. Yet HPVs are also under strong purifying selection in that the viral proteins have already evolved to have optimal structure and function, and the small, compact genomes with multiple overlapping reading frames and RNA processing elements are unlikely to tolerate much change. The finding that clinical lesions contain HPVs with substantial genetic variation (possibly due to mutagenesis by the antiviral APOBEC cytidine deaminases) supports this hypothesis, which indicates that mutations occur frequently but are only very rarely fixed in viral genomes[17].

The HPV infectious cycle

HPVs infect and replicate in the mucosal and cutaneous epithelia of their hosts. These stratified epithelia contain a basal layer of self-renewing cells that divide symmetrically to replenish the basal layer, and asymmetrically to generate daughter cells that make up the differentiated layers of the tissue. The HPV life cycle takes advantage of this process by establishing a reservoir of persistent infection in the self-renewing basal cells, and only generating virion particles in the terminally differentiated cells. Virions are released into the environment in squames (dead cells) that are sloughed from the surface of the epithelium. Different HPV types infect diverse anatomical regions of the cutaneous or mucosal epithelia, but they all have a similar differentiation-dependent life cycle (FIG. 2). In some cases, the infection is inapparent, and in others the enhanced proliferation of the layers of the epithelium results in a benign skin tumour (a papilloma). This strategy of establishing low-level persistent infection in self-renewing cells, while restricting productive infection to terminally differentiated cells, is highly effective and promotes long-term infection and immune evasion.

Tissue tropism. HPVs are highly specialized viruses. Initial infection, and progression of the viral life cycle, occurs only in the distinct layers of stratified, differentiating epithelia. This strategy ensures a reservoir of viral infection with low-level viral activity in cells that are monitored by the immune system, and high-level production of virions in terminally differentiated cells destined to be exfoliated into the environment. At each step, host cell pathways are hijacked and manipulated to facilitate the viral life cycle. For example, this strategy requires viral DNA to be produced in differentiated cells that have normally exited the cell cycle and are therefore unable to replicate DNA. Instead, HPVs induce DNA damage signalling in these differentiated cells and the host responds by delivering DNA repair machinery to the viral genome, to what it perceives is a site of DNA damage[18].

Although the overall differentiation-dependent strategy of all HPVs is similar, some ('low-risk' Alphapapillomaviruses) stimulate cell cycle entry only in the upper differentiated layers to support viral genome replication whereas others (high-risk Alphapapillomaviruses) drive the cell cycle in both lower and upper layers[19]. Most HPVs also delay keratinocyte differentiation to promote viral replication, but different HPV types achieve this by manipulating distinct cellular pathways[19].

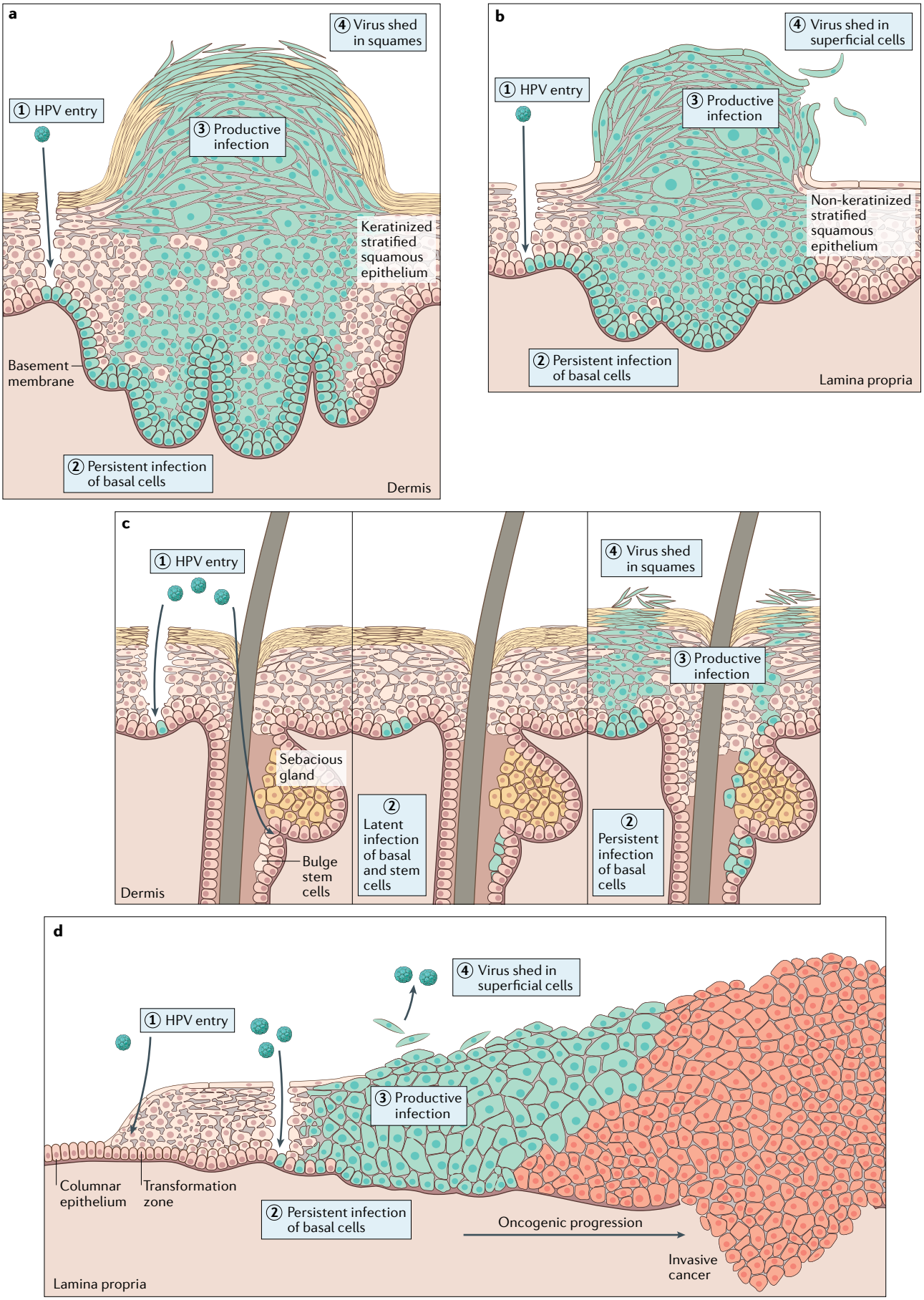
HPV infection is not only restricted by the differentiation-dependent infectious cycle. HPVs from different genera will only productively replicate in anatomically distinct regions of cutaneous or mucosal epithelia of the host (FIGS 1 and 2). This tropism is not dependent on viral entry receptors but is thought to be related to the ability of each virus to transcribe RNA in different types of keratinocytes. Historically, attempts to propagate HPVs in classic virological culture systems failed[20], but even today there are challenges in propagating non-Alphapapillomaviruses in highly specialized keratinocyte culture systems. The primary restriction of HPV infection is not at the stage of cell entry; HPVs can enter and deliver genetic material to many cell types. Alphapapillomavirus HPV genomes will readily transcribe viral genes and replicate in keratinocytes, but non-Alphapapillomavirus genomes are more reticent

p53
A tumour suppressor protein that restricts cell growth when cells are damaged; p53 is very often mutated in non-human papillomavirus (non-HPV) cancers.

PDZ-binding domain
The complementary recognition module of PDZ domains, which are interaction modules found in many proteins.

Persistent infection
A long-term viral infection that maintains a reservoir of host cells containing the viral genome, evades immune clearance and often produces virus particles.

Keratinocyte
An epithelial cell of the epidermis that produces keratin.


a
④ Virus shed in squames
① HPV entry
③ Productive infection
Keratinized stratified squamous epithelium
Basement membrane
② Persistent infection of basal cells
Dermis
b
④ Virus shed in superficial cells
① HPV entry
③ Productive infection
Non-keratinized stratified squamous epithelium
② Persistent infection of basal cells
Lamina propria
c
① HPV entry
Sebacious gland
Bulge stem cells
Dermis
② Latent infection of basal and stem cells
④ Virus shed in squames
③ Productive infection
② Persistent infection of basal cells
d
④ Virus shed in superficial cells
① HPV entry
③ Productive infection
Columnar epithelium
Transformation zone
② Persistent infection of basal cells
Oncogenic progression
Invasive cancer
Lamina propria

◀ Fig. 2 | **Human papillomavirus infectious cycles in stratified epithelia of different host tissues.** All human papillomaviruses (HPVs) have similar infectious cycles; they establish a reservoir of low-level persistent infection in the proliferating basal cells of a stratified epithelium and limit productive infection to the superficial differentiated cells. Shown are four examples of HPV infection with key anatomical features and stages of HPV infection indicated: common skin wart, often caused by Mupapillomavirus infection (panel **a**); anogenital wart infected with low-risk Alphapapillomaviruses (panel **b**); infection of the skin and hair follicles by Betapapillomavirus (panel **c**); cervical infection by high-risk Alphapapillomaviruses (panel **d**). In all cases, virus enters the epithelium through a microfissure (step 1) and establishes a persistent infection in the dividing basal cells (step 2). As the infected cells differentiate, the productive phase of viral infection is activated, with amplification of viral DNA and synthesis of capsid proteins (step 3). Virus is shed from the surface of the epithelium in exfoliating squames or superficial cells (step 4). Infection can give rise to benign growths in the cutaneous (panel **a**) or mucosal (panel **b**) epithelium due to thickening of each cell layer. Betapapillomavirus infections are asymptomatic (panel **c**). Here, virus also enters through a microabrasion (step 1) and establishes infection in the basal skin layer (step 2), but stem cells in the hair follicle bulge can also be infected and act as a reservoir of infection (step 2). Betapapillomavirus infections are commonly latent with only sporadic productive infection (step 3) and release of virus (step 4). Cervical infections (panel **d**) are only grossly visible after acetic acid treatment, which reveals white plaques. Here, virus can infect the cervix through a microfissure, but can also access the basal layer of cells at the transition zone where columnar and stratified epithelium meet. Initially, cervical infections with high-risk HPVs are similar to other infections with a reservoir of infection in the basal cells and viral production and egress in the superficial layers. However, if not cleared by the immune system, the infections can become chronic and eventually undergo oncogenic progression (dark orange cells). These cells fail to differentiate and do not produce virus, and in a small number of cases can progress to invasive cancer.

