Université de Montréal

Determination of viral load and integration status of HPV 16 in normal and LSIL exfoliated cervical cells

par
Otelinda de Morais

Département de microbiologie et immunologie
Faculté de Médecine

Mémoire présenté à la Faculté de Médecine
en vue de l'obtention du grade de Mâitrise
en Microbiologie et Immunologie

Montreal, QC
Novembre 2009

University of Montreal
Faculty of Superior Studies

This entitled memory:

Determination of viral load and integration status of HPV 16 in normal and LSIL exfoliated cervical cells

Presented by:
Otelinda de Morais

Was evaluated by a jury composed from the following professors:

Dr. Claude Lemieux-président rapporteur
Dr. Francois Coutlée--directeur de recherche
Dre. Lise St-Jean-membre du jury

Memory accepted September 2010

RÉSUMÉ

L'intégration du génome du virus papilloma humain (VPH) a été reconnu jusqu''a récemment comme étant un événnement fréquent mais pourtant tardif dans la progression de la maladie du col de l'utérus. La perspective temporelle vient, pourtant, d'être mise au défi par la détection de formes intégrées de VPH dans les tissus normaux et dans les lésions prénéoplasiques.

Notre objectif était de déterminer la charge virale de VPH-16 et son état physique dans une série de 220 échantillons provenant de cols uterins normaux et avec des lésions de bas-grade. La technique quantitative de PCR en temps réel, méthode Taqman, nous a permis de quantifier le nombre de copies des gènes E6, E2, et de la B-globine, permettant ainsi l'évaluation de la charge virale et le ratio de E6/E2 pour chaque spécimen. Le ratio E6/E2 de 1.2 ou plus était suggestif d'intégration. Par la suite, le site d'intégration du VPH dans le génome humain a été déterminé par la téchnique de RS-PCR.

La charge virale moyenne était de 57.5±324.6 copies d'ADN par cellule et le ratio E6/E2 a évalué neuf échantillons avec des formes d'HPV intégrées. Ces intégrants ont été amplifiés par RS-PCR, suivi de séquençage, et l'homologie des amplicons a été déterminée par le programme BLAST de NCBI afin d'identifier les jonctions virales-humaines. On a réussi `a identifier les jonctions humaines-virales pour le contrôle positif, c'est-à-dire les cellules SiHa, pourtant nous n'avons pas detecté d'intégration par la technique de RS-PCR dans les échantillons de cellules

cervicales exfoliées provenant de tissus normaux et de lésions de bas-grade. Le VPH-16 est rarement intégré dans les spécimens de jeunes patientes.

Mots Clés : Virus Papilloma Humain, LSIL, Intégration HPV, Charge Virale, PCR en temps réel, RS-PCR, PCR-séquençage, HPV16.

ABSTRACT

Integration of human papillomavirus (HPV) has, until recently, been a frequent but late event in cervical carcinogenesis. The temporal view has, however, been challenged lately as integrated forms of HPV have been detected even in normal and preneoplastic lesions.

Our objective was to describe HPV 16 load and physical state in a series of 220 normal and low grade cervical samples. We used quantitative real-time PCR, Taqman method, targeting E6, E2 and B-globin to calculate the HPV 16 load and the E6/E2 ratio in each sample. An E6/E2 ratio of 1.2 was used as a surrogate marker of integration. The site of integration was determined by restriction site PCR.

Results show that the average viral load was 57.5±324.6 copies of DNA per cell, while E6/E2 ratio identified 9 samples with integrants. These integrants underwent amplification by restriction site PCR, followed by sequencing and nucleotide blast to identify the human-viral junctions. In conclusion, although it was possible to identify viral-host junctions with the integration positive control, that is, the SiHa cell line, the exfoliated cells of normal and low grade cervical lesions were negative for integration site by RS-PCR. HPV-16 is seldom integrated in specimens from young patients.

Key Words : Human Papillomavirus 16, LSIL, HPV integration, Viral load, real-time PCR, RS-PCR, PCR sequencing, HPV16.

SCIENTIFIC PUBLICATIONS
FROM THE AUTHOR

Impaired learning and LTP in mice expressing the carboxy terminus of the Alzheimer amyloid precursor protein.

Nalbantoglu J, Tirado-Santiago G, Lahsaïni A, Poirier J, Goncalves O, Verge G, Momoli F, Welner SA, Massicotte G, Julien JP, Shapiro ML.

Nature. 1997 May 29;387(6632):500-5.

Interleukin 1 enhances growth factor-dependent proliferation of the clonogenic cells in acute myeloblastic leukemia and of normal human primitive hemopoietic precursors.

Hoang T, Haman A, Goncalves O, Letendre F, Mathieu M, Wong GG, Clark SC.

J Exp Med. 1988 Aug 1;168(2):463-74.

Interleukin-6 enhances growth factor-dependent proliferation of the blast cells of acute myeloblastic leukemia.

Hoang T, Haman A, Goncalves O, Wong GG, Clark SC.

Blood. 1988 Aug;72(2):823-6.

The genetic consequences of the Thy- mutation to CHO cells.

Meuth M, Gonçalves O, Trudel M.

Basic Life Sci. 1985;31:297-312. Review.

Structural alterations of the aprt locus induced by deoxyribonucleoside triphosphate pool imbalances in Chinese hamster ovary cells.

Goncalves O, Drobetsky E, Meuth M.

Mol Cell Biol. 1984 Sep;4(9):1792-9

Structure of mutant alleles at the aprt locus of Chinese hamster ovary cells.

Nalbantoglu J, Goncalves O, Meuth M.

J Mol Biol. 1983 Jul 5;167(3):575-94.

A selection system specific for the Thy mutator phenotype.

Meuth M, Gonçalves O, Thom P.

Somatic Cell Genet. 1982 Jul;8(4):423-32.

TABLE OF CONTENTS

LIST OF FIGURES

Literature Review

Article

LIST OF TABLES

Literature Review

Article

LIST OF ABBREVIATIONS

AIDS:	Acquired immune deficiency syndrome
APOT:	Amplification of papillomavirus oncogene transcript
ATPase:	Adenosine triphosphatase
ASCUS:	Atypical squamous cells of undetermined significance
CIN:	Cervical intraepithelial neoplasia
CIS:	Carcinoma in situ
COPV:	Canine oral papillomavirus
DIPS:	Detection of integrated papillomavirus sequences
DNA:	Deoxyribonucleic acid
E6-AP:	E6 Associated protein
EGF:	Epidermal growth factor
EV:	Epidermodysplasia Verruciformis
GAPDH:	Glyceraldehyde 3-phosphate dehydrogenase
HIV:	Human immunodeficiency virus
HLA:	Human leukocyte antigen
HPV:	Human papillomavirus
HR-HPV:	High risk human papillomavirus
HR-E6:	High risk E6 protein
HSIL:	High-grade squamous intraepithelial lesions
hTERT:	Human telomerase reverse transcriptase
ICC:	Invasive cervical carcinoma
IFN:	Interferon
LSIL:	Low-grade squamous intraepithelial lesions
LCR:	Long control region
mRNA:	messenger ribonucleic acid
MHC:	major histocompatibility complex
ORF:	Open reading frame
ORF(E):	Early open reading frame
ORF(L):	Late open reading frame
Pap test:	Papanicolaou test
PCR:	Polymerase chain reaction
pRB:	Retinoblastoma protein
PV:	Papillomavirus
QT-PCR:	Quantitative PCR
ROPV :	Rabbit oral papillomavirus
RS-PCR:	Restriction site polymerase chain reaction
RSOs:	Restriction site oligonucleotides
SCC:	Squamous cell carcinoma
SIL:	Squamous intraepithelial lesion
S-phase:	Synthesis phase
T-cell:	Thymus cell
URR:	Upstream regulatory region

DEDICATE

This memoir is dedicated to my beloved husband, and to the love of our lives, Nancy, Katie and Kevin, in appreciation for the constant encouragement, help and support provided (without a complaint), during the course of my study.

To my dear parents for having instilled in me the love of learning.

ACKNOWLEDGMENTS

My deepest thanks to my research director, Dr Coutlée, for his availability, his useful comments and expert guidance throughout the program.

Muito Obrigada!

My gratitude goes also to the attending research staff, Helene Voyer, Pierre Forest, and Simon Gagnon for introducing me to the lab, answering my questions, and for the helpful discussions.

LITERATURE REVIEW

Introduction

Over 100 types of human papillomavirus (HPV) have been identified based on DNA sequence analysis [1]. Of these, 40 infect the anogenital region and have been classified according to their oncogenic potential into high-risk and low-risk types. High-risk papillomaviruses have been found to be the single most important risk factor of cervical cancer [2], with HPV 16, the most frequent oncogenic type, accounting for over 50% of cervical cancer[3, 4]. Most women are infected with the HPV virus shortly after sexual debut, with prevalence reaching a maximum around 25 years of age. Prevalence decreases rapidly thereafter, as most HPV infections become latent or are cleared by the host immune system [5-7].

Thus while genital HPV infection is the most common sexually transmitted infection, cancer of the cervix is an uncommon outcome of a high-risk HPV infection. Recent research suggests viral load and viral integration as potential markers for cervical disease progression [8, 9].

Association between increasing viral load of HPV 16 and increasing severity of cervical lesions has been found [10-13]. Conversely, smaller amount of HPV16 DNA in women with HSIL compared to those with LSIL has also been reported [14, 15].

The physical status of high-risk HPV also promises to be a risk marker to evaluate progression of cervical lesions to cancer [16-18]. HPV integration into the cellular genome usually disrupts the E1 and E2 genes, E2 being the preferential site of

integration [19, 20]. The disruption of the E2 regulatory gene, due to integration, results in lack of expression of the E2 protein with subsequential upregulation of the oncogenic E6 and E7 proteins [18, 21]. The continuous overexpression of the E6 and E7 proteins contributes to malignant transformation.

The role that integration plays in malignant transformation is still being questioned. Initial studies on viral integration found viral DNA to be integrated into the host genome in nearly all cases of cervical carcinomas and cervical carcinoma cell lines [22-26], whereas the HPV genome was usually in episomal form in benign and low-grade cervical intraepithelial lesions[16, 25, 27].

This temporal view of integration has, however, been challenged recently as some investigators [15, 28-30] have detected HPV integration even in preneoplastic lesions. In these studies, viral load and viral integration were assessed with qualitative or quantitative real-time PCR targeting the E2 and E6 gene.

The aim of this study was to quantitatively assess, by real time PCR, amplification of E2/E6 sequences in exfoliated cells from normal and LSIL cervical specimens, from which to evaluate viral load and E6/E2 ratio. An E6/E2 ratio of 1.2 or greater was suggestive of integration in nine samples. Restriction site PCR (RS-PCR), a technique that allows retrieval of human–viral junctions, followed by DNA sequencing, however, did not confirm integration site in these potential integrants, although it did identify integration in the positive control of SiHa cultured cells, and some integration artefacts.

1. Basics of Human Papillomavirus Virology

1.1 History

The papillomaviruses are a very diverse family of non-enveloped double-stranded DNA viruses. These small DNA tumour viruses are found in a wide variety of higher vertebrates including mammals, reptiles, and birds [31, 32]. Papillomaviruses infect both mucosal and epithelial cells and induce cellular proliferation giving rise to malignant or benign tumours (warts, papillomas).

The common wart has been described since ancient times, and is characteristic of cutaneous and mucosal epithelial infections. Ciuffo at the beginning of the 20th century demonstrated cell-free filtrates from warty lesions to transmit the disease leading him to conclude that warts are related to an infectious agent [33].

The first papillomavirus was identified in cottontail rabbits in 1933, but progress on the study of human papillomavirus (HPV) infection was delayed for many decades because the virus could not be propagated in cell culture[34].

In 1956, Koss described the morphological aspects of cells from warty lesions of the cervix, coining the term koilocytic atypia [35]. It took another 20 years, however, before researchers were able to demonstrate that this morphological appearance was due to HPV infection [36-38]. The morphological features of koilocytic atypia, which include perinuclear cytoplasmic clearing, peripheral condensation of cytoplasmic filaments, with nuclear enlargement and

hyperchromasia, have since been confirmed to be diagnostic for effect of the HPV virus and as the direct result of the viral genome replication.

The papillomavirus (PV) were studied less intensively in the 1950s and 1960s. Nevertheless, there were two important advances, namely, the physiochemical analysis of the virions and the demonstration that papillomavirus replication was closely associated with the differentiation process of the infected epithelium [39]. Papillomaviruses have indeed proven difficult to propagate in vitro because these viruses replicate in stratified squamous epithelium, which is not mimicked in monolayer cultures. Also, the species specific nature of HPV has thus far also prevented the adaptation of authentic HPV infection to experimental animals, although some useful animal papillomavirus models have since been described.
With the development of molecular cloning technique in the 1970s, however, investigators were able to study the biological and biochemical properties of papillomavirus genomes. Sequencing of the cloned Papillomavirus genomes identified open reading frames and the function of the viral genes was determined by reverse genetics, and this resulted in a revived interest in papillomavirus research [40, 41]. Since, DNA sequence analysis has led researchers to recognize that papillomaviruses are a very diverse group with over 100 human members [31].

During the past decade, it has been determined that a subset of HPV types is closely linked with certain human cancers, most notably, cancer of the cervix. Interest has therefore been focused on this specific subgroup of HPVs which are associated with genital lesions. Of the 40 HPV which infect the anogenital tract, approximately 15 have been found in cervical cancers in a higher percentage than controls, while others are found rarely in cervical cancer, and this has given rise to the distinction between high-risk and low-risk HPV types.

Whilst studies have determined the interaction between HPV and the epithelial host cell, have identified the HPV protein functions and recognized the molecular targets of infected cells, ongoing research seeks to understand the natural history of the infection, to determine the biological properties of the different HPV types, and to identify the role of the nonviral and viral factors in the pathogenesis of cervical disease that may influence the outcome of an HPV infection.

1.2 Taxonomy of papillomaviruses

1.2.1 Family Classification

Early systems of classification lumped papillomavirus, polyomavirus and simian vacuolating virus, into a family grouping collectively known as the papovavirus family. This was based on identification of a common genetic structure: all 3 have a small, circular, double-stranded DNA genome (episome) that replicates in the host cell nucleus, and releases a non-enveloped virion with an icosahedral protein capsid as in Figure 1[31].

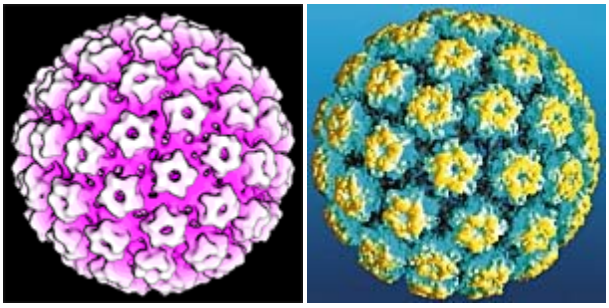


Figure 1: Atomic Model of the Human papillomavirus showing the arrangement of capsid proteins Later it was recognized, however, that the papillomaviruses were distinct from the other 2 members of this group. The papillomavirus genome

ranges from 6900 up to about 8000 base pairs in length, 60% larger than the polyomavirus genome. As such, the capsid is 55 nm in diameter, rather than 40 nm. The genomes are organized differently and except for the helicase motif of the PV E1 protein, do not share any major nucleotide or amino acid sequences [31]. These discoveries led to the reclassification of papillomaviruses as a distinct family by the International Committee on the Taxonomy of Viruses. HPVs are now officially recognized as members of the Papillomaviridae family.

1.2.2 Genotype Classification

The common warts and lesions of Epidermodysplasia verruciformis (EV), which contain large quantities of viral particles, provided enough material to isolate viral DNA genomes. Initially, as more and more types of viruses were identified, researchers in the field agreed on a taxonomic system based on numbering, with each subsequent type receiving the next higher number[31]. For instance, HPV1 is an abbreviation for human papillomavirus type 1. An isolate was accepted as a new type based on liquid hybridisation analysis.

Since the early 1980s, however, when the first full genomes of several papillomaviruses were cloned, nearly all known papillomavirus genomes have been sequenced and compiled into a database such as GenBank and EMBL databases[42] This has allowed for a new classification, at the International Papillomavirus Workshop in 1995, based on nucleotide sequence of the L1 gene.

This new classification, based on DNA sequence, includes to date over 100 genotypes. Genotypes are defined as sharing between 71% and 89% identical nucleotide sequences with other HPV types in the L1 open reading frame, which is the most conserved ORF in the papillomavirus genome [43]. Subtypes have between 90% and 98% sequence identity to a prototype sequence, and variants of a genotype have <2% sequence difference in the coding regions [44].

Further conventional cloning of complete genomes has been difficult due to either limited amounts of sample available or because the viral DNA sequence is toxic to the vector systems used in cloning. This has led to an increased use of PCR amplification of overlapping fragments to obtain viral DNA genomes. These are distinguished by the mention cand, as for example, HPVcand (number.) PCR amplification with degenerate primers mainly of the L1 ORF has identified a few hundred potentially PV novel types [45].

1.2.3 Phylogenetic Classification

Phylogenetic analysis of the L1 sequences has identified clear groupings of the papillomavirus types based on how closely related they are genetically [31], Figure 2 below.

