Human papillomavirus: an overview

Sanjiv Rughooputh
Pamela Greenwell
School of Biosciences

This is an electronic version of an article published in The Biomedical Scientist, 47, pp. 1070-1071, October 2003.

ARTICLE

The result of human papillomavirus infection was recorded over 2000 years ago yet confirmation of its sexually transmitted nature only came in the 1950s. Here, Sanjiv Rughooputh and Pamela Greenwell look in a little more detail at this common infective agent.

Human papillomavirus: an overview

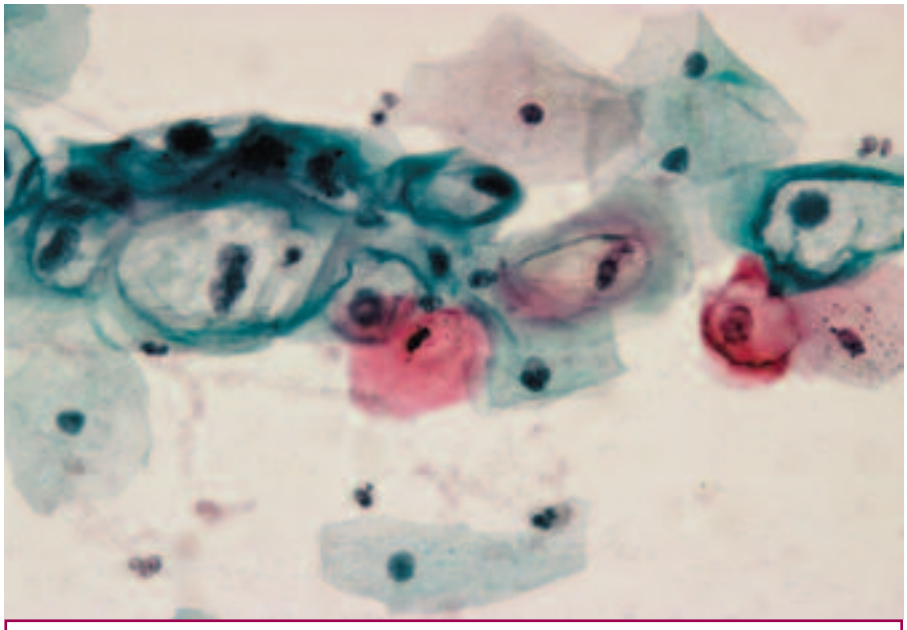
Human papillomavirus (HPV) comprises a group of heterogeneous viruses containing many genotypes that can be divided into high-risk and low-risk types, depending on their association with disease. Some 70 HPV types that can infect epithelial surfaces differentially have been identified, including a number that have been implicated in the aetiology of cervical cancer, the most common form of malignant tumour in women worldwide.

HPV infection, in the form of skin warts, was first recognised as long ago as 500 BC; however, there was no real interest in these viruses until they were linked with genital cancer. For centuries, transmission of HPV, characterised by the presence of genital warts, was associated with sexual intercourse, but it was only in 1954 that Barret, Silber and McGinley confirmed the sexual transmission of HPV in their study of American servicemen who developed penile warts after having intercourse with Korean women during the war in their country.

Feeling squamous

HPV, one of the two genera of papovaviridae, has the ability to infect and replicate in both keratinised and mucosal squamous epithelial cells. Most genotypes infect cutaneously but a smaller number are mucotropic, infecting the oral cavity and the urogenital tract (Fig. 1).

HPV can multiply on virtually any part of the body surface and their DNA may then integrate into the human genomic DNA. Although HPV infection is quite common, not all cases result in malignancy. In fact, infection can lead to a variety of problems that range from common warts and verruca to malignant disease such as cancer of cervix and larynx.



Eric Grave/Science Photo Library

Fig 1. Squamous cells scraped from the uterine cervix showing the mucotropic effects of HPV infection, the most notable being koilocytosis (halos) around the cell nuclei.

Open reading frames

HPV consists of a capsid that shows icosahedral symmetry and 72 capsomeres, each with an average diameter between 52–55 nm (Fig. 2). The virus contains double-stranded DNA (dsDNA) molecules with a genome size of 8 kb, coding for proteins with an estimated weight of $5x10^6$ Da. The molecular organisation of the HPV genome is well conserved between viruses of various types.

The open reading frame (ORF) is divided into areas that code for early (E) and late (L) proteins (Fig. 3). Owing to the overlapping nature of the ORFs, the messenger RNA (mRNA) transcribed is complex and it is not clear which transcript codes for which protein. The ORFs are situated on and are transcribed from the same strand and so are all read in the same direction, from 5' to 3'.