Basement membrane
A thin, non-cellular layer that lies between the dermis and the epidermis of the skin.

Transformation zone
The junction between squamous and columnar epithelial cells in the cervix, anus and similar tissues. Also known as the squamocolumnar junction or transition zone.

Columnar epithelium
A single-layer, glandular epithelium that lines the endocervix.

Latency
The viral genome is dormant in cells and does not produce virus.

Retromer complex
A protein complex that sorts and traffics proteins between the endosome and trans-Golgi network.

PML nuclear bodies
Small, interferon-inducible nuclear bodies involved in many cell processes; they consist of a scaffold of PML protein but contain many other proteins.

Oncoproteins
Proteins that dysregulate the host cell, and by doing so promote carcinogenesis.

and will only replicate when their replication proteins are co-expressed from expression vectors in those cells[21]. This indicates a transcriptional block, and a pivotal study showed that this is primarily due to strong repression from the virally encoded E8^E2 repressor protein[21]. Mutation of this viral repressor results in a virus that can freely replicate and transcribe its genome in keratinocytes[21].

Highly specialized, tissue-tropic viruses often evolve to take full advantage of differential codon usage in different tissues. A study proposed[22] that expression of HPV late genes was finely tuned to the codon usage and tRNA availability in differentiated keratinocytes, and this translational adaptation has been validated recently using sophisticated genomic analyses[23]. This indicates that the sequences of the L1 and L2 genes have evolved to use the tRNAs that are most abundant in the differentiated cell layers to optimize synthesis of these highly expressed proteins in the cells in which viral particles are assembled.

Viral entry. To initiate infection of a stratified epithelium, the HPV virion must access the dividing basal cell layer (FIG. 2). This is usually by means of a microabrasion that exposes heparin sulfate proteoglycans on the basement membrane[24]. Binding to these glycoproteins induces a conformational change in the capsid that promotes association with a secondary receptor on the keratinocytes[25]. Although less well understood, the virus can also gain access to the lower basal cells in the transformation zone of the cervix, where the single-layer columnar epithelium of the endocervix meets the stratified layers of the ectocervix (FIG. 2). The transformation zone is particularly susceptible to infection and carcinogenesis by oncogenic HPVs[26]. Another site of access, and persistence, is provided by the stem cells in the bulge of the hair follicle, and these cells are a reservoir for Betapapillomavirus infection[27] (FIG. 2). Infection of slow-cycling stem cells promotes persistent infection and, possibly, latency[28].

After binding to the keratinocyte, the virion is internalized by endocytosis and hijacks the retromer complex for retrograde transport to the trans-Golgi network[29]. During this process, the minor capsid protein L2 penetrates the endosomal membrane and cloaks the virion particle in a membrane vesicle[30], in part to evade innate immune detection[31]. When the host cell undergoes mitosis, the membrane-encased virions can associate with the host mitotic chromosomes until the nuclear envelope reforms[32,33]. PML nuclear bodies (which disassemble during mitosis) reform in association with the encased virions prior to egress of the viral DNA from the capsid[34]. This step is important for establishment of the infection, as the PML nuclear bodies have antiviral properties that are counteracted by most nuclear DNA viruses.

Early transcription and replication. The initial stages of viral transcription and replication take place at the PML nuclear bodies[34,35]. Many DNA viruses initiate infection at these bodies as they coordinate many advantageous nuclear processes, but they are also important for antiviral defence[36]. Not surprisingly, viruses manipulate these bodies by taking advantage of factors that support viral processes while displacing those that inhibit viral infection[37].

The HPV E1 and E2 replication proteins are the first to be expressed and initiate a few rounds of viral DNA synthesis to establish the genomes as extrachromosomal plasmids[38]. The genomes are maintained at low copy number in the self-renewing cells and are partitioned to daughter cells by interaction with the host mitotic chromosomes[39]. Only low levels of viral genes are expressed in the persistently infected cells. The E5, E6 and E7 proteins promote the viral life cycle by manipulating the balance of cellular proliferation and differentiation to provide an environment conducive to both persistent infection in the basal cells and productive infection in differentiated cells. Cells normally exit the cell cycle upon leaving the basal layer, but HPVs induce cell cycle re-entry and delay differentiation so that the infected cells can support viral replication in the upper layers of the epithelium. Late gene expression and virion production require fully differentiated keratinocytes, and so the overall goal is delay, and not complete inhibition, of terminal differentiation. The viral proteins from different HPV types modulate the balance of proliferation and differentiation in various ways (TABLE 1). For example, Betapapillomavirus E6 proteins hinder differentiation by binding to MAML1 and inhibiting Notch signalling; in contrast, the same region of Alphapapillomavirus E6 binds to E6AP to target p53 for degradation and this could also inhibit Notch signalling by a different mechanism[40].

Simultaneously, the E5, E6 and E7 accessory proteins interfere with host immune detection, and it has been proposed that in oncogenic viruses the primary function of the E6 and E7 proteins (oncoproteins) is to evade

host immunity, with the unintended consequence of increasing the susceptibility of cells to oncogenesis[41].

Late transcription and replication. When infected basal cells divide asymmetrically and begin the process of terminal differentiation, viral transcription switches to an intermediate phase and it is thought that increased levels of the E1 and E2 mRNAs and proteins are expressed[42–44]. This induces amplification of viral DNA in replication factories in the nucleus of differentiated cells[31]. These cells have completed cellular DNA replication and do not express factors required for replicative DNA synthesis. Resourcefully, the virus mimics DNA damage in viral replication and induces DNA damage signalling and repair pathways, which results in recombination-directed replication of viral DNA[45].

Late transcripts are expressed in the most superficial cells of an infected lesion and encode the major and minor capsid proteins, L1 and L2, respectively. When overexpressed in non-keratinocyte cells, L1 and L2

Table 1 | **Key functions of human papillomavirus proteins**

Protein	Function and characteristics	Role in infection
E1	Origin-binding DNA helicase[117] Site-specific DNA-binding protein[118,119]	Initiates viral DNA replication and recruits cellular replication machinery[120,121]
E2	Dimeric DNA-binding protein that binds E2-binding sites in viral DNA[119] Interacts with and loads the E1 DNA helicase onto the replication origin[122] Tethers viral DNA to the host chromatin[123]	Regulates viral transcription[124] Supports viral DNA replication[121] Partitions viral genomes to daughter cells[125]
E8^E2	Site-specific DNA-binding protein that binds E2-binding sites in viral DNA[56] Interacts with cellular transcriptional repressor factors[56]	Represses viral transcription and replication to maintain low-level persistent infection[56]
E1^E4	Highly expressed late protein[59] Induces G2 cell cycle arrest[126] Disrupts and reorganizes keratin filaments[58]	Promotes viral genome amplification[59] Causes fragility in keratinocyte squames to promote viral egress[58]
E5	Hydrophobic membrane proteins[60] Four types[53] Encoded only by Alphapapillomaviruses	Reduces immune detection[61] Promotes cell proliferation and productive stages of infection[60]
E7	Binds and inactivates Rb protein (pRb) and related pocket proteins to promote cell cycle entry[62] Degrades pRb (high-risk HPVs only)[62] Induces DNA damage response signalling in differentiated cells[127] Abrogates the innate immune response[55]	Promotes cell proliferation and viral DNA amplification in differentiating cells[55] Decouples response to oncogene-induced senescence[55] Promotes genome amplification[18] Reduces immune detection[55]
E6	In Alphapapillomaviruses: Inhibits p53 function[55,108] Binds E6-associated protein E3 ubiquitin ligase[128] Degrades p53 (high-risk HPVs only)[55,129] Degrades PDZ domain-containing proteins involved in cell polarity (high-risk HPVs only)[108] Activates telomerase (high-risk HPVs only)[64]	Prevents growth arrest and apoptosis[55] Decouples response to oncogene-induced senescence[55] Modulates cell polarity to promote viral genome replication and maintenance[55]
	In Betapapillomaviruses: Inhibits p53 function[110] Inhibits DNA damage response[110] Binds MAML1 to inhibit NOTCH signalling[63,65]	Prevents replicative senescence[64] Abrogates immune signalling[55] Prevents growth arrest and apoptosis in response to UV irradiation[66] Inhibits keratinocyte differentiation[63,65] Abrogates immune signalling[55]
L1	Capsid protein that self-assembles into capsids consisting of 360 L1 proteins[69]	Major capsid protein[69]
L2	Between 12 and 72 L2 proteins per capsid[130] Traffics viral genome into nucleus, associates with host chromosomes and PML nuclear bodies in early infection[80] Packages viral genome into capsids at late stages of infection[105]	Minor capsid protein[131] Viral genome chaperone[67]

HPV, human papillomavirus.