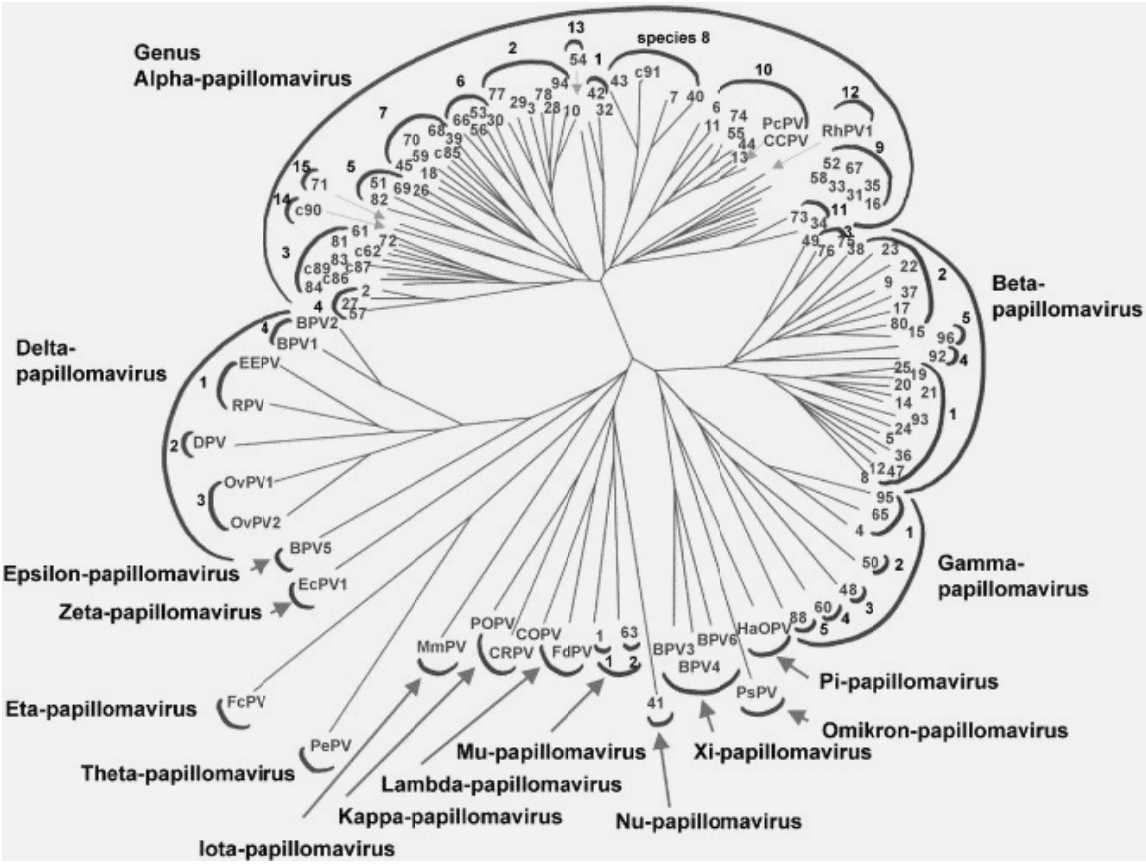


Figure 2: Phylogenetic tree containing the sequences of 118 papillomavirus types. The L1 ORF sequences were used in a modified version of the Phylip version 3.572 and based on a weighted version of the neighbor-joining analysis. The tree was constructed using the Treeview program of the University of Glasgow. The numbers at the ends of each of the branches identify an HPV type; c-numbers refer to candidate HPV types. All other abbreviations refer to animal papillomavirus types. The outermost semicircular symbols identify papillomavirus genera, e.g. the genus alpha-papillomavirus. The number at the inner semicircular symbol refers to papillomavirus species. To give an example taken from the upper part of the figure, the HPV types 7, 40, 43, and cand91 together form the HPV species 8 in the genus alpha-papillomavirus.

Phylogenetic criteria have thus led to the taxonomic levels of family, genus, species, and the previously defined types, subtypes and variants. The sixteen different genera have less than 60% similarity of the L1 gene sequence. Species share between 60% and 70% of L1 ORF identical sequences.

These groupings are relatively consistent with observable papillomavirus phenotypes, including species of origin, tissue tropism, and association with benign versus malignant lesions.

1.2.3.1 Genus or Site of Infection

Based on DNA sequence and protein homologies, the relations between HPV genotypes can be expressed in the form of phylogenetic trees. Tissue tropism of the HPVs is reflected in the grouping of species within a genus. As such, genital mucosal human papillomaviruses are grouped into the genus Alpha-Papillomavirus, although the genus contains a few viruses that are tropic for cutaneous sites and cause common warts. The Alpha-Papillomavirus, however, share certain life cycle features common to this genus that differ from that of other cutaneotropic viruses.

The Beta papillomaviruses are evolutionary distinct from the Alpha genus and seemingly cause asymptomatic infections in the general population. The Beta Papillomavirus have been most frequently isolated from cutaneous epithelium, particularly among patients affected with a rare inherited disorder termed epidermodysplaqsia verruciformis (EV).

While the taxon genus encompasses PV types that have adapted to a particular tissue type and location, this is not absolute. Thus, within the genus Alpha-Papillomavirus, HPV16 is found not only in the genital mucosa, but can also be found in the mucosa of the oropharynx, and in genital cutaneous epithelium.

1.2.3.2 Species: Oncogenic versus non-oncogenic HPV types

The genus Alpha-Papillomavirus which groups genital HPV genotypes is further divided into evolutionarily related subgroups called species [44]. Thus, within a genus, distinct genomic sequences having identical or very similar biological and pathological properties belong to the same species. The sequence-based taxonomy therefore groups the HPV types with known cancer association at the species level.

The HPV types that have most often been associated with cervical cancer and its precursor lesions have been evolutionarily clustered into species 5, 7, and 9. Fifteen HPV types are considered to be carcinogenic or high-risk types: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, and 82. Three are classified as probable high-risk types: 26, 53 and 66. In contrast, HPV types in species 10 have almost no association with invasive cancer. Consequently, these HPV types have been called low-risk types. HPV types 6, 11, 40, 42, 43, 44, 54, 61, 70, 72, 81, and HPV 89 are classified as low-risk types [4].

Further, there is good agreement between epidemiologic classification and the classification based on phylogenetic grouping as seen in Table I below.

Phylogenetic and Epidemiologic Classification of HPV Types.

Phylogenetic Classification	Epidemiologic Classification	
	High risk	Low risk
High risk	16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 82, 26,* 53,* 66*	70
Low risk	73	6, 11, 40, 42, 43, 44, 54, 61, 72, 81, CP6108

TableI: Phylogenetic and Epidemiologic Classification of HPV types. The epidemiologic classification of these types as probable high-risk types is based on zero controls and one to three positive cases.

1.2.4 Clinical Association: Host-Site-Disease

Papillomaviruses have often been classified primarily according to the host species they infect and the sites or diseases with which they are associated. Of the more than 100 human papillomaviruses types that have been identified, they fall into two groups, cutaneous and mucosal HPVs. Mucosal types are associated with oropharyngeal and cervical lesions.

1.3 Genetic Organization

1.3.1-Virion Structure

All papillomaviruses share a number of characteristics. Papillomaviruses are small, nonenveloped, icosahedral DNA viruses that replicate in the nucleus of squamous epithelial cells, mucosal and cutaneous.

The virion particles are 52-55 nanometers in diameter and contain a single molecule of double stranded circular DNA of approximately 8000 base pairs in size bound to cellular histones and contained within a capsid (or spherical protein coat), composed of 72 pentameric capsomers.

A linear representation of the HPV genome is depicted in figure 3 below, organized in open reading frames (ORF'S) and a non-coding region which has been variably referred to as the upstream regulatory region (URR) or the long control region (LCR).

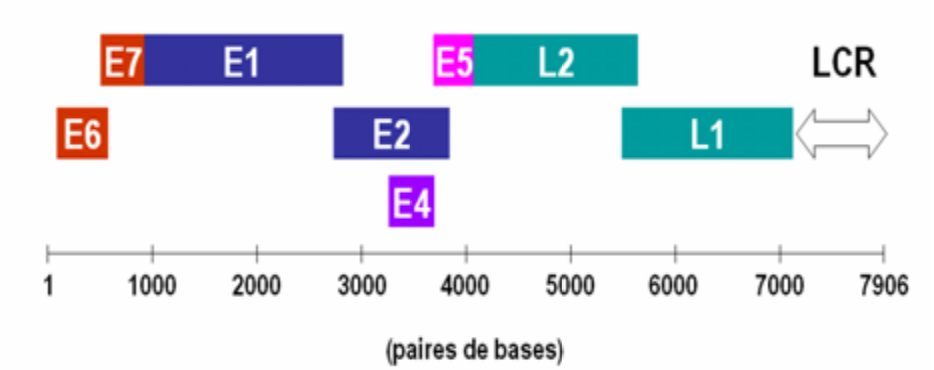


Figure 3: Linear representation of the HPV16 genome; early genes (E1-E6), late genes (L1-L2), long control region (LCR).

The capsid proteins (L1 and L2) are virally encoded by the late open reading frames. The L1 protein is the major structural element, and has a molecular weight of approximately 55 kDa. The L1 protein represents approximately 80% of the total viral protein, whereas L2 is a minor virion protein component, and has a molecular size of approximately 70 kDa. Infectious virions contain 360 copies of the L1 protein organized into 72 capsomeres[46]. A single L2 molecule may be present in the centre of the pentavalent capsomeres at the virion vertices [46, 47]. Both proteins play an important role in mediating efficient virus infectivity.

1.3.2-Genome Structure and Organization

The genomes of the more than 100 human papillomaviruses types have been molecularly cloned and sequenced in their entirety. As the genomic structure of papillomaviruses shares many common features, the genetic map of HPV16 in figure 4 below illustrates the overall genetic organization of HPV genomes.

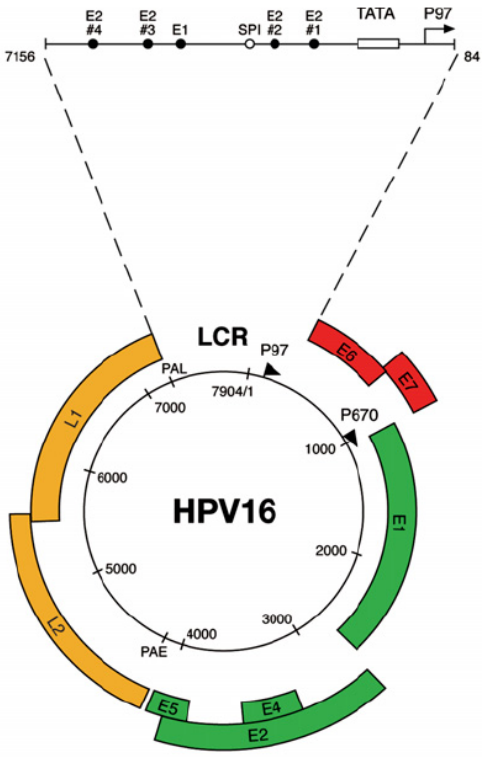


Figure 4: Genome organization of human papillomavirus type 16

The HPV16 genome (7904 bp) is shown as a black circle with the early (p97) and late (p670) promoters marked by arrows. The six early ORFs [E1, E2, E4 and E5 (in green) and E6 and E7 (in red)] are expressed from either p97 or p670 at different stages during epithelial cell differentiation. The late ORFs [L1 and L2 (in yellow)] are also expressed from p670, following a change in splicing patterns, and a shift in polyadenylation site usage [from early polyadenylation site (PAE) to late polyadenylation site (PAL)]. All the viral genes are encoded on one strand of the double-stranded circular DNA genome. The long control region (LCR from 7156–7184) is enlarged to allow visualization of the E2-binding sites and the TATA element of the p97 promoter. The location of the E1- and SP1-binding sites is also shown.

All of the viral open reading frames (ORFs) are transcribed by one strand.The coding strand contains approximately 10 designated translational ORF that are classified as either early (E) or late (L) ORF, based on their location within the

genome[44]. The early region of the viral genome encodes for proteins E6, E7, E1, E2, E4, E5, which are implicated in DNA replication, transcription, and cellular transformation. The late ORF composed of L1 and L2 encode the viral capsid proteins. Downstream of the late L2 capsid gene ORF is the ~850bp LCR that contains no ORF but contains the sequence elements required for regulation of gene expression, replication of the genome and its assembly into virus particles.

The viral E proteins are transcribed from the early promoter whereas the L proteins are transcribed principally from the late promoter. Viral early genes are expressed in undifferentiated and intermediately differentiated keratinocytes, whereas the products of the late genes, the capsid proteins L1 and L2, are expressed only in productively infected differentiated cells [48]. The function of the individual ORF, whose properties have been well characterized, is described in more detail below in section 1.4.4 entitled HPV Protein Functions.

1.4-Normal Infectious Cycle

The papillomaviruses are highly species-specific and also have a specific tropism for squamous epithelial cells. Therefore, all papillomaviruses obligatorily complete their life cycle in the epithelial tissue that they infect.

The human papillomaviruses establish productive infections only within stratified epithelium of the anogenital tract (and of skin and the oral cavity), eventually producing virions from the lysis of dying epithelial surface. As the infected cell moves towards the epithelial surface, the different stages of the virus life cycle are tightly linked to the differentiation program of the epithelial tissue and there is a coordinated timely expression of the different viral gene products.

1.4.1 Attachment, Entry, and Uncoating

The epithelial basal layer of the uninfected epidermis contains cells that are mitotically active. As the surface cells exfoliate, it is the continual division of the basal cells that allows for renewal of the epidermis.

It is believed that papillomavirus infection begins when PV particles gain access to the basal keratinocytes or cervical epithelial cells. This occurs most likely through microwounds or damage of the epithelial sheet [49, 50], although some papillomaviruses are thought to infect sites where access to the basal layer is already naturally facilitated, as at the base of the hair follicle, or sites where the columnar and stratified epithelial cells meet each other (such as the cervical or anal transformation zone).

The receptor by which papillomaviruses bind and enter the cells has not been clearly identified, however, alpha6- integrin has been proposed as a candidate receptor and heparin sulphate may also be involved [51, 52]. Following binding, papillomaviruses are taken into the cell relatively slowly, and for HPV 16, the virus seems to penetrate the cell by clathrin-dependent receptor-mediated endocytosis [53-55].

Inside the cell, there is papillomavirus uncoating and release of the virion occurring most likely by the disruption of intracapsomeric disulfide bonds (in lysosomes). The L2 minor capsid protein facilitates the transfer of the viral DNA to the nucleus where it undergoes transcription and replication.

1.4.2 Virus Replication and Life Cycle

1.4.2.1 Productive Infection

Papillomaviruses establish productive infections only within stratified epithelium and the viral life cycle is closely linked to the differentiation program of the infected epithelial cell, as depicted in figure 5. Productive infection occurs in warts and in CIN 1 lesions of the cervix. The productive infection of cells by the papillomaviruses can be divided into early and late stages.

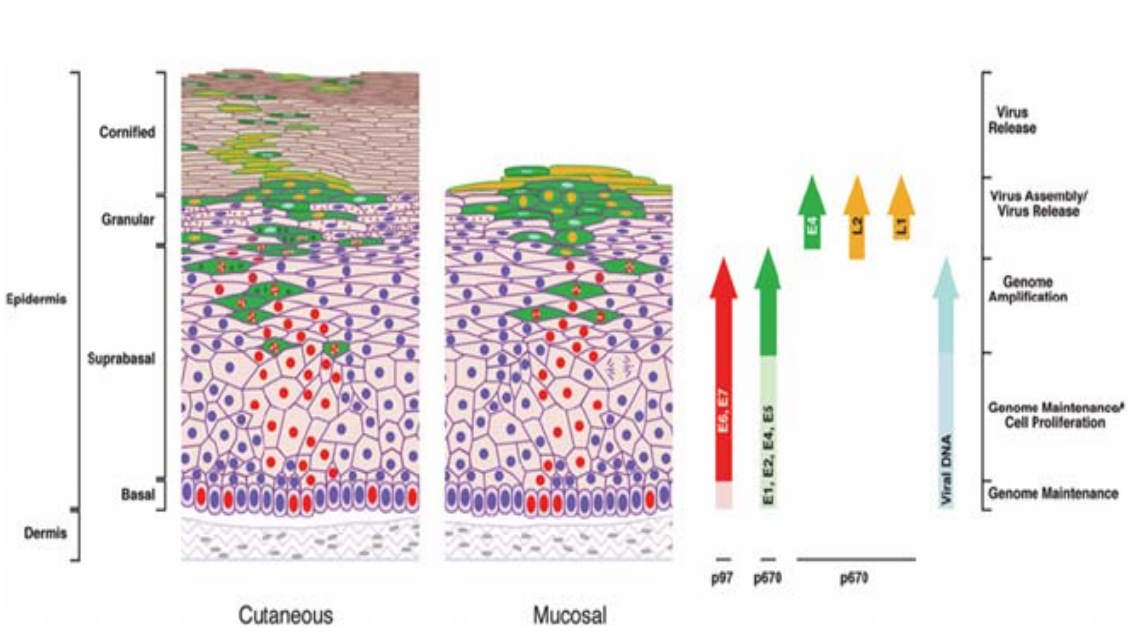


Figure 5: HPV genome and its expression within the epithelium

The key events that occur following infection are shown diagrammatically on the left. The epidermis is shown in colour with the underlying dermis being shown in grey. The different cell layers present in the epithelium are indicated on the left. Cells in the epidermis expressing cell cycle markers are shown with red nuclei. The appearance of such cells above the basal layer is a consequence of virus infection, and in particular, the expression of the viral oncogenes, E6 and E7. The expression of viral proteins necessary for genome replication occurs in cells expressing E6 and E7 following activation of p670 in the upper epithelial layers (cells shown in green with red nuclei). The L1 and L2 genes (yellow) are expressed in a subset of the cells that contain amplified viral DNA in the upper epithelial layers. Cells containing infectious particles are eventually shed from the epithelial surface (cells shown in green with yellow nuclei). In cutaneous tissue, this follows nuclear degeneration and the formation of flattened squames. The timing and extent of expression of the various viral proteins are summarized using arrows at the right of the Figure. The consequence of expressing viral gene products in this ordered way is shown on the far right.