Early or late opening

At least seven proteins are encoded by the early region of the HPV genome, the most frequently studied being E6 and E7, which are important for the efficient immortalisation of human keratinocytes and are produced in the basal or parabasal cells.

E4 protein is produced suprabasally in differentiating cells and is found early in infection, but can accumulate in the late phase. E5 is a membrane-associated hydrophilic protein that has cell transformation potential – a role that is enhanced by the presence of epidermal growth factors.

E2 has two roles in the life cycle of HPV and it has been suggested that the disruption of the *E2* gene in high-risk HPV 16 is an important precursor in the process of cervical carcinogenesis. E2 can regulate transcription

> ‘Infection can lead to a variety of problems that range from common warts to malignant disease such as cancer of cervix and larynx’

from the early promoter and contains a transactivating domain at the 5’ end and a DNA binding domain at the 3’ end. However, the role of E2 during keratinocyte differentiation is poorly understood.

The late proteins L1 and L2 are the major and minor capsid proteins of the virion. L1 is found in the highly differentiated upper spinous layers. As these proteins are made in the epithelial cells, it is difficult for the immune system to recognise them.

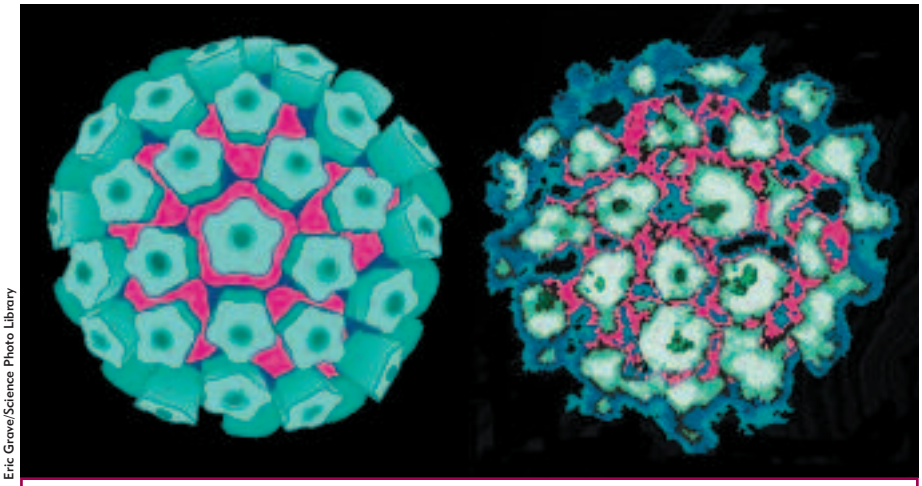
Process of infection

HPV penetrates the squamous epithelium by infecting the basal cell layer. Following entry into the host cell and uncoating, the viral genome migrates to the nucleus where transcription, DNA replication and virion assembly take place.

Early genes are copied from a single promoter and the transcripts subjected to differential splicing to generate mRNA for the seven early proteins. These include regulatory proteins, some with transactivating properties, that depress the genes for certain cellular enzymes and stimulate cellular DNA synthesis.

The viral genome replicates only very slowly as an autonomous plasmid in the nuclei of these cells. However, in the differentiated cells that comprise the outer layers of the squamous epithelium, full expression of the viral genome and DNA synthesis occur.

Late genes are transcribed from a second promoter. The late genes encode structural proteins L1 and L2, which, after various post-translational modifications, are directed to the nucleus via their nuclear localisation signals to be assembled into virions. ■



Eric Grave/Science Photo Library

Fig 2. Transmission electron micrograph of an HPV particle (right), together with computer artwork (left) showing the capsid in red, studded with surface proteins (blue).

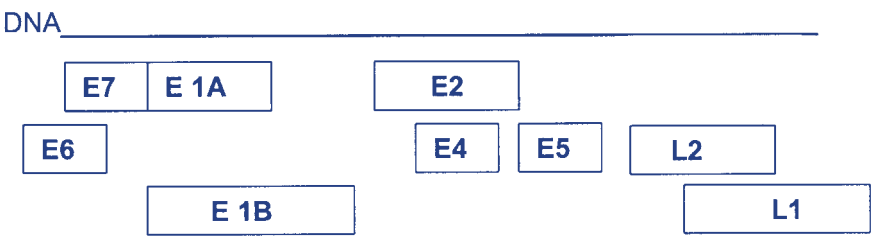


Fig 3. Schematic representation of the overlap seen in HPV early (E) and late (L) open reading frames.

Sanjiv Rughooputh and Pamela Greenwell are members of the Molecular and Medical Microbiology Research Group in the School of Biosciences, University of Westminster, London.