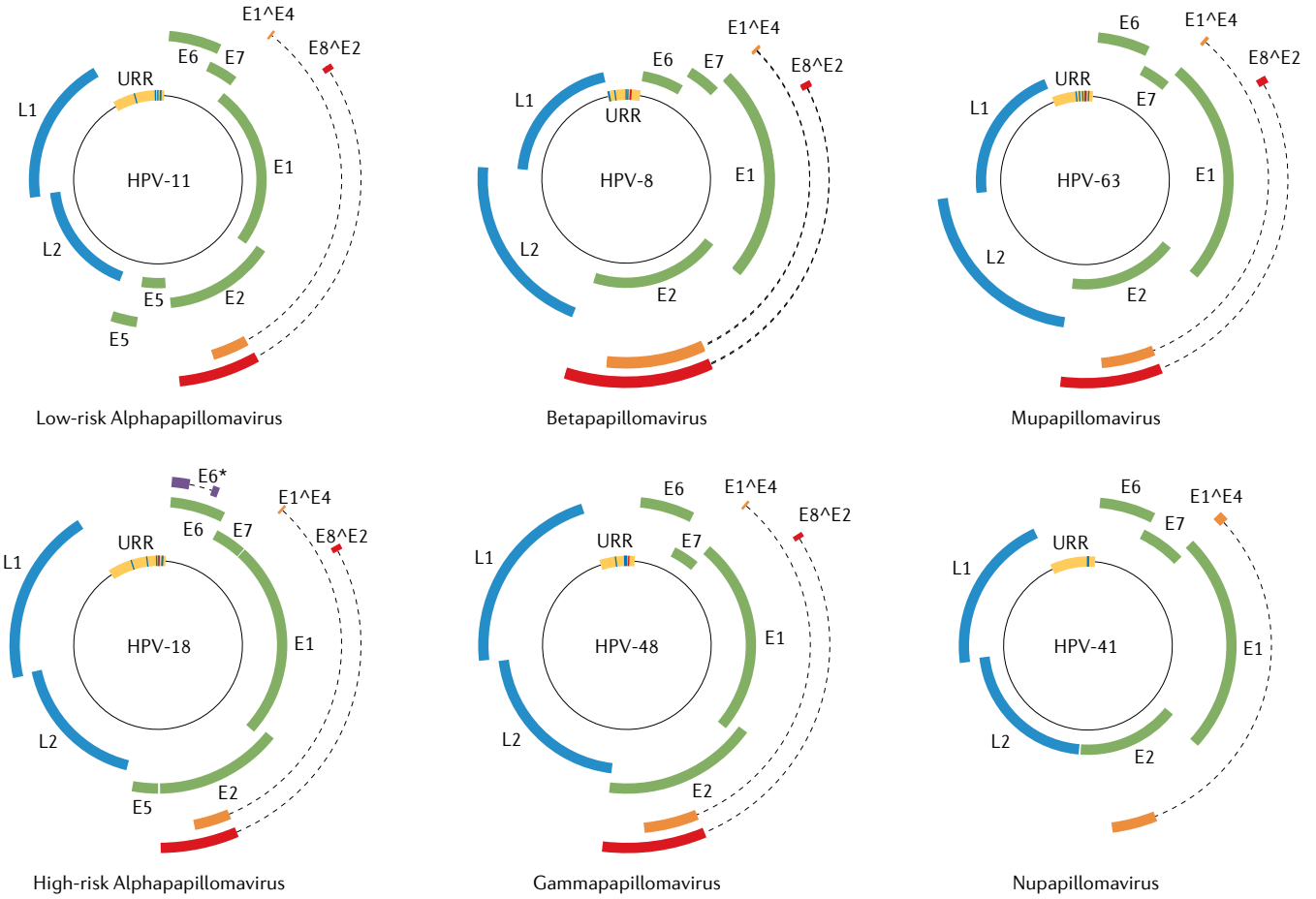


Fig. 3 | **Genomic organization of human papillomaviruses from different genera.** Genome maps of six different human papillomavirus (HPV) genomes are shown (one each from Betapapillomavirus, Gammapapillomavirus, Mupapillomavirus and Nupapillomavirus genera, and examples of genomes of 'low-risk' and 'high-risk' Alphapapillomaviruses). Early genes are shown in green and late genes in blue. Major spliced transcripts shown for E1^E4 (orange), E8^E2 (red) and E6* (purple; only expressed in high-risk HPVs). The upstream regulatory region (URR) is shown in yellow with E2-binding sites (blue) and an E1-binding site (red) indicated. Genome maps adapted from https://pave.niaid.nih.gov/.

efficiently package the viral minichromosomes[46], but very little is known about the assembly process in HPV-infected tissues. HPVs are unusual in that the viral genome is assembled in host nucleosomes inside the viral particle. Moreover, the viral chromatin is enriched in histone post-translational modifications that are usually associated with active chromatin[47]. It has been postulated that these modified histones might enhance packaging of viral chromatin, or may assist in early steps of infection of a new host[47].

Egress and transmission. Arrays of virion particles are present in desquamated cornified cells that are sloughed from the epithelial surface as efficient vehicles to transmit high concentrations of virions to susceptible hosts[48]. Virion particles are very stable, much more so when contained within exfoliated squames[49–51]. Most studies of viral infection and entry have used purified virions generated in the laboratory, but there is now renewed appreciation of studying transmission mediated by squames[49,51].

Host to host papillomavirus transmission has been studied in animal models[52], but this is much harder to evaluate in humans.

Minichromosomes
Small circular DNA molecules that are assembled in chromatin.

Viral genes and host interactions

Genome organization. All HPVs have circular double-stranded DNA genomes of ~7–8 kbp that are assembled into chromatin using host histones at all stages of the infectious cycle (FIG. 3). The genomes consist of three regions: the upstream regulatory region (URR) that contains viral promoters, enhancers and the replication origin; the early coding region containing genes expressed early in the infectious cycle and in the lower layers of an infected lesion; and the late coding region that encodes the major and minor capsid proteins, L1 and L2. Early and late transcription initiate from promoters in the URR or early coding region and terminate at polyadenylation sites at the end of the early and late regions, respectively. Viral transcription is strictly regulated by RNA processing, and alternative splicing generates a diverse series of transcripts[44].

The Alphapapillomaviruses are unique in that they have an additional coding segment between the early and late regions that encodes an E5 gene. This DNA segment is likely to have been integrated into an ancestral papillomavirus genome ~200 Ma, about the time that the Alphapapillomaviruses emerged[12]. In fact,

each Alphapapillomavirus encodes one (or two) of four different types of E5 protein (E5α, E5β, E5γ and E5δ); they are all hydrophobic, but otherwise unrelated[53]. The oncogenic Alphapapillomaviruses are also unique in that the E6 and E7 proteins are transcribed from a polycistronic, alternatively spliced mRNA, whereas E6 and E7 are transcribed from separate promoters in low-risk Alphapapillomaviruses. Splicing of the oncogenic polycistronic E6–E7 transcript also gives rise to a novel protein, termed E6* (FIG. 3). Furthermore, a small insert (30–60 bp) in the 3′ end of the E6 open reading frame (ORF), which is found only in the oncogenic HPVs, encodes a PDZ-binding domain[14].

Viral gene functions. All papillomaviruses encode four highly conserved core proteins: the E1 and E2 replication proteins and the L1 and L2 capsid proteins. The complex HPV life cycle requires each viral protein to interact with a multitude of cellular factors and processes (TABLE 1). The smaller, less well conserved proteins E5, E6 and E7 can be regarded as evolutionary adaptations.

E1 is a helicase that specifically binds to, and unwinds, the origin of replication; E2 binds both to the E1 protein and to E2-binding sites in the origin to form a high-affinity initiation complex. Viral DNA is synthesized by the cellular replication machinery, but E1 and E2 assist by recruiting these factors to the origin[39]. E1 and E2 initiate transient replication from the minimal origin, but additional sequence elements from the URR are required to both augment transient replication and support stable, long-term genome replication. In dividing cells, the low viral genome copy number is maintained by tightly regulated expression and localization of the E1 and E2 proteins[39,54]. Furthermore, the E2 protein retains the genomes in the nucleus and partitions them to daughter cells by tethering them to host chromosomes[39]. This strategy maintains persistent infection, and ensures that the genomes do not trigger an innate immune response in the cytoplasm[55].

The E2 protein is also a transcriptional regulator that can activate, or repress, viral transcription by binding to E2-binding sites in the viral regulatory regions. Essential to persistent HPV infection is a truncated E2 protein that contains a short peptide from the E8 ORF (overlapping the E1 ORF) fused to the carboxy-terminal half of E2. In the resultant E8^E2 protein, the E8 domain recruits transcriptional co-repressors and the E2 DNA binding and dimerization domain competes for binding with full-length E2 (REF.[56]). E8^E2 represses both viral transcription and replication, thus maintaining low-level viral activity and genome copy number in the lower layers of a lesion.

The E4 protein is encoded by the E1^E4 spliced message. Although encoded by the early region of the viral genome, it is expressed at late times and can constitute ~30% of total protein in some warts[57]. The E4 ORF overlaps the hinge region of the E2 protein and is quite divergent among different viral types. The E1^E4 protein causes G2 cell cycle arrest in differentiated cells, presumably to augment viral replication. E1^E4 also interacts with, and collapses, keratin filaments and this contributes to the fragility of desquamating cornified cells and subsequent release of virions[48,58]. E4 may also function to protect virions released into the environment in exfoliated squames[59].

The E5 protein, which is only encoded by the Alphapapillomaviruses, is not absolutely necessary for the viral life cycle but promotes the replicative productive stages[60]. E5 is a transmembrane protein that localizes to the endoplasmic reticulum and promotes cellular proliferation by activating pathways such as epidermal growth factor receptor signalling. E5 evades both adaptive and innate immunity by preventing presentation of viral peptides by major histocompatibility complex molecules, by suppressing IFNκ (a keratinocyte-specific interferon) and by crosstalk with growth signalling pathways[60,61].

A key function of HPV E6 and E7 proteins is to promote a cellular environment conducive to viral replication by modulating the balance of proliferation and differentiation while maintaining the reservoir of infected basal cells. E6 and E7 proteins each bind a myriad of host cell factors; many of these differ among the different genera of papillomaviruses, and in the high-risk HPVs E6 and E7 are oncogenes (TABLE 1). The differences observed are related to both the HPV type and the target epithelium. The E7 proteins from all genera bind to the Rb protein (pRb; a tumour suppressor protein and major G1 checkpoint regulator) to induce entry into S phase[62]. However, only the oncogenic E7 proteins degrade pRb and the ensuing, aberrant proliferation induces p53. All Alphapapillomavirus E6 proteins form a complex with E3 ligase E6-associated protein (E6AP) and disrupt p53 transactivation[63], but only the oncogenic E6 proteins co-opt the E3 ubiquitin ligase function of E6AP to degrade p53. In the absence of p53, cells do not undergo growth arrest upon DNA damage and, over time, this promotes carcinogenesis. Oncogenic E6 proteins also bind and degrade PDZ polarity proteins, and it has been proposed that this could disrupt the balance of symmetrical and asymmetrical cell division to regulate viral DNA replication[55]. Moreover, oncogenic E6 proteins upregulate telomerase to maintain telomere length and prevent senescence in continually proliferating cells[64]. Conversely, the E6 proteins from the Betapapillomaviruses, Gammapapillomaviruses, Mupapillomaviruses and Nupapillomaviruses do not bind to E6AP, and disrupt differentiation by binding and blocking the MAML1 transcriptional co-activator downstream from Notch signalling[63,65]. Betapapillomavirus E6 proteins also inhibit apoptosis following UV-induced DNA damage and promote the development of squamous cell skin cancer[66].