1.4.2.1.1 Early Stage of Productive Infection

Following access to the basal layer (cycling cells), the viral genome will replicate with the cellular DNA during S-phase. The genomes will be divided equally between daughter cells, and each infected basal cell thus accumulates a stable but low copy number of episomes, in the order of 50-100 copies per cell. This type of non-vegetative DNA replication is thought to require the expression of the viral replication proteins, E1 and E2, and possibly E5. Papillomavirus gene expression of the immediate early proteins E6 and E7 is maintained at minimal in the dividing basal cells.

1.4.2.1.2 Late Stage of Productive Infection

In the normal epithelium, suprabasal cells normally complete the cell cycle and begin the process of differentiation in order to produce the protective barrier of the skin or mucosa[56]. However, in HPV-infected epithelium, the cells in the suprabasal layers continue dividing and lose the normal differentiation phase [57]. The expression of E6 and E7 viral proteins is upregulated in the HPV-infected suprabasal cells, in contrast to the basal cells. The high level of expression of the E6 and E7 proteins, both of which exhibit pleiotropic effects (as discussed below), induces the host DNA replication machinery. This allows vegetative DNA replication to occur, followed by expression of the virus capsid proteins (L1 and L2) in the highly differentiated cells, producing genomes to be packaged into

capsids [58]. Also, although nuclei are degraded in normal differentiating epithelia, in HPV- infected epithelium, nuclei are present in all layers.

In this phase, E1 and E2 play a critical role. E2 protein is required for the initiation of viral DNA replication and genome segregation. In addition, E2 can also act as a transcription factor and can regulate the viral early promoter P97 in HPV16 and control expression of the viral oncogenes E6 and E7. The E7 of HPV16 has been shown to be necessary and sufficient to induce suprabasal DNA synthesis. The E5 oncoprotein also contributes quantitatively to this property.

The mechanism(s) which upregulate the switch from plasmid maintenance to vegetative viral DNA replication are not known. The switch may involve the presence or absence of controlling cellular factors in differentiating keratinocytes. In addition, or alternatively, the relative levels of viral factors, such as E1 or E2, or their modification, may change in terminally differentiating keratinocytes. Few studies have examined the mode of vegetative viral DNA replication in differentiated cells.

1.4.3 Regulation of Viral Gene Expression

1.4.3.1 Viral Transcription

Papillomavirus transcription is tightly regulated by the differentiation state of the infected squamous epithelial cell.

Papillomavirus transcription is complex. Whilst multiple promoters generate the various mRNA species, the mRNAs also undergo alternate and multiple splice patterns, resulting in diverse mRNA species in different cells. The major promoter active for HPV 16, in nonterminally differentiated cells, is P97 which directs the expression of E6 and E7 as well as several other early gene products.

1.4.3.2 The Long Control Region (LCR)

Each papillomavirus LCR (also referred to as the URR) contains constitutive enhancer elements that have some tissue or cell type specificity. These constitutive enhancer elements are responsive to cellular factors as well as to virally encoded transcription regulatory factors. Binding of these factors to the URR modulates viral replication and viral gene transcription. Binding sites have been identified for the virally encoded E2 regulatory proteins and the origin of DNA replication that binds the E1 replication factor, as well as for the cellular factors AP1, Oct1, and YY1, among others.

The cis-responsive elements play an important role in initial expression of the viral genes after virus infection and may otherwise be important in the maintenance of viral latency.

1.4.4 HPV Protein Functions

As mentioned previously, the viral genome is divided into early and late open reading frames (Table 2). The early open reading frames encode 6 proteins related to regulation of DNA replication and cell proliferation [59, 60]. The early open reading frames are E1, E2, E3, E4, E5 and E6. The late open reading frames are L1 and L2, and encode proteins related to the viral capsid [60].

The roles of the viral gene products have been most thoroughly worked out for the Alpha HPV types, in particular, the high-risk types associated with cervical cancer. The functions of each of the early and late virally encoded proteins are summarized in Table II and discussed in more detail in the appropriate sections below.

HPV gene	Functional category	Primary role
E1	Replication	Episomal maintenance Up-regulation of genome replication
E2	Replication, transcription	Episomal maintenance Repressor for E6 and E7
E4	Replication, transcription	Production of L2 Disruption of cytokeratin filaments
E5	Replication, transcription, transformation	Stimulation of host growth factor receptors
E6	Replication, transformation	Up-regulates cell cycle Down-regulates p53, Bak, Bax Up-regulates telomerase
E7	Replication, transformation	Up-regulates cell cycle Down-regulates pRb Up-regulates p21 and p27
L1	Viral assembly	Major capsid protein
L2	Viral assembly	Minor capsid protein

Table II: Function of HPV Proteins

1.4.4.1 Regulatory Proteins E1 & E2

The regulatory proteins, E1 and E2, modulate transcription and replication.

1.4.4.1.1 In Viral Replication

Establishment of the viral genome as a stable episome in the proliferating basal cell layer requires the expression of the viral replication proteins E1 and E2. The molecular basis for the role of E1 and E2 in replication is well understood. The E1 gene product is a 73 kDa protein and is expressed at very low levels in the basal cells. The E1 protein binds weakly to the six E1 specific DNA binding sites located within the viral origin of replication. The E2 protein associates with E1 primarily through its N-terminus and also binds to DNA as a dimer through its C-terminus. The complexing of E2 with E1 increases the affinity of the E1 protein to the E1 binding sites in the LCR. The resultant E1-E2 complex induces localized distortion at the viral origin. As additional E1 molecules are recruited at the viral origin, the E2 protein is eventually displaced. This gives rise to a hexameric complex with helicase activity. Subsequently, the DNA unwinds providing the template for DNA synthesis.

The replicating proteins, E1 and E2, are also necessary for the replication of the viral episomes above the basal cell layer. As the infected cell migrates to the epithelial surface, activation of the late promoter (P670 in HPV 16), dependent on cellular differentiation, results in increased levels of E1 and E2. As the levels of E1 and E2 proteins increase, viral genome amplification occurs in the suprabasal cells, producing genomes to be packaged into infectious virions.

1.4.4.1.2 In Genome Segregation

The E2 proteins are well conserved among the papillomaviruses. The E2 protein consists of a transactivating domain at the N-terminal and of a sequence specific DNA binding and dimerization domain located in the carboxy terminal region of the protein. These two domains are separated by an internal hinge region.

The DNA binding domain of E2 recognizes a palindromic motif in the long control region (LCR) of the viral genome. In the case of HPV 16, there are four such E2 specific binding sites in the non coding region of the viral genome.

In addition to the full length E2's critical role in viral DNA replication, the product of the E2 ORF is also important in genome segregation. As the basal cells of the epithelium undergo mitosis, it is thought that the viral genome replicates in synchrony with the cellular DNA during S-phase. It has been reported that E2 plays an important role in anchoring the viral episomes to mitotic chromosomes or to the mitotic spindle (for the high risk genotypes) thereby ensuring correct division of the episomes between the daughter cells [61, 62]. E2's crucial role in segregation thus allows episomes to be maintained long term within replicating cells at a constant level.

1.4.4.1.3 In Viral Transcription

E2 transcriptional regulation has been well studied for HPV infecting the genital tract. E2 acts as a transcriptional factor, activating or repressing the viral early promoter (P97 in HPV16), thus controlling expression of the viral oncogenes E6 and E7. At low levels, E2 acts as a transcriptional activator. At E2 high levels, E2 represses oncogene expression by displacing SP1 transcriptional activator from a site adjacent to the early promoter.

The ability of E2 to either repress or activate early viral gene expression according to its abundance is thought to result from differences in the affinity of E2 for its various binding sites [63]. High levels of E2 acts to downregulate the expression of E6 and E7 genes in experimental systems. In HPV16 it is thought that binding site 4 is the primary site that is occupied when E2 is present at low levels and that binding to this site and to binding site 3 leads to promoter activation [64]. As E2 increases in abundance, occupancy of the remaining sites leads to the displacement of basal transcription factors, such as Sp1 and TBP (TATA-box-binding protein), that are necessary for promoter activation[65]. It appears that the increase in E2 expression that is important in stimulating viral genome amplification will lead eventually to the down regulation of the E6/E7 expression and to the eventual loss of the replicative environment necessary for viral DNA synthesis.

In addition to binding at its cognate sites, the E2 transcriptional activation function is required for E2 mediated promoter repression. Specific conservative point mutations within the E2 transactivation domain that eliminates E2 mediated transcriptional activation, also eliminates E6/E7 promoter repression [66, 67]. The

specific cellular transcription or chromatin remodelling that may mediate the repression has not yet been identified.

1.4.4.2 Proliferatory Role of E6 and E7 in HPV Productive Life cycle

1.4.4.2.1 Basics of E6 & E7

The first open reading frames in the HPV early region, E6 and E7, comprise the two main oncogenes of HPV.

The E6 proteins, from both the low and the high risk types, are approximately 150 amino acids in size and contain two zinc fingers with the characteristic motif Cys-X-X-Cys. Following HPV infection of the epithelial basal cell, the high risk E6 protein is one of the first early viral genes to be expressed, and can be found both in the nucleus and in the the cytoplasm.

The E7 protein of the high risk HPV is a small nuclear protein of 100 amino acids which has been shown to bind zinc through its single binding motif and is phosphorylated by casein kinase II.

1.4.4.2.2 HPV-Infected Epithelial Differentiation

The basal cells of the normal epithelium are mitotically active cells. As the basal or first parabasal cell divides, one cell maintains the basal population, while the other migrates upward to become the superbasal cell layer. The suprabasal cells exit the cell cycle and begin the process of differentiation to become the protective barrier that is normally provided by the skin or mucosa [56].

A number of model systems have been used to examine the papillomavirus productive life cycle during in vivo infection. Following experimental inoculation of mucosal epithelial tissue by ROPV (rabbit oral papillomavirus) or COPV (canine oral papillomavirus), there is an increase in cell proliferation in the basal and as well in the suprabasal cells [68, 69], leading to mature wart formation within 4 weeks post infection. In HPV-infected keratinocytes, there is stimulation of cell cycle progression, and as a result, expected normal terminal differentiation of the epithelium does not occur [57]. Following natural HPV infection, there is minimal activity of the E6 and E7 genes in the basal cell layer. The low activity of the E6 and E7 viral proteins drives the infected basal cell to divide, producing a small number of infected basal cells. The increase in proliferation of infected basal cells and the viral stimulation of suprabasal cells to re-enter the cell cycle, subsequently increases the number of virus producing cells.

The basic mechanism by which papillomaviruses stimulate cell cycle progression is well known. Basically, the E6 and E7 gene products target an abundance of cellular functions, with the most important interactions being what may be termed the E6-p53 and E7-pRb model.

1.4.4.2.3 Role of E6 & E7 in Cell Cycle Progression

Vegetative papillomavirus replication occurs in the more differentiated cells of the epithelium. These cells, however, are no longer dividing. Although the E1 and E2 proteins necessary for viral replication are coded by the virus, the virus is dependent on cell for all other enzymes necessary for its replication. These proteins are normally only expressed in S-phase during cellular DNA replication. Papillomaviruses have thus evolved, through E6 and E7 oncoproteins, a mechanism that activates the cellular genes necessary for their replication. E7 inactivates retinoblastoma tumor suppressor and related pocket-proteins which results in increased levels of p53, followed by G1 cell cycle arrest or apoptosis. E6 by promoting p53 degradation counters the acitivity of E7 and allows for activation of the cell DNA machinery necessary for viral replication.

1.4.4.2.3.1 Role of E7

1.4.4.2.3.1.1 E7–pRb Model

The main cellular target of E7 is the tumour suppressor protein pRb. Normally, the hypophosphorylated form of pRb binds to and inactivates the transcriptional regulator E2F. As a transcriptional regulator, the E2F molecule is important in the activation of genes necessary to enter S-phase. In normal cells, complex formation between pRb and E2F thus prevents the cell from entering the S-phase.

As a result of papillomavirus infection, however, the HPV E7 protein binds to the protein pRb [70], resulting in dissociation of the pRb protein from the E2F

transcriptional factors [71]. The released E2F transcriptional factors stimulate cells to pass from the G1 phase of the cell cycle to the stage of DNA replication. Thus E7 binding to pRb results in loss of pRb function which leads to E2F release, and subsequently basal and parabasal cell proliferation in the absence of external growth factors.

Apart from the dissociation of the pRb/E2F complexes, the binding of E7 to the protein pRb also causes a sharp decrease in the stability of the pRb protein and its rapid proteosomal degradation [72].

As a result of E7-pRb interaction, cell cycle progresses, and the tumour suppressor protein p53 also increases. The p53 tumour suppressor protein has numerous functions. Its principal role, however, is that of a transcriptional regulator required for the expression of a number of genes involved in cell cycle regulation and apoptosis.

1.4.4.2.3.1.2 E7 associates with other cellular proliferation proteins

In addition to pRb, E7 complexes with the pRb related pocket proteins, p107 and p130 [73], thereby exerting its transforming activities.

E7 also associates with other proteins involved in cellular proliferation, such as histone diacetylases [74], components of the AP1 transcription complex [75] and the cyclin-dependent kinase inhibitors p21 and p27 [76].

Although the property of the E7 viral oncoprotein to complex pRb would appear to account, at least in part, for induction of DNA synthesis and cellular proliferation, genetic studies indicate, however, that complex formation between E7 and the pocket proteins, including pRb, is not sufficien for its immortalization and transforming functions, suggesting the existence of additional E7 cellular targets relevant to cell transformation.

Figure 6 below provides a list of additional targets to which E7 has been shown to bind, although the relevance of such interactions is not yet clear.

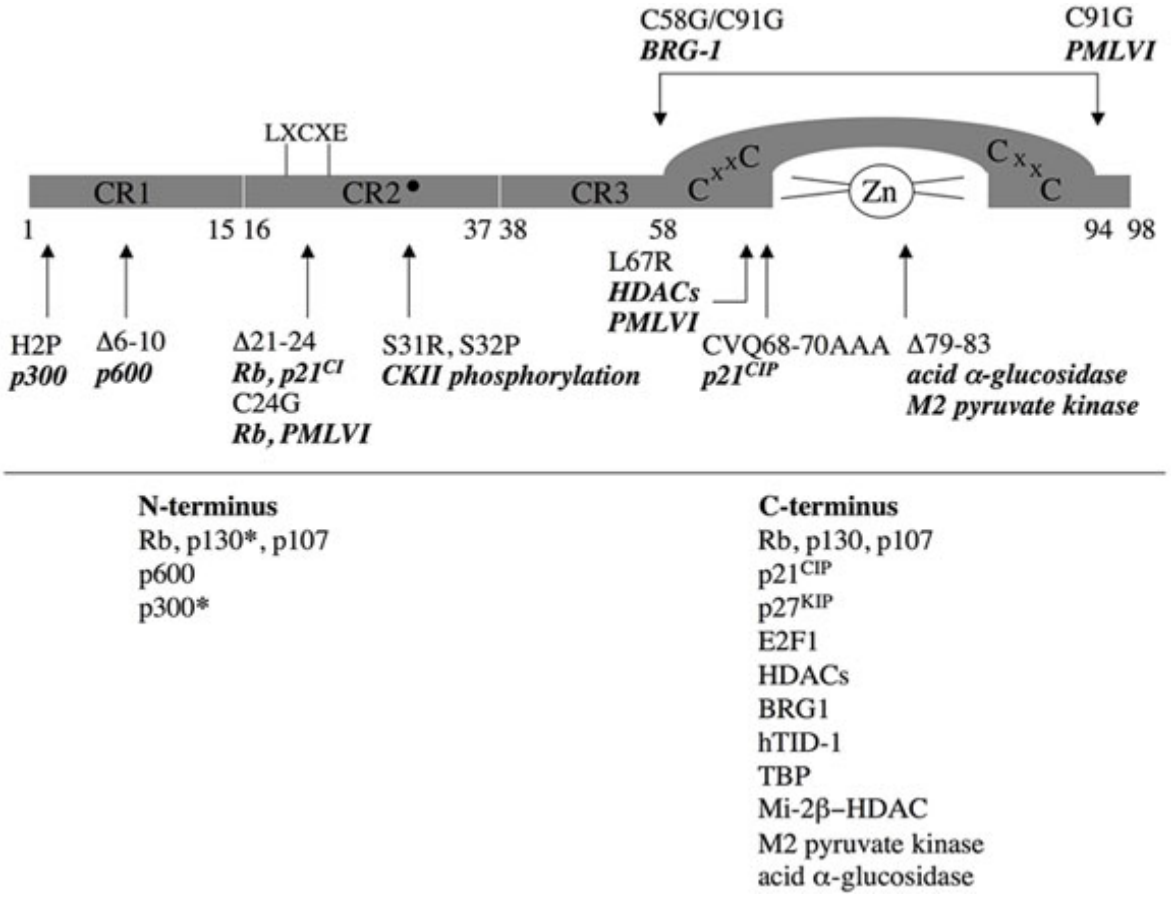


Figure 6. Schematic representation of the HPV 16 E7 protein and interaction of E7 with cellular proteins. Conserved regions (CR) 1-3 are indicated and exhibit homologies with Adenovirus E1A and SV40 large T antigen. A consensus casein kinase phosphorylation site within CR2 is denoted by a black dot. Regions that harbor binding domains for cellular proteins are indicated. These include within the N-terminus a strong interaction domain for the retinoblastoma protein family, as well as domains for the binding of p300 and p600 . The C-terminus contains a weak interaction domain for the retinoblastoma protein family and an E2F binding domain, as well as domains for the binding of the $p21^{CIP1}$ and $p27^{KIP1}$ cyclin/cdk inhibitors, hTID-1 (168), BRG1 (169), TATA binding protein (170), Mi-2beta (46) , M2 pyruvate kinase and acid alpha-glucosidase. Asterisks indicate cellular proteins that interact with both high and low risk HPV E7.