The L2 protein is the minor capsid protein, and chaperones the viral genome at both early and late stages of the life cycle[67]. L2 is packaged inside the capsid along with the viral minichromosome; upon entry, a conformational change results in protrusion of specific L2 sequences onto the surface of the virion and this allows it to traffic the virus particle through the endosome on retromer complexes[29]. Additionally, the amino-terminal tail of L2 penetrates the endosomal membranes and encases the virion particle in a

membrane vesicle. During mitosis, the vesicles enter the nuclear space and associate with mitotic chromosomes until the nucleus reforms. The virions are next observed at the PML nuclear bodies, owing to interactions with L2 and the PML protein[68]. The novel strategy of encasing the incoming virus in membrane vesicles enables the virus to escape detection by cyclic GMP–AMP synthase–stimulator of interferon genes (cGAS–STING)[31]. L2 is expressed again late in infection, at which point it assists in packaging viral minichromosomes into virion particles, but very little is known about this process in differentiated cells.

The L1 protein, expressed late in infection, encapsidates the viral minichromosome and L2 protein to form the icosahedral PV virions. L1 proteins readily self-assemble into icosahedral virus-like particles and are the source of the highly successful HPV vaccine[69].

Course of infection

Viruses from each genus have a different course of infection over the lifespan of their human host (FIG. 4). This is due, in part, to the period of initial exposure, but also to the ability of the HPV virus type to cause chronic infection, and to be controlled and cleared by the host immune system.

Viruses from the Betapapillomavirus and Gammapapillomavirus genera cause asymptomatic infections, and individuals are commonly infected in infancy[5]. Individuals can harbour numerous different viruses that persist for many years and are commonly shared by family members[4,70]. Thus, these viral types can be considered commensals and part of the skin microbiome. However, it is noteworthy that individuals with certain inherited immunodeficiencies are highly susceptible to HPV infection, and many novel HPVs from the Betapapillomavirus

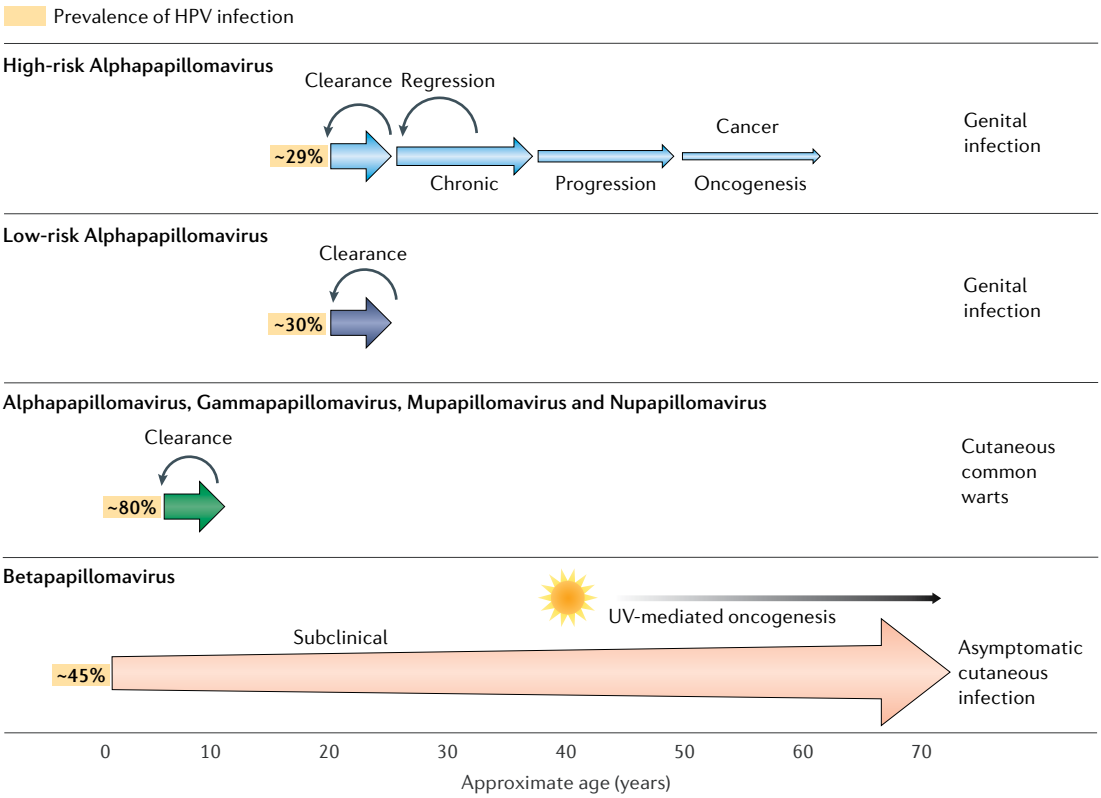


Fig. 4 | **Natural course of infection by different human papillomaviruses.** Timeline showing average course of different types of human papillomavirus (HPV) infection. For each example, HPV genus is shown on the left and the type and location of infection on the right. Position and width of each bar approximates the timing and duration of infection across the lifetime of the host, and the height of the bar represents the prevalence of infection in the general population. Most Alphapapillomaviruses are sexually transmitted with a high frequency of infection at sexual debut. In 2003–2004 (pre-vaccine), prevalence of genital HPV infection (as detected by viral DNA) was highest (~48%) in females 20–25 years old in the United States. In this age group, ~30% infections were low-risk HPV and ~29% were high-risk HPV. In most individuals, both low-risk HPV and high-risk HPV infections are cleared by the immune system within 2 years[133]. About 10% of cervical infections are not cleared and become chronic and susceptible to oncogenesis owing to long-term expression of E6 and E7 oncoproteins. Oncogenic progression is classified in stages as cervical intraepithelial neoplasia 1 (CIN1)–CIN3 and approximately one-third of CIN3 infections will progress to invasive carcinomas within 10–20 years[78]. Common warts are prevalent on hands and feet of elementary and primary school children and can be caused by HPV types from several genera, but most will spontaneously regress within a few months. One study showed that HPV DNA from these types could be detected on ~80% of elementary/primary school children, with 44% harbouring visible common warts[75]. Betapapillomaviruses cause asymptomatic infection and can be considered commensals. HPV DNA can be detected in ~45% of infants[5], and this prevalence increases over the lifespan to ~80% in healthy individuals[4,70]. Infection by some Betapapillomavirus HPV types may predispose to the development of UV-associated squamous cell skin cancer[73].

and Gammapapillomavirus genera have been isolated from such individuals[6,7]. Skin cancer is increased in individuals with immunodeficiencies, and in recipients of organ transplants who are immunosuppressed[71–73]. Based on these findings, it has been proposed that some commensal Betapapillomavirus HPVs can act as a cofactor with UV exposure to promote and predispose to squamous cell skin cancer, particularly in individuals with immunodeficiencies[73]. A provocative study has recently proposed that T cell immunity against commensal HPVs in healthy individuals actually prevents skin cancer by detection of non-self viral antigens[74], but the conclusions of this study have been contested[9].

Exposure to viruses that cause common warts (caused by Mupapillomaviruses, Nupapillomaviruses and some Alphapapillomaviruses) occurs commonly in schools. Common warts are very common in elementary school-aged children, with one study showing that ~44% of Dutch primary school children had cutaneous warts, whereas ~80% harboured DNA from related HPV types on normal skin[75]. Common warts are usually located on the hands (HPV-2, HPV-27 and HPV-57 from Alphapapillomavirus species 4 and HPV-4 and HPV-65 from Gammapapillomavirus species 1) and the plantar surface of the feet (HPV-1 and HPV-63 from the Mupapillomavirus genus). If untreated, common warts regress spontaneously within a few months in most individuals.

For the most part, Alphapapillomaviruses are sexually transmitted with first exposure in early adulthood. Although classified as mucosal, HPV-6 and HPV-11 cause condyloma acuminata (anogenital warts) in both the skin and mucosal epithelia surrounding the genital area[76]. These benign growths often persist for many months before regression, and sometimes remain for many years or recur after ablation[77].

Cervical infection with high-risk Alphapapillomavirus HPVs results in mild abnormal changes in cervical cells and is detected in about one-third of young women within 2 years of sexual debut. About 90% of infections will regress within 2 years; otherwise they become chronic, and infected individuals are at increased risk for cervical intraepithelial neoplasia (CIN) progression[78]. During oncogenic progression, the proportion of proliferating cells steadily increases until they occupy the entire thickness of the epithelium; these stages are defined as CIN1–CIN3. Within 10–20 years, approximately one-third of CIN3 infections will progress to invasive carcinoma[78] (FIG. 2d).

Immune control of HPV infection

Immune detection and clearance. The differentiation-dependent strategy of the HPV life cycle is one of the key elements of immune evasion: the entire life cycle takes place within epithelial cells; there is no cell lysis; and high levels of viral DNA and proteins are present only in the most differentiated layers of cells that are destined to be sloughed from the surface, far from immune defences. Upon entry, the virus traffics to the nucleus in endosomal vesicles to escape innate immune surveillance[31,79]. The early stages of infection take place at PML nuclear bodies, an attractive destination for many DNA viruses, but with antiviral properties that can restrict early infection[34,80,81]. However, the viral L2 protein counteracts this by displacing the repressive components of the nuclear bodies[82]. Moreover, early viral proteins (especially from the oncogenic HPVs) disrupt innate immune signalling pathways and delay adaptive immunity[83].