1.4.4.2.3.2 Role of E6

1.4.4.2.3.2.1 Normal DNA Damage Response

Cells normally respond to DNA damage or to genotoxic agents by increasing the level of p53 protein within the cell. The higher level of p53 within the cell will signal growth arrest in the G1 phase of the cell cycle, or even apoptosis. Therefore, intracellular level of p53 is part of a cell defense mechanism which allows for either the DNA damage to be repaired before initiation of a new round of DNA replication or allows the removal of the cell by apoptosis [77].

1.4.4.2.3.2.2 E6-p53 Model

A primary role of the E6 protein is its association with the cellular tumour suppressor p53. In the case of high risk types, the E6 oncoprotein binds to p53 and stimulates its degradation by forming a complex with an ubiquitin ligase, the human protein E6AP [78]. The degradation of p53 is thought to prevent growth arrest or apoptosis in response to E7 mediated cell cycle entry in the upper epithelial layers.

1.4.4.2.3.2.3 E6 associates with Bak and Bax

The role of E6 protein in proliferation is further emphasized by the finding that it also associates with the proapoptotic proteins Bak [79] and Bax [80]. As an anti-apoptotic protein, E6 allows cellular progression and prevents death of the infected replicating cells.

1.4.4.2.3.2.4 Other E6 cellular targets

A variety of other E6 cellular targets have been identified, however physiological relevance to transformation or immortalization has not yet been clarified. It is possible that the binding of E6 to some of these targets might contribute to the virus host cell functions unrelated to cellular transformation. The E6 has been found to bind to over 12 different cellular proteins, as depicted in figure 7 below. Relevant function of some of these proteins is discussed under the heading viral factors, subheading genotype.

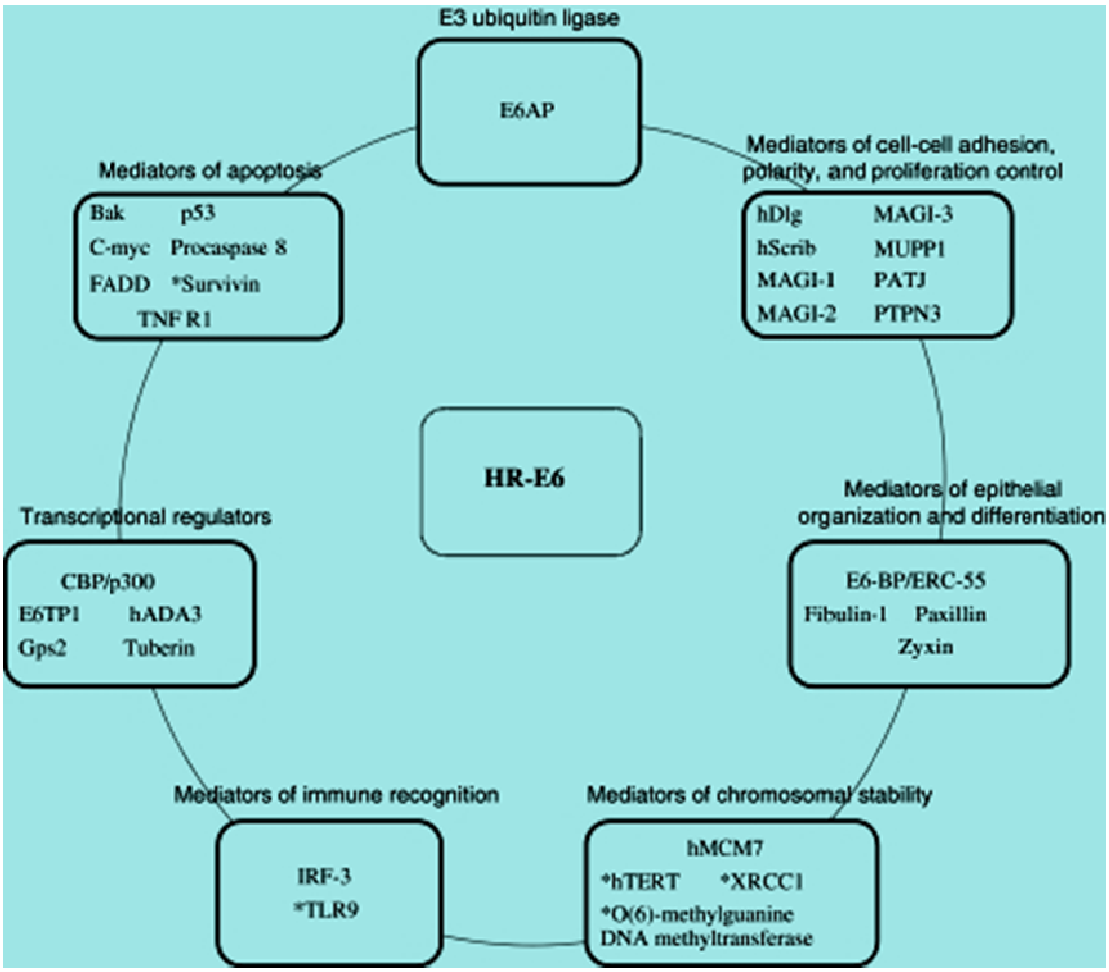


Figure 7: Representation of cellular proteins affected by HR-E6. ‘‘Asterisk’’ designates proteins where the influence of HR-E6 is at the level of transcription; proteins without an asterisk are those to which HR-E6 has been shown to bind.

1.4.4.3 The E4 and E5 Proteins

1.4.4.3.1 The E5 Protein

The HPV E5 proteins are required for optimal growth. In tissue culture, various HPV E5 genes have been shown to have some modest transforming activities. In transgenic mice, HPV16 E5 expressed in basal keratinocytes can alter the growth and differentiation of stratified epithelia and induce epithelial tumors at high frequency.

Although the biochemichal mechanism by which the E5 gene of HPV exerts its growth stimulatory effects have not yet been fully elaborated, it may involve interactions with the EGF receptor or the 16 kd subunit of the vacuolar ATPase, each of which has been shown to bind HPV E5 proteins. Interaction of HPV E5 protein with the 16kd subunit of the vacuolar ATPase can inhibit acidification of endosomes.

The E5 protein also binds to platelet-derived growth factor β receptor and colony stimulating factor 1 receptor [81], and is believed to be necessary for amplification of the viral genome [82] possibly related to the expression of polyadenylation sequences that regulate viral gene expression for all early ORFs [83].

There is also some evidence that E5 helps prevent cell apoptosis after DNA damage [84]. E5, however, is not expressed in most HPV-positive cancers, suggesting that if the E5 gene does stimulate cell proliferation in vivo, it probably functions in benign papillomas and not in cancer.

E5 protein might also participate in the initiation of the carcinogenic process or in some other aspects of the viral-host interaction relevant to the pathogenesis of the HPV infection. Indeed, some data implicate E5 in the downregulation of major histocompatibility complex (MHC) class II antigen expression which may aid the infected cells to evade the host immune system [85].

1.4.4.3.2 The E4 Protein

Although E4 is located in the early region of the viral genome, it is nevertheless a protein that exerts its action in the viral replication cycle. The expression of E4 is necessary for the production of the L2 protein, one of the 2 capsid structural proteins. The E4 protein is the most abundant protein in benign warts, and is expressed at relatively high levels in differentiated squamous cells [83] where viral packaging and assembly occur. In cultured epithelial cells, the E4 proteins are associated with the keratin cytoskeleton. The HPV16 E4 protein induce collapse of the cytokeratin network causing condensation of tonofilaments at the cell periphery and perinuclear cytoplasmic clearing which results in the morphological appearance of a koilocyte [83]. It is possible that this disruption facilitates the release of viral particles from superficial squamous epithelial cells [86, 87].

In addition, E4 may have a role in supporting amplification of the viral genome [88]. Reiterating, the available data thus point to the possibility that E4 may contribute to vegetative DNA replication or to altering the cellular environment in a manner that may favour virus synthesis or perhaps virus release.

1.4.4.4 Structural Proteins L1-L2

The structural proteins L1 and L2 compose the viral capsid [89].

1.5 Virus Assembly and Release

Little is yet known of papillomavirus assembly and release. The virus is not cytolytic. Virus particles are only observed in the granular layer of the epithelium and not at lower levels. Release of the virion particles occurs in the granular layers of the mucosal epithelium or the cornified layers of the keratinized epithelium. Viral release probably follows cell death thereby increasing the invisibility of HPVs to the immune system.

1.6 Abnormal Proliferative Infection

Ocasionally, the tight regulation between viral gene expression and epithelium differentiation is lost. In contrast to a differentiated and virally productive phenotype as that which occurs in warts and low grade lesions, in a proliferative infection there is apparent morphological evidence of increased abnormal proliferation of the basal cells. E6 and E7 are overexpressed in proliferating basaloid cells that overtake the epithelium and produce lesions. Ongoing research seeks to identify among viral, host, and environmental factors the mechanism that mediates loss of E2 control of E6/E7 expression.

2-Natural History and Epidemiology of Cervical HPV Infection

2.1 HPV acquisition and transmission

HPV is acquired by sexual transmission and this has been strongly confirmed by studies involving initially virginal women [90]. HPVs in the anogenital tract are transmitted mainly by skin-to-skin or mucosa-to-mucosa contact with infected epithelium of cervical, vaginal, vulvar, penile or anal origin. It is presumed that HPV infections are easily transmitted through microscopic lesions in the skin or the mucosa.

Some studies, however, report that on occasion, HPVs are transmitted through a non-sexual mode of transmission, namely, through vertical transmission from parent to unborn child, by fomites and by skin contact [91].

The probability of infection per sexual act is not known. However, a recent study on the McGill Concordia Cohort of young female students, the same cohort being studied for the current report on integration, has estimated the probability of HPV transmission per coital act among newly forming couples by using stochastic computer simulation. The HPV transmission probability per act was found to range anywhere from 5-100%, leading the authors to conclude HPV to be more transmissible than either HIV or herpes virus [92]. There is also evidence to suggest that the amplitude of sexual transmissibility possibly varies among HPV types and also among populations [93].Due to their common transmission avenue, several HPV types can be transmitted from the same partner. This results in a high proportion of simultaneous infections with several different HPV types when individuals of either sex are sampled in the general population.

Epidemiological studies suggest, that in addition to the sexual behaviour of both men and women, genetic and environmental susceptibility factors such as age, use of barrier contraceptives, co-infections, and male circumcision are related to the acquisition and transmission of HPV[94].

2.2 Prevalence of HPV Infection

2.2.1 Definition

HPV prevalence can be defined as the percentage of individuals with detectable infection at a given point or period in time. Because the infection due to HPV is subclinical, prevalence estimates will vary based on the method of detection (cytology, colposcopy, biopsy, or HPV DNA detection). PCR-based methods yield the highest prevalence estimates of HPV DNA in the genital tract, and identify between 1.5%-44.3% of genital HPV infections in otherwise normal Pap smears [95, 96]

However when cervical specimens are taken from these women during follow-up surveys, the majority of infections is found to be transient. Thus total exposure to HPV infection is difficult to measure not only due to detection method but also because HPV DNA detection is usually transient. Serologic assays used to detect serum antibodies to certain viral proteins have also been insensitive. Moreover, titers of antibodies induced by natural infection are quite low. Thus, the true extent of HPV infection is thought to be underestimated.

The prevalence of HPV infection also varies between countries. In the United States, the annual incidence of HPV infection has been reported to approach 6.2

million per year, and has an estimated prevalence of 20 million [97], with genital HPV infection considered to be the most common sexually transmitted viral infection [94, 98].

2.2.2 Age-specific prevalence of HPV DNA

HPV incidence peaks soon after women initiate sexual activity, figure 8 below. Prevalence of infection ranges from approximately 25% to 40% [99, 100] in women 15 to 25 years of age. Subsequently, there is a lower incidence of HPV infection with age perhaps due to immune response, or otherwise due, to decreased HPV exposure and/or developing resistance to infection.

In some populations there is an increase in detection of HPV DNA in women over 60 years of age [101]. It is believed this peak of HPV prevalence around the age of menopause could perhaps represent persistent infections acquired at a younger age, could result from reexposure or otherwise be a cohort effect.

2.2.3 Type-specific prevalence of HPV DNA

In general, high risk types tend to be detected more frequently than low-risk types, and infection with one or more of the more than 40 genital HPV genotypes is a common occurrence among sexually active women [6]. HPV16 is the most common type detected among cytologically normal women [102].

More specifically, for the population under study, the McGill Concordia Cohort, 621 female university students were followed for 24 months at 6-month intervals, and prevalence/ incidence rates were determined. Results for the 10 HPV genotypes with the highest incidence rates in this cohort are in table III below [103].

Table 1 Prevalence and incidence[a] of infection with the most frequently detected HPV types and for groups according to oncogenicity

HPV type	Baseline prevalence (%)	No. of incident cases	Women-months of follow-up	Incidence rate (per 1000 woman-months) (95% CI)
HPV-6	2.7	29	12709	2.3 (1.5–3.3)
HPV-16	7.0	62	11928	5.2 (4.0–6.7)
HPV-18	3.1	24	12735	1.9 (1.2–2.8)
HPV-31	2.6	21	12854	1.6 (1.0–2.5)
HPV-39	1.0	247	13476	1.8 (1.1–2.5)
HPV-51	2.9	43	12588	3.4 (2.5–4.6)
HPV-53	4.3	31	12468	2.5 (1.7–3.5)
HPV-54	2.7	32	12783	2.5 (1.7–3.5)
HPV-56	2.6	19	12842	1.5 (0.9–2.3)
HPV-84	3.8	46	12475	3.7 (2.7–4.9)
Any HPV	29.0	155	8151	19.0 (16.1–22.3)
HR HPV	21.8	131	9344	14.0 (11.4–16.3)
LR HPV	14.8	128	10299	12.4 (10.4–14.8)

[a] The 10 types with the highest incidence rates are shown. Not shown are HPV-11, -26, -33, -35, -40, -42, -45, -52, -55, -57, -58, -59, -66, -68, -82, -83, and MM9.

Table III: Prevalence and incidence rates of infection for the most frequently detected HPV genotypes in the McGill Concordia Cohort under Study

In agreement with results in the general population, HPV-16 is the most prevalent type (7%) at enrollment and also presents the highest incident rate (5.2) in this cohort of young female university students.