In most HPV infections, the cell-mediated immune system eventually recognizes the infected cells, and T cells infiltrate the lesion to resolve the infection[84]. Cell-mediated responses also provide protection against infection with specific HPV types as the humoral immune response is not robust, and many individuals do not seroconvert after natural infection[84,85]. This is in contrast to the humoral response to the HPV vaccine (see below), which produces extremely high levels of antibodies that are exuded in serum and mucus at sites of infection[84,86].

Deficiencies in immune control of HPV infection. Although HPVs have multiple ways to avoid complete clearance by the host immune system, the immune system is also critically important in controlling the infection, thus maintaining a balance between virus and host. Individuals with certain primary immunodeficiencies are highly susceptible to pathological infection by specific HPV types that are well controlled in immunocompetent individuals[72,87]. Each of these deficiencies highlights the many ways in which the immune system controls infection by different HPV types. For instance, individuals with epidermodysplasia verruciformis have recessive mutations in the genes encoding one of three proteins (EVER1, EVER2 and CIB) that form a complex that restricts Betapapillomavirus infection[88]. In the absence of this functional restrictive complex, individuals have disseminated plaques of viral infection with a high propensity to develop into squamous cell skin cancer on sun-exposed areas of skin, which led the International Agency for Research on Cancer (IARC) to classify two Betapapillomavirus types as possibly carcinogenic (Group 2). Organ transplant recipients undergoing immunosuppressive treatment have a 60-fold to 250-fold increase in the incidence of squamous cell skin cancer, further implicating Betapapillomaviruses in carcinogenesis[27]. In patients with WHIM (warts, hypogammaglobulinaemia, infections and myelokathexis) syndrome, a dominant mutation in the chemokine receptor gene *CXCR4* causes retention of neutrophils in the bone marrow (myelokathexis) and extreme susceptibility to both cutaneous and anogenital warts[89]. Mutations in *GATA2* (encodes a haematopoietic transcription factor) or *DOCK8* (AR dedicator of cytokinesis 8) result in severe impairment in the control of HPV infection[87]. When infected by HPV-2 or HPV-4, individuals with inherited CD28 deficiency result in severe 'tree man' syndrome but are otherwise healthy[90]. Individuals with mutations in the Fanconi anaemia DNA repair pathway are highly susceptible to HPV infections and to cancers in sites normally associated with oncogenic HPVs. However, these cancers are often HPV-negative and it has been proposed that the HPV oncogenes promote genetic instability in cells defective

in DNA repair, thus rendering them highly susceptible to cellular mutations that are eventually independent of HPV[91]. Very rarely, infection of the larynx with the low-risk Alphapapillomaviruses HPV-6 and HPV-11 (which normally cause anogenital warts) results in recurrent respiratory papillomatosis in young children. Individuals with respiratory papillomatosis are not particularly susceptible to other infections but, nevertheless, an underlying immune deficiency has been proposed to have a role[87].

Latency. Retroviruses and Alphaherpesviruses have defined latent stages in their life cycle, during which they either integrate into the host genome or lie dormant in a non-dividing cell, respectively. It is not clear whether HPVs frequently undergo true latency in which cells harbour low levels of viral DNA, without viral gene expression or virus production and escape immune detection, or whether such cells are under tight immune control that restricts the virus life cycle. Of note, a second peak of cervical infection with the same HPV type is sometimes observed in older women[78]. This could result from reinfection due to a failure to seroconvert after the primary infection but could also represent reactivation of a latent infection. Focal latent infection has recently been detected in the human cervix by whole tissue mapping[92]. Latency has also been demonstrated in rabbit papillomavirus models showing that cellular stress promotes reactivation[93–95]. Latent infection most likely occurs in infrequently dividing stem cells where the virus can persist with very little gene expression. The likelihood of latent infection occurring probably depends on many factors such as the presence of stem cells at the infection site and the balance between immune detection and evasion. Latency is likely most common in hair follicle stem cells infected with Betapapillomaviruses[96].

Mechanisms of oncogenesis

Almost all cases of cervical cancer are due to HPV infection, as are varying proportions of anal, vaginal, vulvar, rectal, penile and oropharyngeal cancers[97]. Twelve human Alphapapillomaviruses are classified by the IARC as oncogenic (Group 1) and, based either on weaker evidence or their position on the phylogenetic tree, an additional 13 as probably or possibly oncogenic (Groups 2A and 2B)[98] (FIG. 1b). The high-risk mucosotropic HPVs belong to Alphapapillomavirus species 5, 6, 7, 9 and 11 and the low-risk anogenital HPVs are found in Alphapapillomaviruses species 1, 3, 10 and 14 (FIG. 1b).

Betapapillomaviruses may also predispose to skin keratinocyte carcinomas, more specifically squamous cell carcinoma[27,99]. There are two different ways in which HPV infection can promote cancer (FIG. 5).

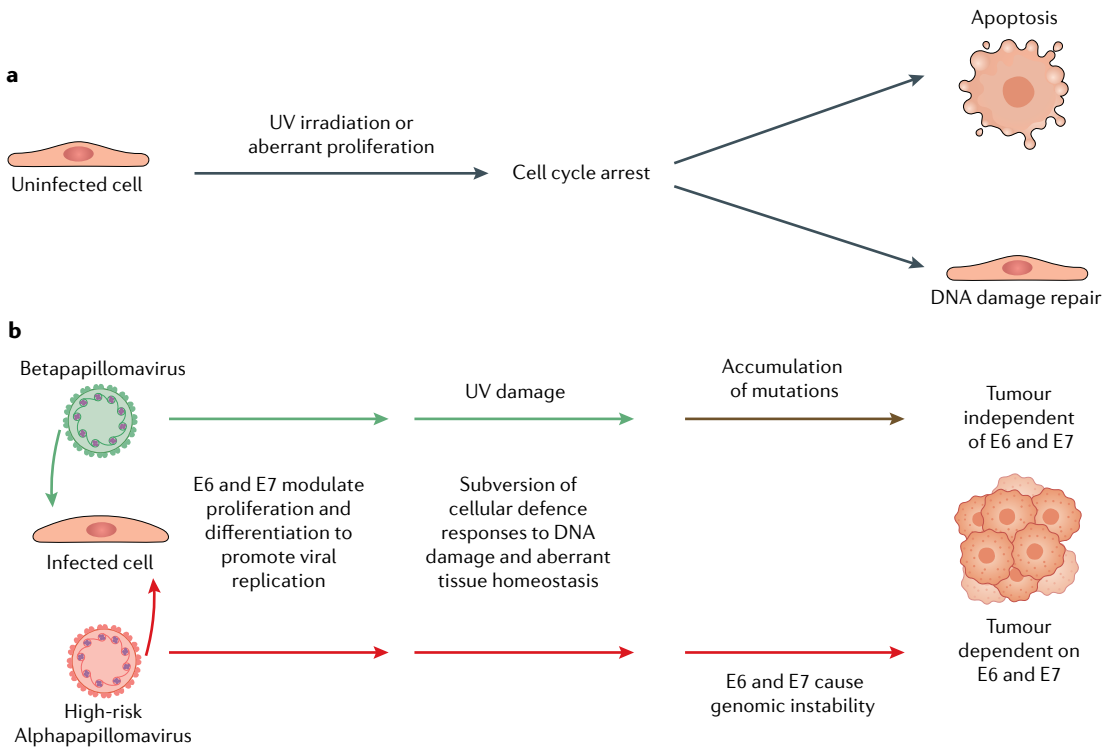


Fig. 5 | **Mechanisms of carcinogenesis. a** | How UV damage or aberrant proliferation leads to cell cycle arrest of uninfected cells and subsequent DNA repair or apoptosis. **b** | Comparison of infection by certain Betapapillomaviruses and high-risk Alphapapillomaviruses. In both cases, viral E6 and E7 proteins modulate host cell growth and differentiation to sustain infection. However, cellular defences that protect uninfected cells are dysregulated in infected cells and the cells either accumulate UV-mediated mutations (Betapapillomaviruses) or become genetically unstable owing to continued degradation of the p53 and Rb protein (pRb) tumour suppressors (Alphapapillomaviruses). E5 proteins are not required for carcinogenesis: Betapapillomaviruses do not encode E5 proteins, and E5 genes are often disrupted by integration of Alphapapillomavirus genomes.

It is well established that infection with the oncogenic Alphapapillomaviruses can result in anogenital and oropharyngeal cancers that are driven by, and dependent on, continued expression of the E6 and E7 oncoproteins. There is also growing evidence that some Betapapillomaviruses predispose to squamous cell skin cancer by a 'hit and run' mechanism. In this case, the virus interferes in the repair and response to UV damage. The resulting cellular mutations drive carcinogenesis but the tumours are no longer dependent on the virus[99] (FIG. 5).

Cells in epithelial transformation zones seem to be most susceptible to HPV infection and carcinogenesis[100]. Squamous cell carcinoma of the cervix could develop by at least two different courses. Classically, high-grade lesions and cancers are thought to develop and progress from the infected stratified epithelium of the ectocervix (FIG. 2d). However, cervical cancer frequently originates in the cervical transformation zone and it has been proposed that this results from infection of residual embryonic, squamocolumnar junction cells[100,101]. It is thought that viral gene expression is dysregulated in these cells and that they can develop directly into high-grade lesions without productive viral infection or a low-grade intermediate lesion[102,103]. The anorectal junction is also prone to HPV oncogenesis and contains similar cell populations[101]. Similarly, oropharyngeal cancers arise most often in the tonsil crypts, which are deep invaginations in the epithelium that increase surface area, are less well differentiated and are highly reticulated[104].

The overarching strategy of the virus is to establish a reservoir of persistently infected cells in the basal cells of a stratified epithelium, while also producing viral particles from the overlying differentiating cells. As each HPV type has adapted to a specific epithelial niche, the virus has fine-tuned the way in which it manipulates the cells in the tissue to achieve this goal, which in some cases has greatly increased the susceptibility of the host cells to DNA damage and carcinogenesis.

The oncogenic Alphapapillomavirus E6 and E7 proteins are both necessary and sufficient for virus-mediated carcinogenesis. All HPVs stimulate cell proliferation and delay differentiation in the upper layers of a lesion to promote vegetative viral DNA replication. However, oncogenic Alphapapillomavirus HPVs also disrupt this balance and inactivate cell cycle checkpoints in the less well differentiated lower layers[105], which leads to a high frequency of mutations (genetic instability) in the host genome and malignant progression. The distinguishing functions of the oncogenic HPV E6 and E7 proteins are strong immune evasion, degradation of pRB family members, p53 and PDZ-binding domain proteins, and induction of telomerase[105–108]. Furthermore, the HPV genome becomes integrated into the host genome (perhaps due to genetic instability of many cancers). Integration at favourable sites, such as those enriched in active enhancer elements, leads to deregulated expression of the viral oncogenes and further genetic instability[109]. Conversely, Betapapillomavirus E6 and E7 proteins inhibit the repair and response to UV-mediated DNA damage, leading to cellular mutations[110], but in the long term these proteins are not required to sustain the tumour phenotype.

Impact of HPV vaccines

Three highly successful HPV vaccines have been introduced worldwide since 2006, and they are projected to be on the national vaccine schedule of more than 150 countries by the end of 2021 (REF.[111]). These prophylactic bivalent, tetravalent and nonavalent vaccines consist of highly immunogenic virus-like particles formed by self-assembly of the L1 protein from two, four or nine HPV types, respectively (FIG. 1a). The highly ordered structure of the particles induces extremely high and durable levels of protective serum antibodies that are transuded to mucosal epithelial surfaces[86]. A meta-analytic study of 60 million individuals in high-income countries showed an 80% reduction in HPV-16 and HPV-18 infections (the most prevalent high-risk HPVs) and 70% reduction in anogenital wart diagnoses in the 8 years after vaccination[111,112]. In female-only vaccination programmes, herd effects were already observed in males[112]. It has been calculated that the current nonavalent vaccine has the potential to eliminate 90% of cases of cervical cancer and 50% of cancers that occur at sites associated with HPV infection, worldwide. However, some challenges remain[113]; for example, the implementation of a single-dose immunization, gender-neutral immunization, the manufacture of lower-cost vaccines and HPV vaccines bundled with other paediatric vaccines as well as increased vaccine uptake in countries where vaccination rates are suboptimal[113–115].

Concluding remarks

Papillomaviruses are masters of manipulation. During their long evolutionary association with their host they have adapted to and hijacked a myriad of cellular pathways to support and promote each stage of infection. Each disrupted or disturbed pathway represents engagement between the virus and host, and over time each has counteracted, modulated or tolerated the response of the other. What is remarkable, but poorly understood, is the diversity of HPVs and how each one has found a specific anatomical and ecological niche in the human epithelium to establish infection. For many decades, oncogenic viruses have helped to identify key host factors important for carcinogenesis[107].

The availability of hundreds of complete HPV DNA genome sequences[116] enables comparative genomics. Advances in protein structure prediction will also facilitate comparative HPV protein modelling to promote understanding of viral–host protein interactions in different viruses, and the design of antiviral therapeutics. Currently, single-cell analysis is often used for the study of many viral infections; however, papillomavirus infections are perfectly laid out, both temporally and spatially, in a stratified epithelium. Rapid advances in in situ microscope-based techniques in genomics, epigenetics, transcriptomics and proteomics will greatly increase our understanding of HPV infection at the single-cell level, both temporally and spatially, in different infected tissues.