2.3 HPV clearance versus persistence

2.3.1 HPV Clearance

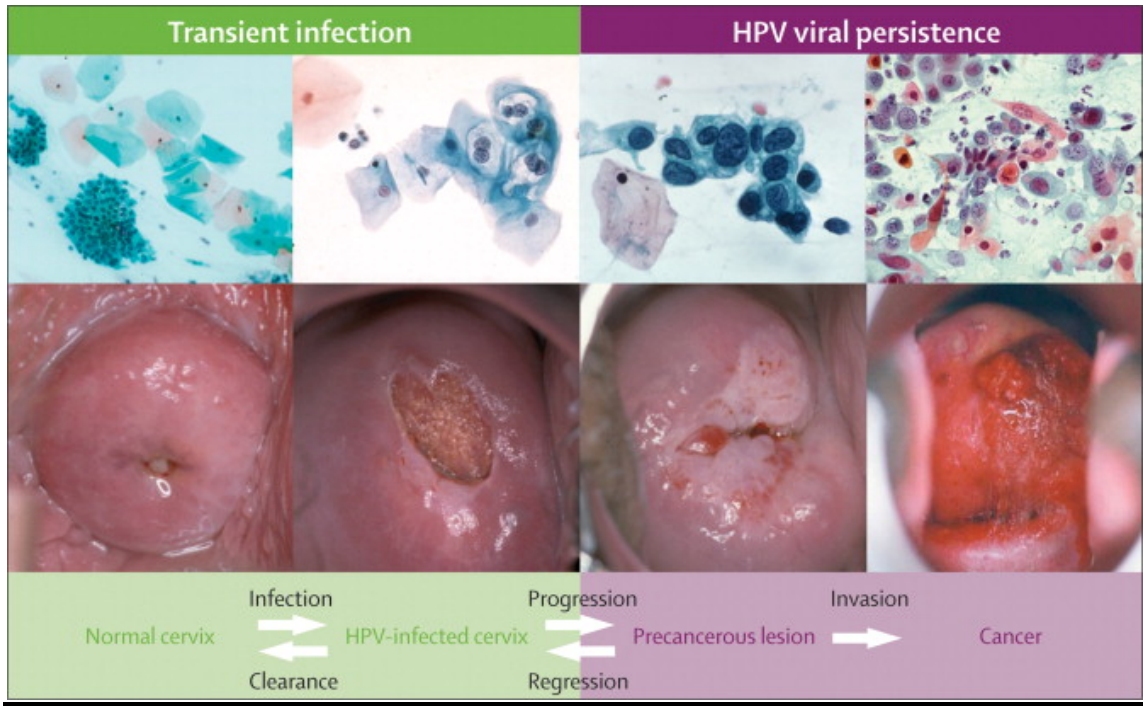


Figure 8: Major steps in the development of cervical cancer

Top row shows cytology, bottom row colposcopy

Cervical HPV infection follows one of two patterns: HPV infection is frequently self limiting with no clinical manifestations, or conversely, 30% of HPV-infected women will develop detectable CIN lesions (infection being accompanied by cytological changes, such as koilocytes) usually occurring within 3-6 months, that may progress to cervical cancer, as in figure 8. Most cervical HPV infections, with cytological abnormality or not, however, are transient. Dependent on their stage and host immune system, 50% of low-grade lesions regress within 1 year, while only a smaller proportion of high grade lesions regress [104].

Overall, as detected by sensitive PCR assays, 85% of the virus and of the cytologic abnormalities are cleared or suppressed by the host immune response within 1-2 years of exposure [105-107].

It is believed that low-grade lesions generally give rise to higher grade lesions after several years [108, 109], or high-grade lesions can also develop rapidly from a high risk HPV infection without any signs of CIN1 or CIN2 stages [110].

Although less frequent, the process is also reversible by clearance of the HPV infection and regression of the precancerous lesions to normality given that invasion has not yet occurred.

Ongoing cohort studies, with up to 10 years of follow up data, have shown, however, that after clearance, the same HPV type can re-appear [111]. It remains unclear as to whether infections resolve by complete viral clearance or whether the virus remains as a persistent although silent infection. The virus may persist as a latent infection by various viral strategies: the virus may be maintained in a latent state in the basal cells without viral DNA replication, the virus productive life cycle is absent but is reactivated periodically similar to that of herpes simplex virus, or the virus is chronically maintained at low replication.

There is indirect evidence for such a persistent silent state (latent or chronic infection) from immunosuppresed patients (renal transplant and HIV infected) who

are at a higher risk for CIN and cancer, and from pregnant women who develop HPV-related lesions during pregnancy (an immunocompromised state) which

regress postpartum. In addition it has been reported that a high risk HPV infection, but otherwise normal smear, can develop into CIN3 within 4 years [106].

2.3.2 HPV Persistence

Compared to clearance, persistence is uncommon. Persistence can be defined as infection with the same HPV type detected two or more times over a certain period, with the interval usually being 6-12 months.

A small proportion of infected women do have persistent infection, with infection by high risk types being of a longer duration as compared to low risk types. HPV16 persists longer than any other HPV type [112], and it is generally accepted that persistent infection with high risk HPV confers a higher risk for developing high-grade neoplasia and cancer[113].

The molecular virology underlying HPV persistence, progression, and invasion is not yet well understood, but this causal model is supported by epidemiological and laboratory data.

3. Development of cervical cancer

3.1 Pap Test

Current cervical cancer screening in industrialized countries is based on Papanicolaou staining (Pap test) of cervical swab or cytobrush specimens containing exfoliated cervical cells. This cytological staining process enables microscopic detection of cellular changes characteristic of HPV infection

(koilocytosis, dyskariosis) and associated with various stages of the development of ICC.

Women with precancerous or cancerous lesions identified through Pap screening are referred for repeat Pap screening, colposcopy, biopsy and treatment if need be.

3.2 Pre-Cancerous Abnormalities

The microscopic abnormalities result from HPV induced cellular proliferation, and have been classified by pathologists as cervical intraepithelial neoplasia (CIN) of varying grade.

3.2.1 Classification Systems

Many classification systems for cervical cytology have been proposed over the years in different health systems, including the Papanicolaou terminology, the Munich classifications, and the histology oriented WHO classification frequently used in the UK. The most widely used system, however, is the two-tiered Bethesda classification in that abnormal cells are classified as low grade or high grade squamous intraepithelial lesions, that is, LSIL, HSIL respectively [114]. A substantial number of atypical specimens not attributed to either one of these are referred to as atypical squamous cells of undetermined significance (ASCUS). Lesions are graded according to a multistage cytopathology model (see figure 9 below for comparative terminology), and assessment may be visual (colposcopy), microscopic (via cytology and histology), or by molecular detection methods.

3.2.2 Microscopic Lesion Progression

Precancerous lesions undergo a defined set of dysplastic alterations, over the time course of several years, before progressing to malignant cell carcinomas[115].

The severity of the lesion (microscopic abnormality) is determined by the degree to which the squamous epithelium is replaced by basaloid cells, with the entire thickness being replaced in the most severe dysplasias (Figure 9).

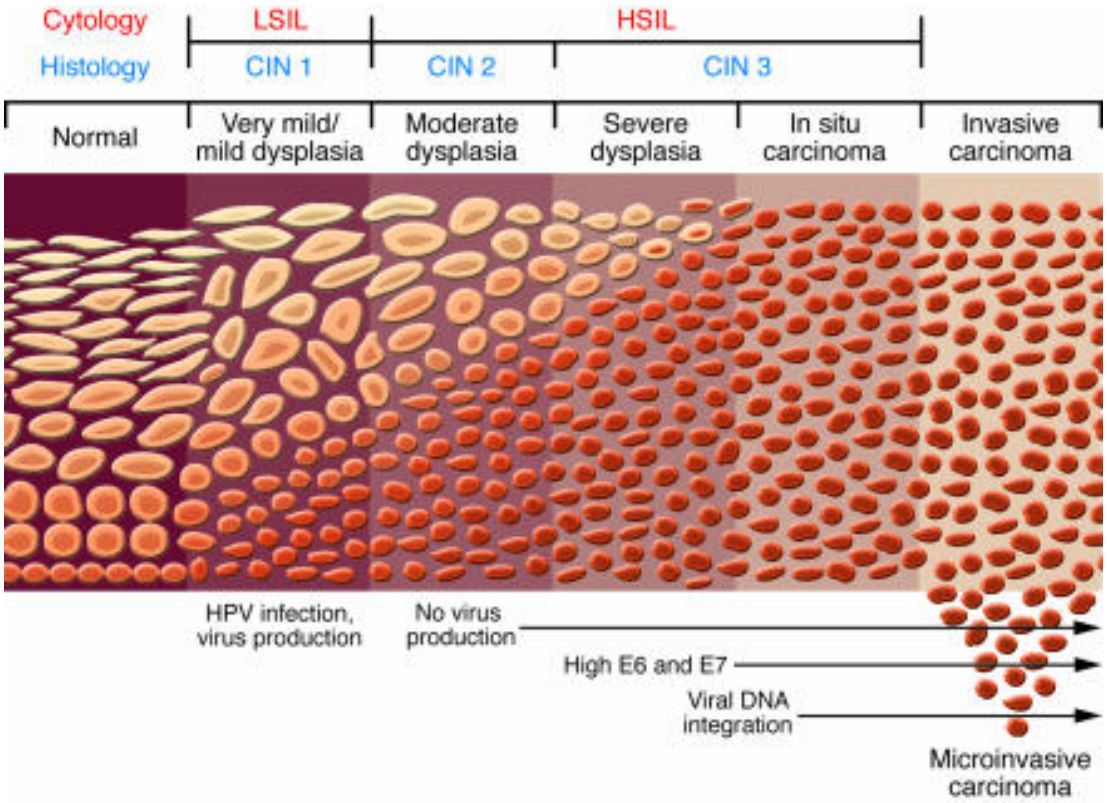


Figure 9:Progression from a benign cervical lesion to invasive cervical cancer. Infection by oncogenic HPV types, especially HPV16, may directly cause a benign condylomatous lesion, low-grade dysplasia, or sometimes even an early high-grade lesion. Carcinoma in situ rarely occurs until several years after infection. It results from the combined effects of HPV genes, particularly those encoding E6 and E7, which are the 2 viral oncoproteins that are preferentially retained and expressed in cervical cancers; integration of the viral DNA into the host DNA; and a series of genetic and epigenetic changes in cellular genes. HSIL, high-grade squamous intraepithelial lesion; LSIL, low-grade squamous intraepithelial lesion.

3.2.2.1 Low-Grade CIN

The most characteristic histological feature of anogenital HPV infection is nuclear atypia. HPV-related nuclear atypia is due to heteroploidy, which results from mitotic spindle abnormalities, leading to replication without cytokinesis. This results in bi and multinucleated cells and enlarged atypical nuclei. In low-grade lesions the nuclei are mainly diploid and polyploid. Normal mitotic figures are generally increased in low-grade lesions, but these remain restricted to the lower third of the epithelium, as are undifferentiated basal type cells.

3.2.2.2 High-Grade CIN

High-grade lesions (CIN 2 and CIN3) are substantially more atypical than low-grade CIN, have a higher degree of disorganization and have undifferentiated cells that extend beyond the lower third of the epithelium. In high-grade lesions, the characteristic koilocyte of low- grade lesions is absent or considerably decreased, and the high-grade lesions have abnormal mitotic figures which clearly distinguish them from the low-grade lesions.

Accurate identification of lesion grade has prognostic significance. About 20% of CIN 1 will progress to CIN2 and around 30% of CIN2 will progress to CIN3 if left untreated. Approximately 40% of untreated CIN3 lesions will progress to cervical cancer [116].

It is important to note here, however, that microscopic diagnoses are susceptible to subjectivity and lack interobserver reproducibility.

3.2.3 Molecular Progression of Lesions

After a normal HPV infectious cycle, the E6 and E7 proteins are expressed at low levels. However, during cancer progression, and at an as of yet undefined point, normal regulation of the papillomavirus life cycle is lost. This results in changes in the viral gene expression pattern, with an increase in levels of E6 and E7 transcripts in cervical cancer cells [117] with E6/E7 expressed in the full thickness of the epithelial lesion [118].

CIN1 or (LSIL) lesions generally resemble productive lesions as previously described under the heading normal infectious cycle. The order of events is generally similar to that seen in productive lesions with virus coat proteins expressed at the epithelial surface.

In CIN2 and CIN3, the order of events remains the same, however, E7 expression is increased [119-121] and the onset of late events is retarded. Production of infectious virions becomes restricted to smaller and smaller areas close to the epithelial surface.

Integration of HPV sequences into the host cell genome can accompany these changes and can lead to further deregulation in the expression of E7 (and the loss of the E1 and E2 replication proteins). In cervical cancer the productive stages of the virus life cycle are no longer supported and viral episomes are usually lost.

4. Cancer of the Cervix

4.1 Cause of Cervical Cancer

In the mid 1970's, Meisels and Fortin recognized that human papillomavirus (HPV) infection of the cervix occurred frequently and that, histologically, the infection often displayed characteristics of mild cervical intraepithelial neoplasia (CIN) [122].

Meanwhile, ongoing clinical and epidemiological research pointed to cervical cancer as a sexually transmitted disease. This led zur Hausen in 1976 to propose HPV as the sexually transmitted agent responsible for cervical cancer [36]. In the early 1980s, zur Hausen then went on to identify HPV16 and HPV18 [22, 123] and these two types together have been found to account for about 70% of cervical cancer.

Since then, epidemiological studies have permitted researchers to identify the tight relationship between the virus and the disease. Under optimal testing conditions HPV DNA can be identified in 99.7% of all invasive cervical cancer specimens, in at least 70% of CIN1, 80% of CIN2, and 96% of CIN3 precursor lesions.

Indeed, it is now well established that the vast majority of cervical carcinomas and its precursor lesions are caused by persistent infections with certain types of human papillomaviruses which have been designated as high risk types [3, 124, 125].

4.2 Burden of Cervical Cancer to Humanity

Despite being a theoretically preventable disease, cervical cancer remains the second most common cancer in women worldwide, after breast cancer [126]. Of approximately 2 274 000 women affected in 2002, there were 500 000 newly diagnosed cases of cervical cancer, and of these, about 275 000 deaths, being equivalent to approximately one tenth of all cancers in women [127]. Despite its worldwide distribution, the frequency of cervical cancer varies considerably, being about ten times more common in some countries than in others (Figure 10 below).

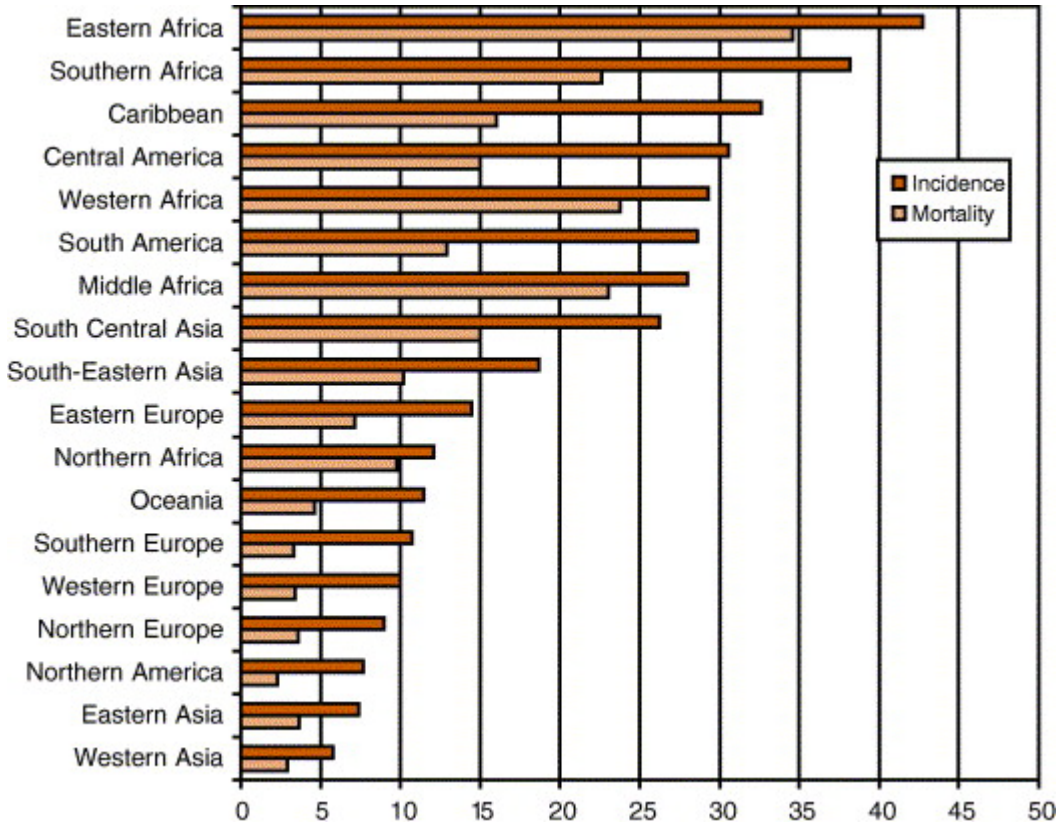


Figure 10: Cancer of the uterine cervix: age-standardised (world) incidence and mortality rates per 100 000 (all ages) in 18 world regions

Approximately 80% of the burden of cervical cancer occurs in developing countries of Africa, South and Central America and in the Caribbean, where it is the major cause of cancer-related death among women. In developing countries, cervical cancer is often detected at later stages due to non-existent or inadequate screening, and also because the standard treatment options are often absent or unaffordable, thus making it the most prevalent and important cancer in women.

By contrast, in industrialised nations, the widespread use of the Papanicolau (Pap) screening test has reduced cervical cancer rates by nearly 80% in the last 50 years [127]. Nevertheless, in Western Europe, approximately 33 500 new cases of cervical cancer are diagnosed each year, and of these, 15 000 women die from the disease. In the United States, an estimated 13 000 new cases of cervical cancer, and 4 000 deaths occurred in 2003 [128]. Therefore, despite recent advances in treatment, cervical cancer remains nonetheless an important cancer in women from a public health perspective.