Published online 14 September 2021

1. de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H. U. & zur Hausen, H. Classification of papillomaviruses. *Virology* **324**, 17–27 (2004).
2. Karamanou, M., Agapitos, E., Kousoulis, A. & Androutsos, G. From the humble wart to HPV: a fascinating story throughout centuries. *Oncol. Rev.* **4**, 133–135 (2010).
3. zur Hausen, H. Condylomata acuminata and human genital cancer. *Cancer Res.* **36**, 794–794 (1976). **This paper describes the association of HPV and cancer that led to the Nobel Prize in 2008**.
4. Antonsson, A., Forslund, O., Ekberg, H., Sterner, G. & Hansson, B. G. The ubiquity and impressive genomic diversity of human skin papillomaviruses suggest a commensalic nature of these viruses. *J. Virol.* **74**, 11636–11641 (2000). **This paper is the first indication that HPVs are ubiquitous and commensal**.
5. Antonsson, A., Karanfilovska, S., Lindqvist, P. G. & Hansson, B. G. General acquisition of human papillomavirus infections of skin occurs in early infancy. *J. Clin. Microbiol.* **41**, 2509–2514 (2003).
6. Tirosh, O. et al. Expanded skin virome in DOCK8-deficient patients. *Nat. Med.* **24**, 1815–1821 (2018).
7. Pastrana, D. V. et al. Metagenomic discovery of 83 new human papillomavirus types in patients with immunodeficiency. *mSphere* **3**, e00645-18 (2018). **Together with Tirosh et al. (2018), this paper presents the recent discovery of many new HPV types**.
8. de Martel, C., Plummer, M., Vignat, J. & Franceschi, S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int. J. Cancer* **141**, 664–670 (2017).
9. Lambert, P. F., Munger, K., Rosl, F., Hasche, D. & Tommasino, M. Beta human papillomaviruses and skin cancer. *Nature* **588**, E20–E21 (2020).
10. Van Doorslaer, K. Evolution of the Papillomaviridae. *Virology* **445**, 11–20 (2013).
11. Shah, S. D., Doorbar, J. & Goldstein, R. A. Analysis of host–parasite incongruence in papillomavirus evolution using importance sampling. *Mol. Biol. Evol.* **27**, 1301–1314 (2010).
12. Bravo, I. G. & Felez-Sanchez, M. Papillomaviruses: viral evolution, cancer and evolutionary medicine. *Evol. Med. Public Health* **2015**, 32–51 (2015).
13. Chen, Z. et al. Niche adaptation and viral transmission of human papillomaviruses from archaic hominins to modern humans. *PLoS Pathog.* **14**, e1007352 (2018).
14. Auslander, N., Wolf, Y. I., Shabalina, S. A. & Koonin, E. V. A unique insert in the genomes of high-risk human papillomaviruses with a predicted dual role in conferring oncogenic risk. *F1000Res* **8**, 1000 (2019).
15. Burk, R. D., Harari, A. & Chen, Z. Human papillomavirus genome variants. *Virology* **445**, 232–243 (2013).
16. Pimenoff, V. N., de Oliveira, C. M. & Bravo, I. G. Transmission between archaic and modern human ancestors during the evolution of the oncogenic human papillomavirus 16. *Mol. Biol. Evol.* **34**, 4–19 (2017).
17. Hirose, Y. et al. Within-host variations of human papillomavirus reveal APOBEC signature mutagenesis in the viral genome. *J. Virol.* **92**, e00017-18 (2018).
18. Moody, C. A. & Laimins, L. A. Human papillomaviruses activate the ATM DNA damage pathway for viral genome amplification upon differentiation. *PLoS Pathog.* **5**, e1000605 (2009). **This paper shows that HPVs activate the DNA damage response to replicate their genomes in differentiated cells**.
19. Egawa, N. & Doorbar, J. The low-risk papillomaviruses. *Virus Res.* **231**, 119–127 (2017).
20. Rowson, K. E. & Mahy, B. W. Human papova (wart) virus. *Bacteriol. Rev.* **31**, 110–131 (1967). **This paper is a fascinating documentation of early human transmission studies**.
21. Dreer, M. et al. Interaction of NCOR/SMRT repressor complexes with papillomavirus E8^E2C proteins inhibits viral replication. *PLoS Pathog.* **12**, e1005556 (2016). **This paper describes how HPV E8^E2 protein restricts replication of many HPV types**.
22. Zhou, J., Liu, W. J., Peng, S. W., Sun, X. Y. & Frazer, I. Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability. *J.Virol.* **73**, 4972–4982 (1999).
23. Hernandez-Alias, H.-A., Benisty, H., Schaefer, M. H. & Serrano, L. Translational adaptation of human viruses to the tissues they infect. *Cell* **34**, 108872 (2021).
24. Roberts, J. N. et al. Genital transmission of HPV in a mouse model is potentiated by nonoxynol-9 and inhibited by carrageenan. *Nat. Med.* **13**, 857–861 (2007).
25. Day, P. M. & Schelhaas, M. Concepts of papillomavirus entry into host cells. *Curr. Opin. Virol.* **4**, 24–31 (2014).
26. Herfs, M., Soong, T. R., Delvenne, P. & Crum, C. P. Deciphering the multifactorial susceptibility of mucosal junction cells to HPV infection and related carcinogenesis. *Viruses* **9**, 85 (2017).
27. Quint, K. D. et al. Human Beta-papillomavirus infection and keratinocyte carcinomas. *J. Pathol.* **235**, 342–354 (2015).
28. Ryser, M. D., Myers, E. R. & Durrett, R. HPV clearance and the neglected role of stochasticity. *PLoS Comput. Biol.* **11**, e1004113 (2015).
29. Xie, J., Zhang, P., Crite, M. & DiMaio, D. Papillomaviruses go retro. *Pathogens* **9**, 267 (2020).
30. DiGiuseppe, S. et al. Incoming human papillomavirus type 16 genome resides in a vesicular compartment throughout mitosis. *Proc. Natl Acad. Sci. USA* **113**, 6289–6294 (2016).
31. Uhlorn, B. L. et al. Vesicular trafficking permits evasion of cGAS/STING surveillance during initial human papillomavirus infection. *PLoS Pathog.* **16**, e1009028 (2020).
32. Pyeon, D., Pearce, S. M., Lank, S. M., Ahlquist, P. & Lambert, P. F. Establishment of human papillomavirus infection requires cell cycle progression. *PLoS Pathog.* **5**, e1000318 (2009).
33. Aydin, I. et al. Large scale RNAi reveals the requirement of nuclear envelope breakdown for nuclear import of human papillomaviruses. *PLoS Pathog.* **10**, e1004162 (2014).
34. Guion, L., Bienkowska-Haba, M., DiGiuseppe, S., Florin, L. & Sapp, M. PML nuclear body-residing proteins sequentially associate with HPV genome after infectious nuclear delivery. *PLoS Pathog.* **15**, e1007590 (2019).
35. Day, P. M., Baker, C. C., Lowy, D. R. & Schiller, J. T. Establishment of papillomavirus infection is enhanced by promyelocytic leukemia protein (PML) expression. *Proc. Natl Acad. Sci. USA* **101**, 14252–14257 (2004).
36. Corpet, A. et al. PML nuclear bodies and chromatin dynamics: catch me if you can! *Nucleic Acids Res.* **48**, 11890–11912 (2020).
37. Hofmann, S., Stubbe, M., Mai, J. & Schreiner, S. Double-edged role of PML nuclear bodies during human adenovirus infection. *Virus Res.* **295**, 198280 (2021).
38. Ozbun, M. A. Human papillomavirus type 31b infection of human keratinocytes and the onset of early transcription. *J. Virol.* **76**, 11291–11300 (2002).
39. Coursey, T. L. & McBride, A. A. Hitchhiking of viral genomes on cellular chromosomes. *Annu. Rev. Virol.* **6**, 275–296 (2019).
40. White, E. A. Manipulation of epithelial differentiation by HPV oncoproteins. *Viruses* **11**, 369 (2019).
41. Moore, P. S. & Chang, Y. Why do viruses cause cancer? Highlights of the first century of human tumour virology. *Nat. Rev. Cancer* **10**, 878–889 (2010).
42. Xue, Y. et al. HPV16 E2 is an immediate early marker of viral infection, preceding E7 expression in precursor structures of cervical carcinoma. *Cancer Res.* **70**, 5316–5325 (2010).
43. Klumpp, D. J. & Laimins, L. A. Differentiation-induced changes in promoter usage for transcripts encoding the human papillomavirus type 31 replication protein E1. *Virology* **257**, 239–246 (1999).
44. Johansson, C. & Schwartz, S. Regulation of human papillomavirus gene expression by splicing and polyadenylation. *Nat. Rev. Microbiol.* **11**, 239–251 (2013).
45. Sakakibara, N., Chen, D. & McBride, A. A. Papillomaviruses use recombination-dependent replication to vegetatively amplify their genomes in differentiated cells. *PLoS Pathog.* **9**, e1003321 (2013).
46. Pyeon, D., Lambert, P. F. & Ahlquist, P. Production of infectious human papillomavirus independently of viral replication and epithelial cell differentiation. *Proc. Natl Acad. Sci. USA* **102**, 9311–9316 (2005).
47. Porter, S. S. et al. Histone modifications in papillomavirus virion minichromosomes. *mBio* **12**, e03274-20 (2021).
48. Bryan, J. T. & Brown, D. R. Transmission of human papillomavirus type 11 infection by desquamated cornified cells. *Virology* **281**, 35–42 (2001).
49. Egawa, N. et al. Dynamics of papillomavirus in vivo disease formation & susceptibility to high-level disinfection — implications for transmission in clinical settings. *EbioMed.* **63**, 103177 (2021).
50. Roden, R. B. S., Lowy, D. R. & Schiller, J. T. Papillomavirus is resistant to desiccation. *J. Infect. Dis.* **176**, 1076–1079 (1997).
51. Ozbun, M. A. et al. Infectious titres of human papillomaviruses (HPVs) in patient lesions, methodological considerations in evaluating HPV infectivity and implications for the efficacy of high-level disinfectants. *EbioMed.* **63**, 103165 (2021).
52. Spurgeon, M. E. et al. A novel in vivo infection model to study papillomavirus-mediated disease of the female reproductive tract. *mBio* **10**, e00180-19 (2019).
53. Bravo, I. G. & Alonso, A. Mucosal human papillomaviruses encode four different E5 proteins whose chemistry and phylogeny correlate with malignant or benign growth. *J. Virol.* **78**, 13613–13626 (2004).
54. Bergvall, M., Melendy, T. & Archambault, J. The E1 proteins. *Virology* **445**, 35–56 (2013).
55. Wu, S. C. et al. in *Human Papillomavirus: Proving and Using a Viral Cause for Cancer* (eds Jenkins, D. & Bosch, F. X.) 53–65 (Academic, 2020).
56. Dreer, M., van de Poel, S. & Stubenrauch, F. Control of viral replication and transcription by the papillomavirus E8^E2 protein. *Virus Res.* **231**, 96–102 (2017).
57. Doorbar, J., Campbell, D., Grand, R. J. & Gallimore, P. H. Identification of the human papilloma virus-1a E4 gene products. *EMBO J.* **5**, 355–362 (1986).
58. Doorbar, J., Ely, S., Sterling, J., McLean, C. & Crawford, L. Specific interaction between HPV-16 E1–E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network. *Nature* **352**, 824–827 (1991).
59. Doorbar, J. The E4 protein; structure, function and patterns of expression. *Virology* **445**, 80–98 (2013).
60. DiMaio, D. & Petti, L. M. The E5 proteins. *Virology* **445**, 99–114 (2013).
61. Scott, M. L. et al. Human papillomavirus 16 E5 inhibits interferon signaling and supports episomal viral maintenance. *J. Virol.* **94**, e01582-19 (2020).
62. Roman, A. & Munger, K. The papillomavirus E7 proteins. *Virology* **445**, 138–168 (2013).
63. Brimer, N., Drews, C. M. & Vande Pol, S. B. Association of papillomavirus E6 proteins with either MAML1 or E6AP clusters E6 proteins by structure, function, and evolutionary relatedness. *PLoS Pathog.* **13**, e1006781 (2017).
64. Katzenellenbogen, R. Telomerase induction in HPV infection and oncogenesis. *Viruses* **9**, 180 (2017).
65. Brimer, N., Lyons, C., Wallberg, A. E. & Vande Pol, S. B. Cutaneous papillomavirus E6 oncoproteins associate with MAML1 to repress transactivation and NOTCH signaling. *Oncogene* **31**, 4639–4646 (2012).
66. Jackson, S. & Storey, A. E6 proteins from diverse cutaneous HPV types inhibit apoptosis in response to UV damage. *Oncogene* **19**, 592–598 (2000).
67. Campos, S. K. Subcellular trafficking of the papillomavirus genome during initial infection: the remarkable abilities of minor capsid protein L2. *Viruses* **9**, 370 (2017).
68. DiGiuseppe, S., Bienkowska-Haba, M., Guion, L. G. & Sapp, M. Cruising the cellular highways: how human papillomavirus travels from the surface to the nucleus. *Virus Res.* **231**, 1–9 (2017).
69. Buck, C. B., Day, P. M. & Trus, B. L. The papillomavirus major capsid protein L1. *Virology* **445**, 169–174 (2013).
70. Hsu, J. Y., Chen, A. C., Keleher, A., McMillan, N. A. & Antonsson, A. Shared and persistent asymptomatic cutaneous human papillomavirus infections in healthy skin. *J. Med. Virol.* **81**, 1444–1449 (2009).
71. Birkeland, S. A. et al. Cancer risk after renal transplantation in the Nordic countries, 1964–1986. *Int. J. Cancer* **60**, 183–189 (1995).
72. Leiding, J. W. & Holland, S. M. Warts and all: human papillomavirus in primary immunodeficiencies. *J. Allergy Clin. Immun.* **130**, 1030–1048 (2012).
73. Rollison, D. E., Viarisio, D., Amorrortu, R. P., Gheit, T. & Tommasino, M. An emerging issue in oncogenic virology: the role of beta human papillomavirus types in the development of cutaneous squamous cell carcinoma. *J. Virol.* **93**, e01003-18 (2019).
74. Strickley, J. D. et al. Immunity to commensal papillomaviruses protects against skin cancer. *Nature* **575**, 519–522 (2019).
75. de Koning, M. N. et al. High prevalence of cutaneous warts in elementary school children and the ubiquitous presence of wart-associated human papillomavirus on clinically normal skin. *Br. J. Dermatol.* **172**, 196–201 (2015).