4.3 The Cervical Transformation Zone

The transformation zone of the cervix is defined as the area where the columnar cells of the endocervix form a junction with the stratified squamous epithelium of the ectocervix. Microscopically, the ectocervix can be divided into basal, intermediate, and superficial zones. Cell proliferation is normally limited to the basal zone. As the basal or first parabasal cell divides, one cell maintains the basal cell population while the other migrates upward and differentiates in a highly controlled manner. Most cervical cancers and its precursors originate from the cervical transformation zone. The reason for this is not well defined. However, as

the basal cells are both exposed directly and undergoing rapid turnover at the endo-ecto cervical junction, they are particularly vulnerable to HPV infection, without the need for microtrauma. Similar sensitivity to HPV infection has been demonstrated in squamocolumnar junctions of the rectum, larynx, and nasal cavity. Infection of the squamous epithelia in the basal layer gives rise to benign or malignant lesions[129].

4.4 Histopathology

Most studies have focused on the more common squamous cell cancers, however, most adenocarcinomas, adenosquamous carcinomas, and carcinomas with neuroendocrine differentiation, although rare, also contain HPV DNA [130, 131]. Most cervical carcinomas are of squamous origin, with HPV-16 being the most common viral type. HPV infection is also ascociated with cutaneous, anorectal and oropharyngeal cancers, as shown in Figure 11 below. And similar to the cervix, the high-risk types are those most often associated with their transformation.

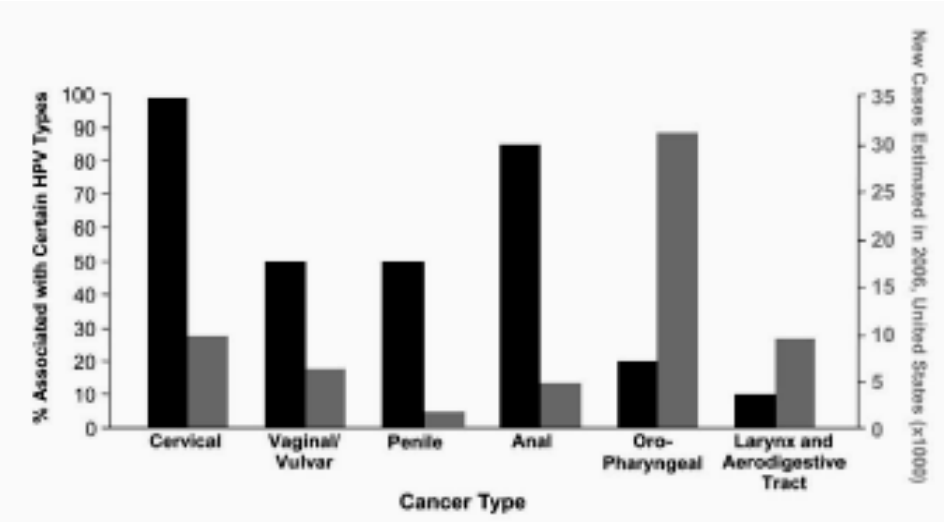


Figure 11: High-risk types of HPV have been identified in a wide range of malignancies.

4.5 Cancer of the Cervix & HPV Types

Certain types of human papilloma viruses, the so-called high risk types are the aetiological agents of cervical cancer [2, 123, 132]. HPV- negative squamous cell cervical carcinomas are extremely rare [133].

4.5.1 Definition of high risk-low risk HPV

There are over 100 types of HPV defined on the basis of DNA homology. Of these, approximately 40 strains are termed mucosotropic viruses that infect the anogenital and upper digestive tracts. Only some of these mucosotropic types, however, are regularly found in cervical cancers, and in a higher proportion than controls, such that these types have subsequently been termed high risk types. On the other hand, those HPV types less frequently found or never found in tumours as compared to controls, have been termed low risk viruses [31].

4.5.2 Low risk HPV

The low risk HPV types are HPV 6, 11, 34, 42, 44, 53 and 54. Infection with HPV types classified as low or no oncogenic risk, predominantly, HPV 6 and 11, may cause subclinical infection and benign genital lesions, including low-grade CINs and ano-genital warts. Genital warts are small bumps localized in the genital areas of men and women including the vagina, cervix, vulva, penis and rectum.

HPV 6 and HPV 11 are responsible for at least 90% of genital warts. Although low risk types such as HPV11 are rarely associated with cervical cancer, they are still medically important because they cause genital warts, a major sexually transmitted disease in many countries affecting 1-2% of young adults.

4.5.3 High risk HPV

Approximately 15 of the mucosotropic HPV types are considered to be high risk types [112]. HPVs 16, 18, 31, 33, 35, 39, 45, 50, 51, 53, 55, 56, 58, 59, 64 and 68 are considered high risk types and are mainly related to high grade cervical lesions and squamous cervical cancer.

Table IV below summarizes the proportion of cervical cancers caused by the carcinogenic HPV types. The most common HPV types in women with cervical cancer in descending order of frequency are types 16,18,33,45,31,58,52, and 35 [134]. Other types identified in patients with cervical cancer but with a much lower frequency are types 39, 51, 56, 59, 68, 73 and 82. However, the relative frequency of each type will vary from one country to another.

	Proportion of cervical cancers caused	Cumulative total
HPV16	54·6%	54·6%
HPV18	15·8%	70·4%
HPV33	4·4%	74·8%
HPV45	3·7%	78·5%
HPV31	3·5%	82·0%
HPV58	3·4%	85·4%
HPV52	2·5%	87·9%
HPV35	1·8%	89·7%
HPV59	1·1%	90·8%
HPV56	0·8%	92·2%
HPV51	0·7%	92·9%
HPV39	0·7%	93·6%
HPV73	0·5%	94·1%
HPV68	0·5%	94·6%
HPV82	0·2%	94·8%
No type identified	5·2%	100%

Table IV: Proportion of cervical cancer caused by the carcinogenic HPV types [135].

4.5.3.1 Human Papillomavirus Type 16

In almost all areas of the world, HPV16 is the most prevalent HR-HPV in the general population. It accounts for about 20% of infections among cytologically normal women, 20% among women with equivocal lesions, and 26% among those with mild abnormalities. Prospectively, HPV16 persists longer on average than any other type. Persistence is highly associated with precancer, approximately 45% of women with persistent HPV16 being diagnosed within 5 years with precancer. Even though HPV16 is more likely to cause cytological abnormalities than any other carcinogenic type, it also disproportionately causes changes suggesting precancer.

HPV 16 is the major carcinogenic type in almost every country surveyed. HPV16 is associated with 54% of SCCs of the cervix and 41% of adenocarcinomas worldwide, and is the main type that causes other anogenital and oropharyngeal cancers.

Given the cause of cervical cancer is persistent infection with one of the approximately 15 carcinogenic human papillomavirus (HPV) types, if we could eliminate HPV16 infection or reliably identify and destroy all its cytopathological and colposcopic manifestations, we could at the very least prevent half of cancer cases in the world. As HPV16 is the most clinically significant, it has also become one of the most studied, thereby serving as the prototype for understanding HPVs, and in particular, for detecting HPV persistence.

4.6 Invasive cervical cancer

4.6.1 Persistence of HPV Infection

Most studies confirm the association between persistent HPV infection and the increased risk of progression to HSIL and cancer [136]. As persistent infection with a high risk HPV is considered the main risk factor for progression to CIN or invasive cancer and since high risk HPV DNA is found in virtually all cases of cervical cancer [124], it has been concluded that HPV infection is therefore a necessary cause of cervical cancer.

Persistent HPV infection alone, however, is not sufficient to cause cancer because some individuals who have persistent infection with a high risk HPV do not develop serious lesions, and because low-grade HPV types that persist are much less likely to progress than high risk types[112].

Therefore, other factors influence persistence, and these are thought to be viral, host or environmental.

4.6.2 Relationship among incidences of cervical HPV infection, precancer, and cancer

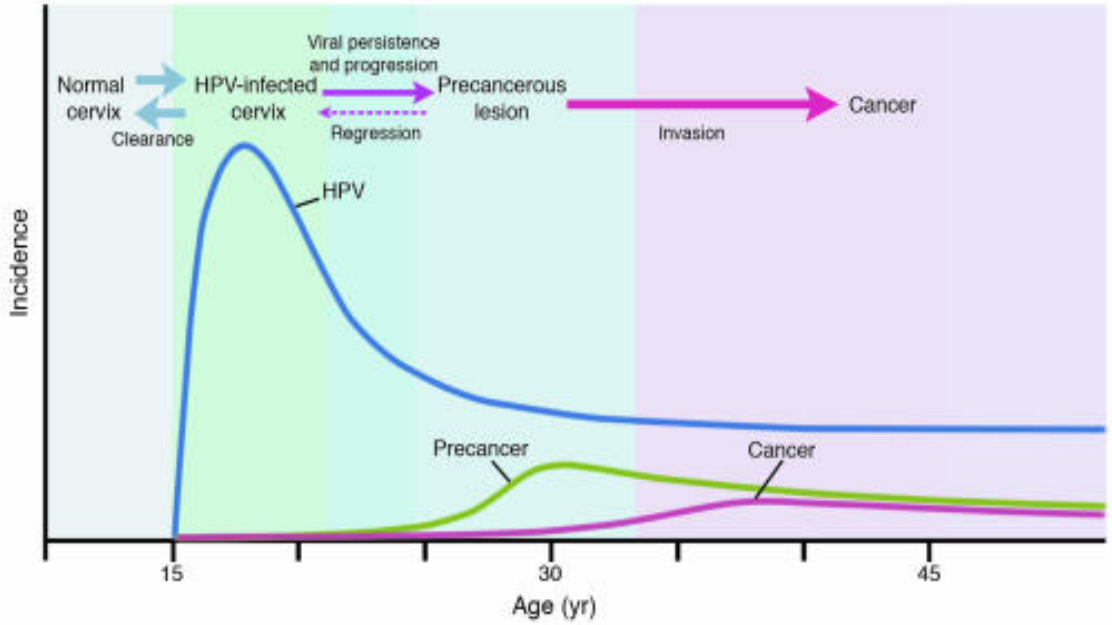


Figure 12: Relationship among incidences of cervical HPV infection, precancer, and cancer. The HPV curve emphasizes the high incidence of infection that develops soon after women initiate sexual activity and subsequent lower incidence because a high proportion of infections are self-limited. The precancer incidence curve follows several years behind the HPV incidence curve and is substantially lower than that of HPV incidence, as there is generally a delay between the acquisition of HPV infection and precancer development, and only a subset of infected women develop precancers. The cancer incidence curve follows several years behind the precancer curve, reflecting the relatively long interval between precancer and progression to invasive cancer. As women approach 40 years of age, the incidence of cancer begins to approach the incidence of precancer. Figure modified with permission from the New England Journal of Medicine53). Copyright © 2009, The American Society for Clinical Investigation.

The precancer incidence curve is considerably lower than the HPV incidence curve, revealing that only a subset of infected women will develop precancers. There are also many more precancers than cancers, suggesting that only a minority invade. Given the high prevalence of genital HPV types in the general population, the incidence of cervical cancers is very low with most infections being either successfully resolved or controlled by the host immune system [115].

There is a lag time of about 7-10 years between the occurrence of HPV infection in the late teens or early twenties and the peak of precancer around 25–30 years of age. The cancer incidence curve follows several years behind the precancer curve, as invasive cancer generally occurs in women 10 years older than women with CIN3. The long interval between initial infection and development of cervical cancer implies that in addition to persistent infection by an appropriate HPV type, additional environmental or host factors may contribute to malignant progression.

5. Determinants of clinical progression of HPV infection

5.1 Environmental Factors

Many of the potential cofactors for cervical cancer and the precursor lesions have not been rigorously evaluated in epidemiological studies, that is, did not restrict their analyses to HPV positive women or these had small sample sizes resulting in conflicting results. Therefore, an incomplete summary follows with focus on infectious agents, multiparity/early parity, tobacco smoking, and hormonal contraceptive use.

5.1.1 Infectious Agents

5.1.1.1 Herpes

It has been suspected since the early 90's that HPV coinfection with certain microorganisms could have an influencing effect on the development of cervical

cancer. Herpes simplex virus 2 is one of the infectious agents that have been most frequently studied as a possible cofactor for cervical cancer. However, despite a positive correlation with herpes simplex 2 infection [137], there has not been consistent epidemiological evidence in support of a herpes role in cervical cancer. [138-140]

5.1.1.2 Chlamydia trachomatis

Another infectious agent, Chlamydia trachomatis, has been associated with clinical cervical hypertrophy [141] and with the induction of squamous metaplasia [142], and additional reports provide evidence in support of C.trachomatis possible role in the etiology of cervical disease [143, 144].

In epidemiological studies that controlled statistically for HPV infection [144], an epidemiological association of Chlamydia trachomatis co-infection with cervical neoplasia and invasive cervical cancer has been found.

Recent studies suggest that C. trachomatis might act indirectly in cervical carcinogenesis by varying the expression of tumour suppressor genes in epithelial cells .

5.1.1.3 HIV/AIDS

The incidence of cervical cancer is significantly increased in women who have human immunodeficiency virus/acquired immune deficiency syndrome

(HIV/AIDS) [145], but it remains uncertain, however, whether there might also be a direct biological interaction between HIV and HPV [146, 147].

5.1.1.4 Cervical inflammation

During C.trachomatis infection, many cytokines are released and these possibly cause tissue damage by inducing apoptosis of uninfected cells [148]. Infiltrating macrophages could cause further tissue damage through release of reactive oxygen species [149]. These effects, together, probably result in partial disruption of the tissue barrier and exposure of the basal cells to HPV infection. Therefore, it is possible that the immune response to pathogen infection plays a role in HPV infection and cervical cancer, thus offering an explanation for the conflicting results obtained in epidemiological studies with a range of microbial agents including herpes viruses, C.trachomatis, and others. Any sexually transmitted infection could possibly act as a co-factor in cervical carcinogenesis by either simply bringing a permissive environment for HPV infection and viral persistence or by promoting genomic changes.

5.1.2 Hormonal Contraceptive Use

Epidemiological data suggest an association between hormonal status that is, age at first intercourse and early parity/multiparity (three or more children), the long term use of oral contraceptives, and the risk for preneoplastic lesions of the cervix and cervical cancer. Progesterone is the major ingredient of oral contraceptives and injectable hormonal contraceptives. The level of progesterone has been shown to increase during pregnancy. In vitro experimental studies indicate that upon

exogenous hormonal stimulation, hormonal recognition elements in the LCR of high risk mucosal HPV will increase the production of the E6 protein [150-152].

The female sex hormone, estrogen, is also considered another possible cofactor to the oncogenic effect of HPVs. During puberty, the high level of circulating estrogen is considered a major influence in the metaplastic changes in the cervical transformation zone.

Additional evidence of a co-carcinogenic role for hormones in cervical cancer results from studies on HPV infected transgenic mice. Steroid hormones were found to increase the transcription of oncoproteins E6 and E7, contributing not only to the formation but also to the maintenance and malignant progression of cervical cancers in HPV16 transgenic mice [153, 154].

Furthermore, the degree of proliferation of SiHa cancer cells was directly proportional to the duration of estrogen therapy [155], and the levels of estrogen were found to be significantly higher in a dose dependent fashion in cervical cancer patients as compared to controls [156] .

There is also experimental evidence that hormones may mediate changes in the immune status of the cervical mucosa [157].

These results suggest that the hormonal environment of the cervical mucosa of an HPV infected woman could contribute to cervical malignancy.

5.1.3 Tobacco Smoking

Cigarette smoking, in the presence of HPV infection, is also being considered a cofactor in the development of preneoplastic lesions of the cervix and invasive cervical cancer. Although secondary genetic changes may occur randomly, the presence of tobacco metabolites in cervical secretions is considered a risk factor in the development of cervical cancer [158], and certain smoking carcinogens are found at significantly higher levels in the cervical secretions of cigarette smoking women [159].

Cigarette smoke contains mutagens, carcinogens and other components that may act as initiators and /or promoters of uterine cervix carcinogenesis. Currently, there appear to be two mechanisms by which tobacco smoking can increase the risk for cervical disease. First, cigarette components can affect the immune function and allow HPV infection to persist and progress: tobacco smoking has been associated with a generalized suppression of the immune system, including a significant decrease in NK cells and NK cell activity, in level of immunoglobin (Ig)G and (Ig)A, and in Langerhans dendritic antigen presenting cells [160, 161]. Second, components of tobacco smoke such as nicotine or cotinine or their metabolites can also act directly as co-carcinogens in cervical tissue [162, 163], or in combination with the immune system.

5.2 Genetic or Host Factors

Host genetic differences that influence the host response to viral infection are also being investigated. These include the immune response and genetic

polymorphisms present in genes related to viral infection (HLA, DNA repair systems, and tumour suppressor genes) [164].

5.2.1 Evidence of Immune Response

In most women, the immune response is able to successfully control the HPV infection [165]. The humoral and cellular components of the immune system have been shown to contribute to the host response in transient infections.

The contribution of the humoral component is expressed via the presence of HPV antibodies. Antibodies against the HPV capsid proteins have been found in the cervical secretion of HPV induced lesions. Antibodies induced by HPV L1 vaccine are nearly 100% protective against incident CIN2, 3.

Evidence that the cellular component of the immune response also plays a role in the control of HPV infections is provided by the observation that most HPV infected tissues show an inflammatory response at the time of regression. Immunomodulation studies involving the use of IFN and imiquimod also highlight the importance of cell-mediated immunity in promoting the regression of HPV lesions.

Additional evidence for the role of the immune system derives from individuals who have genetic or aquired immune deficiency. Immunosuppressed patients are at an increased risk for HPV infections and associated neoplasia. Renal transplant patients, for example, suffering from cell mediated immune suppression are at an increased risk for cutaneous and genital HPV lesions [166, 167]. T-cell immunosuppression has been shown to have a profound effect on the risk for HPV infection and persistence [168], highlighting the significance of the T-cell

response. HIV-positive women have at least double the HPV prevalence of HIV-negative women who report similar risk profiles and the risk for persistent HPV is strongly associated with degree of immunosuppression. These data highlight not only the importance of host responses for resolution of natural infection, but also define the HIV-positive population as high risk group for HPV-associated malignancies.

Therefore, HPV resolution requires an effective host immune response. Most HPV associated lesions regress spontaneously, however, on occasion the lesions progress.

5.2.2 Genetic susceptibility factors

Based on observations that natural daughters and sisters of patients with cervical cancer have a higher risk for developing cancer than adopted daughters or sisters, with half sisters showing approximately half of the risk compared to full sisters [164, 169] genetic predisposition may exist to developing cervical cancer.

Therefore research has focused to identify the host factors that increase one's susceptibility to development of invasive cervical cancer following HPV infection including, among others, the immune major histocompatibility complex (HLA types) and p53 polymorphisms.

5.2.2.1 MHC Complex

Human Leukocyte Antigen (HLA) molecules are responsible for the presentation of foreign antigens to the immune system. The HLA molecules consist of a family of genes within the major histocompatibility complex (MHC), composed of class I and class II genes. The evidence in the literature suggests that members of class II HLA genes are likely to protect against the development of cancer whilst on the other hand no alleles have consistently been associated with an increase risk of disease [113].

5.2.2.2 p53 Polymorphism

Polymorphism at specific loci of several genes is also being investigated, with that of p53 gene being the most extensively studied. The frequent mutation of p53 in human cancers suggests the critical participation of this gene in the carcinogenesis process [170]. Essentially, p53 exists in two main polymorphic forms at codon 72, arginine(72R) or proline(72P) [171], and this polymorphism is balanced [172].

In cervical cancer, HPV E6 oncoprotein is known to contribute to neoplastic progression by inhibiting the p53 pathway. Since it was found that p53Arg is more prone to degradation by HPV16 and HPV18 E6 protein, it was proposed that the polymorphism at codon 72 in the p53 gene could be associated with an increased risk for cervical cancer [173]. Case control studies have compared the frequency of the p53 allele in controls to cervical carcinoma, however results vary by country and HPV genotype [174].

5.3 Viral Factors in Progression to Malignancy

It is well established that the vast majority of cervical carcinomas and its precursors are caused by persistent infections with certain high risk types of human papillomaviruses, the high risk types.

However, within the high risk type group, there are some differences in the frequency of cancer association that are not fully understood [175]. Although HPV16 and HPV31 are closely related at the evolutionary level, HPV 16 is associated with approximately 54% of cervical cancer whereas HPV 31 is associated with only about 3% of cases [135, 175, 176]. Research has therefore focused on viral factors (genotype differences, viral polymorphism, viral load, and HPV integration status) as to elucidate the role of these in cancer progression.

5.3.1 Genotype- High risk-low risk differences

There are more than 40 different HPV genotypes that can infect the cervix. Approximately only 15 of these are associated with cervical cancer [4] and have consequently been called high risk types. Therefore, this has led scientists to believe that the molecular differences between the E6 and E7 proteins of the low and high risk types are important factors that could partially explain the differences in the likelihood of cancer progression between low and high risk viral types.

5.3.1.1 In Vitro Properties of E6 and E7 Proteins

E6 and E7 encode for proteins that allow for immortalization and transformation of the cell that hosts the HPV DNA.

Indeed, a number of assays have been used to evaluate the ability of the E6 and E7 proteins to transform cells in vitro. Using primary rodent cells, primary human fibroblast and keratinocyte cultures, the high risk HPVs have been found to induce transformation [177-183] whereas the low risk viruses do not [182, 184]. These assays have permitted the mapping of the viral genes directly involved in cellular transformation to the E6 and E7 ORFs.

Although the HPV16 E7 is the major transforming gene in established NIH3T3 rodent cells and E7 induces DNA synthesis in quiescent cells and cooperates with an activated ras oncogene to transform primary rodent cells, E6 together with E7 are required for efficient immortalization of primary human fibroblasts or keratinocytes, leading to unterminal differentiation [179, 180, 182].

5.3.1.2 E7 Oncoprotein

As previously stated, the E6 and E7 of both the low risk and high risk HPVs bind to the tumour suppressor proteins p53 and Rb respectively. However, the low risk types bind with a lower affinity than the high risk HPVs.

The high risk E7 protein binds to pRb with a tenfold higher affinity than the low risk E7 proteins [185, 186]. The resultant E7-pRb binding inactivates the pRb tumour suppressor, thereby disrupting a G1 cell cycle control mechanism.
The high risk E7 proteins are also capable of mediating Rb degradation through a proteosome-dependent mechanism [72, 187], which is important for E7-mediated cell transformation.

Unlike the low risk, the high risk E7 proteins are also able to induce centrosome-related mitotic disturbances, as has been detected in cell culture and in transgenic animals. A loss of spindle integrity during cell division increases the risk for chromosome missegregation and aneuploidy. This leads to an increase in genome instability not as yet quite understood [188-190].

5.3.1.3 E6 Oncoprotein

As with E7, the E6 protein also differs in its function between the high and low risk HPV types.

5.3.1.3.1 E6 complexes E6AP and p53

The high risk E6 proteins bind p53 with higher affinity. In addition, the high risk E6 proteins form a complex with both p53 and cellular ubiquitin ligase E6AP (E6-associated protein), which leads to degradation of p53 mediated by proteosomes. Conversely, the E6 of the low risk types don't have either significant ability to bind E6AP nor to stimulate p53 degradation [173, 191].

5.3.1.3.2 p53 dependent repression of transcription

High risk E6 proteins associated with cervical cancer also have the capacity to repress p-53 dependent transcription by binding the transcriptional coactivator p300/CREB-binding protein, providing a second mechanism by which to inhibit p53 expression [192, 193].

5.3.1.3.3 E6 PDZ binding

A number of additional cellular targets have been identified for the high risk E6 proteins. The high risk E6 oncoproteins contain a motif at the extreme C-terminus that can mediate the binding of cellular PDZ domain containing proteins. This motif is unique in the high risk HPV E6 proteins and is not present in the E6 proteins of the low risk HPV types. The E6 protein serves as a molecular bridge between these PDZ domain proteins and E6-AP facilitates their ubiquitylation and mediates their proteolysis. Among the PDZ domain proteins implicated as E6 targets are hDlg, hScrib, both tumour suppressors. Several of the PDZ-containing proteins have been shown to be involved in negatively regulating cellular proliferation. Therefore, through its C-terminal PDZ ligand domain, the E6 protein of the high risk HPV types mediates cell proliferation independently of E7. E6 PDZ binding is reported to mediate suprabasal cell proliferation which may contribute to the development of metastatic tumours by disrupting normal cell adhesion [194-196].

5.3.1.3.4 E6 Activates Telomerase

A p53 independent activity of the E6 of the high risk HPV types is its ability to transcriptionally upregulate telomerase through expression of the catalytic subunit

hTERT. The mechanism by which E6 activates the hTERT promoter is still unclear, although two E6-related mechanisms involving the interaction of E6AP have been proposed.

As normal cells lack telomerase, there is a gradual loss of telomere length upon successive cell divisions in vitro. Once telomere length reaches a critical size, cellular senescence is induced. In HPV infected cells, however, E6-induced telomerase activity synthesizes hexamer repeats at the ends of chromosomes and thus allows for maintenance of telomere length [197, 198], extending the lifespan of cells infected with HPV, and in this way perhaps predisposing to persistent infection.

5.3.2 HPV Viral Polymorphism

There is sufficient nucleic acid heterogeneity within all HPV types to form phylogenetically distinct subgroups (variants) within the same viral type. Sequencing of one or more ORFs

(E6, E7, and L1), as well as the LCR has identified based on DNA homology multiple subtypes and variants for HPV 18 and in particular for HPV16. The HPV16 variants have been broadly categorized into European (E), and non-European variants (NE), according to DNA homology and the region of the world where they were originally isolated. Since the prototype variant of HPV16 was first detected in a cervical cancer specimen from a woman in Europe, this variant established the European lineage. Nonprototype HPV 16 variants are classified as Asian, Asian-American, African-1, or African-2 lineages. The nonprototype HPV 16 variants generally contain multiple nucleotide variations [199].

Multiple studies have demonstrated an increased risk of high grade cervical dysplasia and cervical cancer among women infected with the HPV 16 nonprototypic variant relative to the European variant [200, 201]. Although differences between these variants in terms of viral oncogenicity or immune recognition are not as of yet fully understood, these are under investigation.

5.3.3 HPV Viral Load

Viral load, that is, the quantity of viral DNA, is currently under study. Ongoing research tries to determine the importance of this viral factor in detectable disease development. There is evidence to suggest that in women with cytological abnormalities, a high viral load is associated to high grade CIN [202]. Women with a type specific infection and a high viral load have the highest risk of persistent SIL [203]. Using a sensitive PCR assay, it has been demonstrated that carcinoma in situ associated with HPV16 occurs mainly in HPV16 women who have consistently high long term viral loads [11]. Further, in a cohort study, women with normal cytology as well as those with abnormal cytology, an increased HPV16 viral load conferred an increased risk of developing a cervical lesion [204].

Although these data indicate that women with higher viral loads may be more likely to progress to high grade dysplasia [104, 205] not all studies have supported this conclusion [206]. Moreover, the association between viral load and disease may vary between types.

5.3.4 Integration of Viral DNA

5.3.4.1 In Cervical cancer and in Lesions

In cancer of the cervix, the HPV18 viral genome is always integrated into the human genome, with HPV16 also found integrated in the majority of HPV16-related cancers [18, 22, 23, 207, 208].

In the cervical lesions, however, HPV is usually present in episomal form, but sometimes it is also found integrated in the host cell genome, particularly in premalignant lesions of CIN 2/3 [17, 208]. HPV integration frequency has been reported to increase in parallel with the severity of cervical lesions [209, 210]. In contrast, HPV DNA is commonly found extrachromosomally (in episomal form) in benign and low-grade cervical lesions [16, 25, 211, 212].

Recent studies however suggest that viral integration of HPV DNA into the human genome occurs already in early lesions and even in clinically normal epithelium [15, 28-30, 213, 214]. In addition, low risk types are very rarely found integrated in tumours. However, a recent report provided evidence of integrated genomes of the low risk type HPV11 in cancers patients with early onset recurrent respiratory papillomatosis [215].

Together, these results point to the possible relevance of integration of the viral genome in malignant transformation. Integration, occurring as a rare but early event, may thus play a role in transforming HPV-16 low grade lesions into high grade dysplasia and invasive carcinoma. As a result, integration may be a possible marker for identifying high grade lesions in patients with asymptomatic and low cytologic abnormalities.

5.3.4.2 Model of Integration

During HPV integration, the viral genome usually breaks in the E1/E2 region [18, 207], leading to the loss or fragmentation of the E1 and/or E2 regions, most commonly in the E2 protein hinge region [17]. Full length E2 protein inhibits the basal transcription of E6 and E7 ORFs. Upon HPV integration, the fragmented E2 protein is insufficient to suppress the E6/E7 gene promoter region, and this leads, at least in part, to constitutive expression of E6/E7 proteins [208]. E2 loss of the regulatory control of E6 and E7 proteins, as depicted in figure 13 below, generally results in cellular changes that lead to deregulated cell cycle control and increased genomic instability as found in high grade dysplasia and cervical cancer.

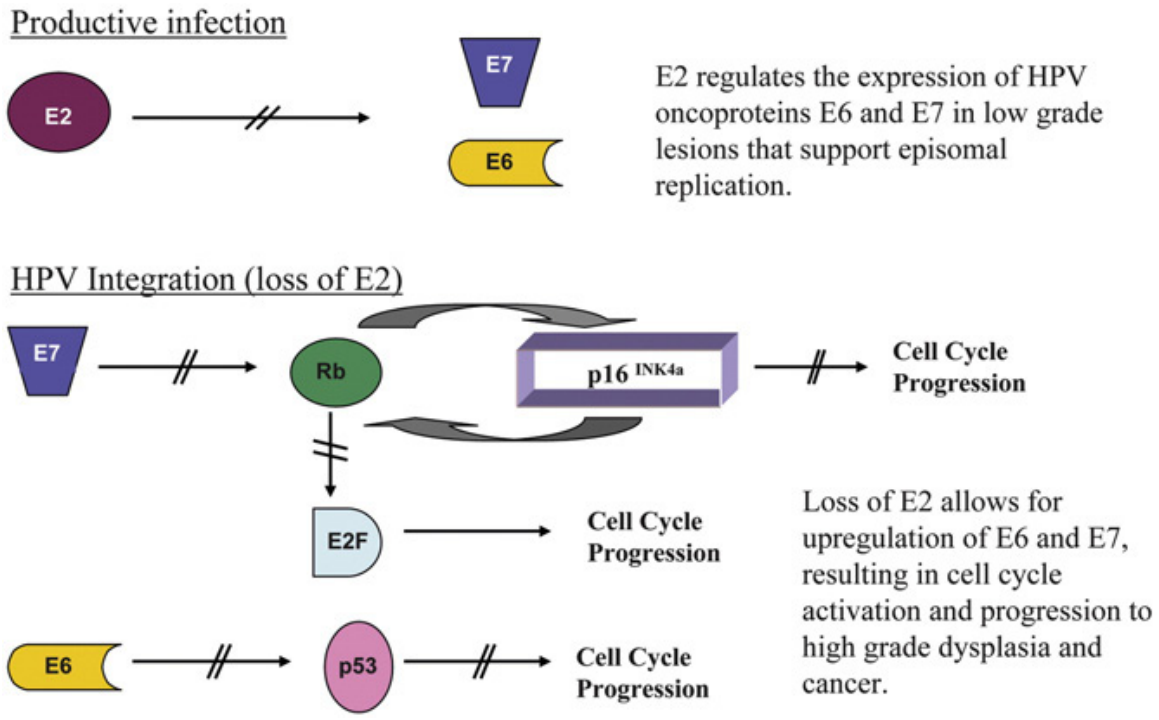


Figure 13: Regulation of effect of HPV transforming proteins

6. Techniques Available to Identify Viral Load & Integration

Integration has traditionally been studied by Southern blotting, two-dimensional gel electrophoresis, multiple displacement amplification, in situ hybridisation and PCR.

6.1 Southern hybridisation

Southern blot has been considered the gold standard for the evaluation of HPV genomes. The technique not only identifies HPV genomes in a specimen accurately and specifically, but it can also determine the physical status of the genomes, thus differentiating between episomal and integrated HPV DNA. The viral load is determined semiquantitatively by comparison to a known amount of viral DNA on gel.

The main drawback, however, with using this hybridisation procedure to detect integration of HPV is that it lacks sensitivity, and therefore requires large amounts (5-10ug) of highly purified and well preserved DNA. While cell lines or biopsy specimens provide sufficient material for Southern or in situ hybridisation, the volume of the premalignant lesion, or of exfoliated cells from the cervix only provides enough nucleic acid for PCR-based methods, leaving scientists no other alternative to determine the physical HPV status in normal and premalignant lesions of the cervix.

6.2 PCR-based methods

PCR, on the other hand, can selectively amplify HPV DNA by a series of reactions leading to an exponential and reproducible increase in viral sequences in the biological specimen. By using specific primers to the DNA template of interest, a PCR reaction can, under optimal reaction conditions theoretically produce from a single double stranded DNA molecule and after 30 cycles of amplification, 10^9 copies of the molecule or gene.

Nowadays, routine collection of specimens in liquid-based cytology solutions allows both morphological and immunohistochemical evaluation, and DNA and RNA studies can be performed for at least 14 days following sampling [216, 217]. Due to its enzymatic nature, however, the PCR-based procedures do not tolerate impurities in the sample well. Amplification of a housekeeping gene such as B-globin, B-actin or GADPH determines the adequacy of the cervical samples for PCR analysis and internal controls screen for the presence of inhibitors in real-time PCR assays. Primer–driven inhibition is eliminated in 90% of diluted cervical samples [218].

6.2.1 Real-Time QT-PCR-an indirect method

Quantitative real-time PCR permits sensitive specific detection and quantification of both RNA and DNA. The real-time PCR 5'-exonuclease Taqman method, releases and quantifies fluorescence at each amplification cycle directly proportional to the amount of amplicon generated.