76. Egawa, N., Egawa, K., Griffin, H. & Doorbar, J. Human papillomaviruses; epithelial tropisms, and the development of neoplasia. *Viruses* **7**, 3863–3890 (2015).
77. Steben, M. & Garland, S. M. Genital warts. *Best Pract. Res. Clin. Obstet. Gynaecol.* **28**, 1063–1073 (2014).
78. Gravitt, P. E. The known unknowns of HPV natural history. *J. Clin. Invest.* **121**, 4593–4599 (2011).
79. DiGiuseppe, S., Bienkowska-Haba, M., Guion, L. G. M., Keiffer, T. R. & Sapp, M. Human papillomavirus major capsid protein L1 remains associated with the incoming viral genome throughout the entry process. *J. Virol.* **91**, e00537-17 (2017).
80. Guion, L. G. & Sapp, M. The role of promyelocytic leukemia nuclear bodies during HPV infection. *Front. Cell Infect. Microbiol.* **10**, 35 (2020).
81. Stepp, W. H., Meyers, J. M. & McBride, A. A. Sp100 provides intrinsic immunity against human papillomavirus infection. *mBio* **4**, e00845-13 (2013).
82. Florin, L., Schafer, F., Sotlar, K., Streeck, R. E. & Sapp, M. Reorganization of nuclear domain 10 induced by papillomavirus capsid protein l2. *Virology* **295**, 97–107 (2002).
83. Ferreira, A. R., Ramalho, A. C., Marques, M. & Ribeiro, D. The interplay between antiviral signalling and carcinogenesis in human papillomavirus infections. *Cancers (Basel)* **12**, 646 (2020).
84. Roden, R. B. S. & Stern, P. L. Opportunities and challenges for human papillomavirus vaccination in cancer. *Nat. Rev. Cancer* **18**, 240–254 (2018).
85. Gravitt, P. E. Unraveling the epidemiology of oral human papillomavirus infection. *Ann. Intern. Med.* **167**, 748–749 (2017).
86. Schiller, J. & Lowy, D. Explanations for the high potency of HPV prophylactic vaccines. *Vaccine* **36**, 4768–4773 (2018).
This paper explains why the HPV vaccine is so immunogenic, and successful.
87. Beziat, V. Human genetic dissection of papillomavirus-driven diseases: new insight into their pathogenesis. *Hum. Genet.* **139**, 919–939 (2020).
88. Venuti, A., Lohse, S., Tommasino, M. & Smola, S. Cross-talk of cutaneous beta human papillomaviruses and the immune system: determinants of disease penetrance. *Philos. Trans. R. Soc. London Ser. B, Biol. Sci.* **374**, 20180287 (2019).
89. McDermott, D. H. & Murphy, P. M. WHIM syndrome: immunopathogenesis, treatment and cure strategies. *Immunol. Rev.* **287**, 91–102 (2019).
90. Beziat, V. et al. Humans with inherited T cell CD28 deficiency are susceptible to skin papillomaviruses but are otherwise healthy. *Cell* **184**, 3812–3828.e30 (2021).
This paper identifies deficiency in the T cell co-stimulatory molecule CD28 as responsible for the 'tree man' syndrome.
91. Park, S., Park, J. W., Pitot, H. C. & Lambert, P. F. Loss of dependence on continued expression of the human papillomavirus 16 E7 oncogene in cervical cancers and precancerous lesions arising in fanconi anemia pathway-deficient mice. *mBio* **7**, e00628-16 (2016).
92. Hammer, A. et al. Whole tissue cervical mapping of HPV infection: molecular evidence for focal latent HPV infection in humans. *Papillomavirus Res.* **7**, 82–87 (2019).
This paper presents an analysis of HPV latency in human cervical infection.
93. Doorbar, J. Latent papillomavirus infections and their regulation. *Curr. Opin. Virol.* **3**, 416–421 (2013).
94. Amella, C. A. et al. Latent infection induced with cottontail rabbit papillomavirus. A model for human papillomavirus latency. *Am. J. Pathol.* **144**, 1167–1171 (1994).
95. Maglennon, G. A., McIntosh, P. & Doorbar, J. Persistence of viral DNA in the epithelial basal layer suggests a model for papillomavirus latency following immune regression. *Virology* **414**, 153–163 (2011).
96. Lanfredini, S. et al. HPV8 field cancerization in a transgenic mouse model is due to Lrig1$^+$ keratinocyte stem cell expansion. *J. Investig. Dermatol.* **137**, 2208–2216 (2017).
97. Viens, L. J. et al. Human papillomavirus-associated cancers — United States, 2008–2012. *MMWR* **65**, 661–666 (2016).
98. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Biological agents. Volume 100 B. A review of human carcinogens. *IARC Monogr. Eval. Carcinog. Risks. Hum.* **100**, 1–441 (2012).
99. Howley, P. M. & Pfister, H. J. Beta genus papillomaviruses and skin cancer. *Virology* **479–480**, 290–296 (2015).
100. Herfs, M. et al. A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer. *Proc. Natl Acad. Sci. USA* **109**, 10516–10521 (2012).
101. Yang, E. J. et al. Microanatomy of the cervical and anorectal squamocolumnar junctions: a proposed model for anatomical differences in HPV-related cancer risk. *Mod. Pathol.* **28**, 994–1000 (2015).
102. Doorbar, J. & Griffin, H. Refining our understanding of cervical neoplasia and its cellular origins. *Papillomavirus Res.* **7**, 176–179 (2019).
103. Reich, O. & Regauer, S. Thin HSIL of the cervix: detecting a variant of high-grade squamous intraepithelial lesions with a p16INK4a antibody. *Int. J. Gynecol. Pathol.* **36**, 71–75 (2017).
104. Roberts, S., Evans, D., Mehanna, H. & Parish, J. L. Modelling human papillomavirus biology in oropharyngeal keratinocytes. *Philos. Trans. R. Soc. London B Biol. Sci.* **374**, 20180289 (2019).
105. Doorbar, J., Egawa, N., Griffin, H., Kranjec, C. & Murakami, I. Human papillomavirus molecular biology and disease association. *Rev. Med. Virol.* **25** (Suppl 1), 2–23 (2015).
106. Klingelhutz, A. J. & Roman, A. Cellular transformation by human papillomaviruses: lessons learned by comparing high- and low-risk viruses. *Virology* **424**, 77–98 (2012).
107. Mesri, E. A., Feitelson, M. A. & Munger, K. Human viral oncogenesis: a cancer hallmarks analysis. *Cell Host Microbe* **15**, 266–282 (2014).
108. Mittal, S. & Banks, L. Molecular mechanisms underlying human papillomavirus E6 and E7 oncoprotein-induced cell transformation. *Mutat. Res.* **772**, 23–35 (2017).
109. McBride, A. A. & Warburton, A. The role of integration in oncogenic progression of HPV-associated cancers. *PLoS Pathog.* **13**, e1006211 (2017).
110. Tommasino, M. HPV and skin carcinogenesis. *Papillomavirus Res.* **7**, 129–131 (2019).
111. de Sanjose, S., Brotons, M., LaMontagne, D. S. & Bruni, L. Human papillomavirus vaccine disease impact beyond expectations. *Curr. Opin. Virol.* **39**, 16–22 (2019).
112. Drolet, M. et al. Population-level impact and herd effects following the introduction of human papillomavirus vaccination programmes: updated systematic review and meta-analysis. *Lancet* **394**, 497–509 (2019).
113. Stern, P. L. & Roden, R. B. Opportunities to improve immune-based prevention of HPV-associated cancers. *Papillomavirus Res.* **7**, 150–153 (2019).
114. Walling, E. B. et al. Interventions to improve HPV vaccine uptake: a systematic review. *Pediatrics* https://doi.org/10.1542/peds.2015-3863 (2016).
115. Robles, C. et al. Determinants of human papillomavirus vaccine uptake by adult women attending cervical cancer screening in 9 European countries. *Am. J. Prev. Med.* https://doi.org/10.1016/j.amepre.2020.08.032 (2020).
116. Van Doorslaer, K. et al. The Papillomavirus Episteme: a major update to the papillomavirus sequence database. *Nucleic Acids Res.* **45**, D499–D506 (2017).
This paper describes the Papillomavirus Episteme, the HPV sequence database.
117. Hughes, F. J. & Romanos, M. A. E1 protein of human papillomavirus is a DNA helicase/ATPase. *Nucleic Acids Res.* **21**, 5817–5823 (1993).
118. Ustav, M., Ustav, E., Szymanski, P. & Stenlund, A. Identification of the origin of replication of bovine papillomavirus and characterization of the viral origin recognition factor E1. *EMBO J.* **10**, 4321–4329 (1991).
119. Androphy, E. J., Lowy, D. R. & Schiller, J. T. Bovine papillomavirus E2 *trans*-activating gene product binds to specific sites in papillomavirus DNA. *Nature* **325**, 70–73 (1987).
120. Melendy, T., Sedman, J. & Stenlund, A. Cellular factors required for papillomavirus DNA replication. *J.Virol.* **69**, 7857–7867 (1995).
121. Ustav, M. & Stenlund, A. Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames. *EMBO J.* **10**, 449–457 (1991).
122. Mohr, I. J. et al. Targeting the E1 replication protein to the papillomavirus origin of replication by complex formation with the E2 transactivator. *Science* **250**, 1694–1699 (1990).
123. Skiadopoulos, M. H. & McBride, A. A. Bovine papillomavirus type 1 genomes and the E2 transactivator protein are closely associated with mitotic chromatin. *J. Virol.* **72**, 2079–2088 (1998).
124. McBride, A. A. The papillomavirus E2 proteins. *Virology* **445**, 57–79 (2013).
125. Piirsoo, M., Ustav, E., Mandel, T., Stenlund, A. & Ustav, M. *Cis* and *trans* requirements for stable episomal maintenance of the BPV-1 replicator. *EMBO J.* **15**, 1–11 (1996).
126. Davy, C. E. et al. Identification of a G2 arrest domain in the E1^E4 protein of human papillomavirus type 16. *J. Virol.* **76**, 9806–9818 (2002).
127. Moody, C. A. Impact of replication stress in human papillomavirus pathogenesis. *J. Virol.* https://doi.org/10.1128/JVI.01012-17 (2019).
128. Huibregtse, J. M., Scheffner, M. & Howley, P. M. A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. *EMBO J.* **10**, 4129–4135 (1991).
129. Adelstein, D. et al. NCCN Guidelines Insights: head and neck cancers, version 2.2017. *J. Natl Compr. Cancer Netw.* **15**, 761–770 (2017).
130. Buck, C. B. et al. Arrangement of L2 within the papillomavirus capsid. *J. Virol.* **82**, 5190–5197 (2008).
131. Wang, J. W. & Roden, R. B. L2, the minor capsid protein of papillomavirus. *Virology* **445**, 175–186 (2013).
132. Cubie, H. A. Diseases associated with human papillomavirus infection. *Virology* **445**, 21–34 (2013).
This paper is a highly cited review of HPV-associated disease.
133. Moscicki, A. B. et al. The natural history of human papillomavirus infection as measured by repeated DNA testing in adolescent and young women. *J. Pediatr.* **132**, 277–284 (1998).
134. Gissmann, L., Pfister, H. & Zur Hausen, H. Human papilloma viruses (HPV): characterization of four different isolates. *Virology* **76**, 569–580 (1977).
135. zur Hausen, H., Meinhof, W., Scheiber, W. & Bornkamm, G. W. Attempts to detect virus-specific DNA in human tumors. I. Nucleic acid hybridizations with complementary RNA of human wart virus. *Int. J. Cancer.* **13**, 650–656 (1974).

Acknowledgements
Work in the author's laboratory is funded by the Intramural Research Programs of the National Institute of Allergy and Infectious Diseases (NIAID) (grant number ZIA1000713). The author thanks R. Kissinger (NIAID) for drawing a draft for Fig. 2.

Competing interests
The author declares no competing interests.

Peer review information
Nature Reviews Microbiology thanks J. Doorbar, S. Graham and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

RELATED LINKS
International Agency for Research on Cancer: https://www.iarc.who.int
International Committee on Taxonomy of Viruses: https://talk.ictvonline.org
International HPV Reference Center: https://ki.se/en/labmed/international-hpv-reference-center
Papillomavirus Episteme: https://pave.niaid.nih.gov