The assay detects integrated and episomal forms of HPV 16 based on measurement of the absolute values of the E2 and E6 ORFs in HPV 16 positive DNA samples. A unique region of the E2 open reading frame that is most often disrupted during HPV 16 integration (the hinge region) is targeted by one set of PCR primers and a probe, and another set targets the E6 ORF.

Integration of HPV 16 into the human genome is detected by subtracting the total copy numbers of E2 representing episomal HPV from the total copy numbers of E6 (E6 detects both episomal and integrated virus). When only episomal forms of HPV are present, the specific primers of E2 and E6 should detect equivalent copy numbers. Otherwise, when only integrated form is present, E2 PCR signal should be undetectable.

PCR amplification of the E2 gene can detect integration even when the amount of sample DNA is very small [28], as that of cervical exfoliated cells, as oppposed to Southern blots and 2D gel electrophoresis which require large amounts of well preserved (high molecular weight) DNA. In addition, real-time quantitative PCR is more sensitive than Southerns, as blot hybridization risks displaying faint signals while assessing minute events of randomly integrated HPV DNA into the host genome. Southern blot and 2D gel electrophoresis techniques, however, may prove essential to confirm the physical status of HPV.

Another advantage of real-time quantitative PCR is its ability to detect the viral load and the physical state of the viral genome simultaneously, making it suitable for both screening and research purposes.

Due to the retention of E2 sequences in both the pure episomal form and in mixed (integrated and episomal) forms, quantitative real-time PCR technique assessment of integration is dependent, however, on the relative ratio of HPV-E6 to HPV-E2 amplification products (E6/E2 ratio). In order to discriminate between the pure episomal form and mixed forms, a cutoff value of the E6/E2 ratio is determined empirically, and this value varies between laboratories, making it difficult to compare integration results between research studies. Moreover, sensitive quantification of the E2 gene (or E6 gene) by the real-time PCR approach requires the target sequences to be extremely small, 87 base pairs (bp) in this study, in comparison to the entire E2 ORF of 1,097 bp. Therefore, misinterpretation of mixed forms as episomal could lead to false negative results of integration. In this research project, an E6/E2 ratio equal to or greater than 1.2 was suggestive of integration.

6.2.2 PCR-based methods to prove integration

A direct proof of HPV integration is laborious since HPV genomes are integrated at random positions in the genome and thus lack a specific sequence that can be amplified.

To identify integrated HPV genomes, new PCR-based methods have been developed recently, namely ligation mediated PCR (DIPS-PCR) [219], the amplification of papillomavirus oncogene transcript (APOT) test [9], and Restriction Site PCR (RS-PCR).

The APOT assay, an RNA-based amplification of viral-cellular fusion transcripts specific for HPV integration, is less laborious than the DNA based integration detection assays (RS-PCR and DIPS-PCR).
The main problem with techniques using RNA, however, is that it requires fresh frozen material with proper RNA quality, and is therefore less available in most biological specimens, depending on the time and type of storage conditions [220]. In addition, the APOT test is highly sensitive in detecting integration but only if there is transcriptional activity from integrants.

The restriction site PCR (RS-PCR) integration detection assay applied to the 9 exfoliated cervical specimens in this study with an E6/E2 ratio equal to or greater than 1.2, uses primers binding to specific restriction enzymes sites for amplification and sequencing of the unknown cellular sequence [221].

6.2.2.1 Restriction Site PCR

Restriction site PCR (RS-PCR) enables amplification of unknown nucleotide sequences adjacent to known nucleotide sequences. The use of RS-PCR to generate PCR products spanning HPV host junction fragments in cervical carcinomas has been used previously [222-224], and in cell cultured W12 cervical keratinocytes [225, 226]. Basically, previous studies, have amplified 100ng of genomic DNA from each sample, beginning with a biopsy or cell line (as start material).

The basic outline followed to isolate and characterize integration in cervical exfoliated clinical specimens is shown in Figure 14 below. Basically, DNA from

cervical samples is amplified by quantitative real-time PCR for the E2 and E6 ORF with their respective primer–probe pair.

Quantification of E2 and E6 DNA products allows determination of the E6/E2 ratio. Clinical samples with an E6/E2 ratio of 1.2 or greater are suggestive of integration and are thus amplified by RS-PCR to identify viral integration into the human genome.

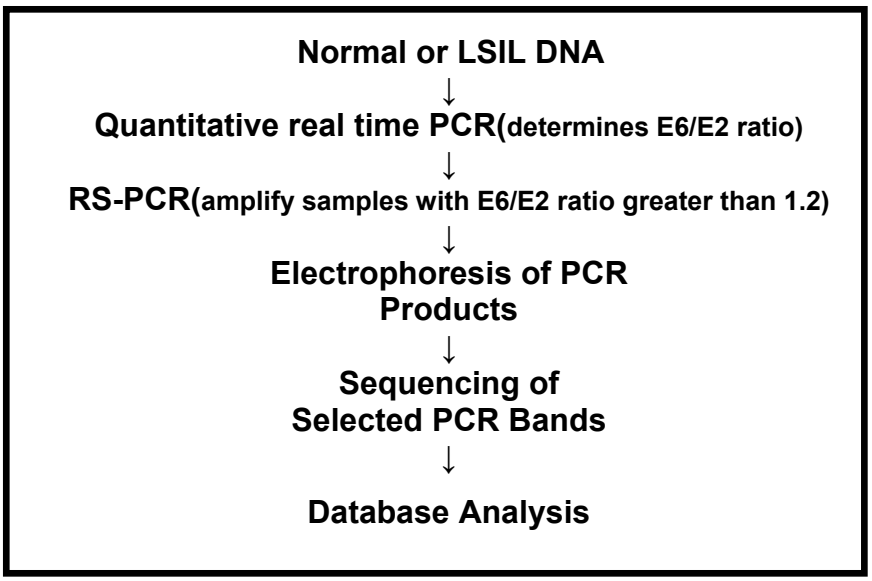


Figure 14: Schematic of the procedure used to detect integration in normal and LSIL cervical specimens with RS-PCR.

In Round 1 RS-PCR, various combinations of eight HPV 16-specific primers and six restriction site oligonucleotide primers (RSOs) are used under low-stringency cyling conditions, Figure 15. The selected RSOs have 0-2 restriction recognition sites per HPV genome.

For every cervical sample, a total of 48 HPV 16 primary RS-PCR reactions are set using every possible combination of the six RSO primers with BamH1(RSO1), SacI(RSO2), NheI(RSO3), SalI(RSO4), SapI(RSO5), ECORI(RSO6) restriction sites and the eight HPV 16-specific primers [224]. In Round 2 RS-PCR, five microliters of the 1st PCR is used as template for the nested PCR reaction, where same restriction site primer is paired with the corresponding internal HPV primer, under high-stringency conditions.

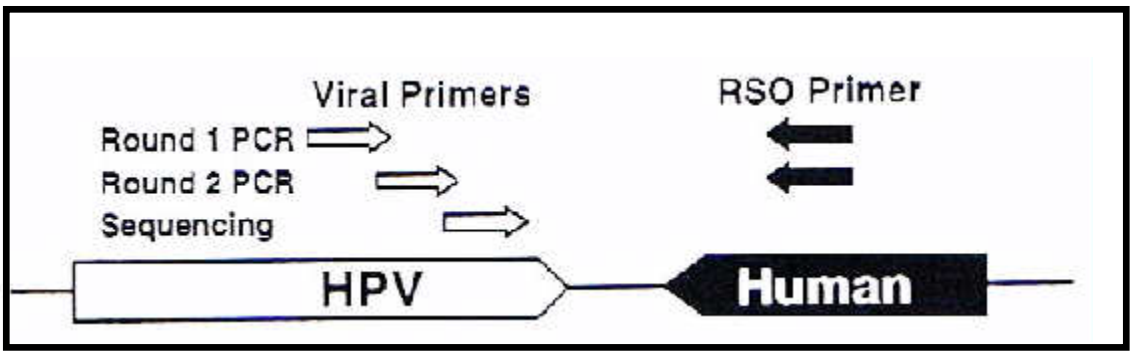


Figure 15: Schematic of RS-CR, involving two rounds of PCR, followed by sequencing of amplicons.

The amplified products from the 48 HPV 16 nested reactions are separated in a 1.5-2% ethidium bromide agarose gel as outlined in figure 14. RS-PCR products of interest, that is, not corresponding to bands of human or HPV 16 DNA are excised in a black room under ultraviolet illumination. Consequently, DNA is eluted from the gel using Qiagen Gel Extraction Kit. DNA sequencing reactions are performed using the Big Dye terminator DNA sequencing kit. As HPV16-specific sequencing primers, internal to the nested primers, were not available, PCR products were sequenced bidirectionally with T7 primer and same nested HPV-specific primer. Sequencing data is then aligned using the NCBI BLASTN-program to the human and nucleotide databases to identify homologous sequences.

STUDY OBJECTIVE

- To set up, by Taqman quantitative real-time PCR, standard curves of plasmid HPV-16 from 10^1 to 10^6 DNA copies to determine the reproducibility of E6 and E2 real-time PCR
- To assess the validity of QT real-time PCR in determining physical status, by validating E2/E6 ratios for known amounts of pBr322 HPV16 and SiHa cells(integrated HPV 16)
- To evaluate the number of copies of E2 and E6 ORFs in each specimen from which to identify and select integrants to be amplified by RS-PCR and from which to determine the episomal, integrated and total HPV 16 viral load in each sample
- To optimize RS-PCR for detection of SiHa cells, the integration positive control and to amplify by nested RS-PCR potential integrants with an E6/E2 ratio of 1.2 or greater
- To separate by agarose gel electrophoresis, sequence and map RS-PCR amplicons using NCBI BLAST alignment services to evaluate the HPV16 integration frequency in a population of young women without lesions
- To determine the association between viral load and persistance of infection
- To determine the risk factors associated with a high viral load

ROLE OF STUDENT

The student accomplished all of the above with the exception of the two last statistical objectives

ARTICLE

Human papillomavirus (HPV) types 16, 18, 31, 45 DNA loads and HPV-16 integration in persistent and transient infections in young women.

Agnihotram V. Ramanakumar [1], Otelinda Goncalves[2,3], Harriet Richardson[1], Pierre Tellier[4], Alex Ferenczy[5], François Coutlée[1,2,3], Eduardo L.F. Franco[1].

1- Division of Cancer Epidemiology, McGill University, Montreal, Que., Canada.
2- Laboratoire de Virologie Moléculaire, Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Qué., Canada.
3- Département de Microbiologie et Immunologie, Université de Montréal, Montréal, Qué., Canada.
4- Department of Family Medicine, McGill University, Montreal, Que., Canada
5- Department of Pathology and Obstetrics & Gynecology, the Sir Mortimer B. Davis-Jewish General Hospital and McGill University, Québec, Canada

Address correspondence to: E.L. Franco, DrPH, Division of Cancer Epidemiology, Department of Oncology, McGill University, 546 Pine Avenue West, Montreal, QC, Canada H2W1S6.

Conflict of interest disclosure : the authors do not have commercial or other associations that might pose a conflict of interest.

Key words: human papillomavirus, viral load, persistent, clearance, HPV-16, HPV-18, HPV-31, HPV-45, HPV integration

Word count: 4,432

Abstract (249 words):

Background. HPV burden is a predictor for high-grade cervical intraepithelial neoplasia and cancer. The natural history of HPV load in young women being recently exposed to HPV is now described.

Methods. A total of 621 female university students were followed for 2 years each 6 month. Cervical specimens containing HPV-16, -18, -31, or -45 DNA by consensus PCR were further evaluated with type-specific and β-globin real-time PCR assays. Proportional hazards regression was used to estimate hazard ratios (HR) of infection clearance. Generalized estimating equations assessed whether HPV loads was predictive of HPV infection at the subsequent visit.

Results. HPV loads were consistently higher among women <25 years old, with multiple sex partners, with multiple HPV type infections or smokers. Infection clearance was faster among women at lower tertiles of HPV-16 (HR=2.8, 95%CI: 1.0-8.1), HPV-18 (HR=3.5, 95%CI: 1.1-11.2) or combined (HR=2.4, 95%CI: 1.8-6.2) DNA loads. The relationship between HPV-16 and HPV-18 DNA loads and infection clearance followed a clear dose-response pattern, adjusting for age and number of sexual partners. Odds Ratios for HPV persistence of the middle and upper tertiles relative to the lower tertile were 2.7 and 3.0 for HPV-16 and 3.8 and 39.1 for HPV-18, respectively. There was no association between HPV-31 or -45 DNA loads and persistence. HPV-16 integration was a rare event in this cohort being encountered only in one sample.

Conclusions. The association between HPV load and persistence is not uniformly found across high-risk genital genotypes. HPV-16 integration was rarely demonstrated in young women.

Most sexually active women are infected by human papillomavirus (HPV). Since most genital HPV infections regress within two years, only a minority of women will develop persistent HPV infection that could eventually cause cervical intraepithelial neoplasia (CIN). High-grade CIN (CIN-2,3) is the immediate precursor lesion to invasive cervical cancer. Considering the natural history of HPV infection over several years, there is a need for understanding various predictors of cancer and their relevance in monitoring an array of viral outcomes.

In several recent studies, HPV-16 DNA load has been independently associated with CIN-2,3 and invasive cancer (1-6). Most studies also reported that HPV load was an ancillary marker for persistent HPV infection (3,7-11). HPV-16 or 18 infections are cleared more slowly than infections caused by other high-risk types (12). Since the biological behavior of HPV types differs, the predictive value for persistence of HPV DNA load may vary between types (13). We still know little about type-specific viral loads and their relation with clearance of HPV infection. Moreover, most studies on HPV viral load have focused on women in older women at risk for CIN. The evolution of HPV viral load in younger women recently infected remains largely unexplored.

High-risk HPV integration is considered to be a key event in the progression of CIN to invasive cancer (14). In the current model of cervical carcinogenesis, HPV-16 is integrated in CIN-2,3 and cancer. Recent data casts doubt on this concept by showing that integrated HPV-16 DNA can be detected in women with CIN-1 or no cervical lesion (15-20), although these results were not confirmed by all (21). It is important to establish whether HPV integration occurs early in the course of HPV infection to assess its contribution to carcinogenesis. Overall, the longitudinal assessment of high-risk HPV load and integration in the natural history of HPV infection considering various viral outcomes such as clearance and persistence has thus received little attention up to now. In order to improve our current understanding of the natural history of HPV infection, it would be useful to assess in young women who have initiated sexual activity the time course of HPV viral loads and occurrence of HPV integration at the early stages of interactions between the virus and the host.

In 1996, we began a prospective cohort study of the natural history of HPV infection in a population of young women attending college in Montreal, Canada, to investigate epidemiologic determinants of persistent and transient cervical HPV infections (22,23). The focus of the analyses reported here was to assess prospectively in this cohort of young women the time

course and association between HPV-16 integration, HPV-16, 18, 31 and 45 DNA loads, and type-specific viral outcomes.

Materials and Methods

Study subjects. Female students attending either the McGill University or the Concordia University Health Clinics were invited to participate if they intended to remain in Montreal for the next 2 years and had not been treated for cervical disease in the last 12 months (22). A total of 621 female university students were recruited between November 1996 and January 1999, and were followed for 2 years every 6 months. Detailed information was obtained at enrollment with a self-administrated questionnaire and changes in life-style factors were obtained at each follow-up visits with an abridged questionnaire, as described previously (22). Two cervical samples were collected with an Accelon cervical biosampler at every visit. A Papanicolaou smear was prepared with the first sampler. The remaining cells along with those collected with the second sampler were processed for HPV testing. Informed consent was obtained from all study participants. The study was reviewed and approved by the Ethics committees from each participating institution.

HPV DNA testing. HPV DNA testing has been described elsewhere (22). Briefly, five µl of sample DNA purified with QIAamp columns (Quiagen Inc, Valencia, CA) were first amplified for β-globin with PC04/GH20 primers to demonstrate the integrity of extracted DNA. β-globin-positive specimens were further tested with the L1 consensus HPV primers MY09/MY11 and HMB01 using *AmpliTaq* gold (Roche Diagnostic Systems, Laval, Canada) and with the Line blot assay for the detection of 27 genital HPV genotypes (22).

HPV-16, 18, 31 and 45 viral loads. A total of 382 specimens collected from 183 participants contained HPV-16, -18, -31 or -45 DNA. Nineteen women were infected concurrently with two or more of these genotypes. Extracted DNA from these samples was first screened for the presence of PCR inhibitors by amplification of internal controls for HPV-16, HPV-18, HPV-31 or HPV-45, and for ß-globin DNA, as described previously (18,24). The presence of PCR inhibitors was suspected when 1000 copies of at least one internal control generated a signal corresponding to less than 700 copies, as previously described (25). All samples were free of inhibitors. Two µl of processed sample were tested in duplicate for quantitation of ß-globin DNA to estimate the cell content of samples (18,24). HPV-16 E6 and HPV-18 E7 DNA was

quantitated using a standard protocol (26). HPV-31 L1 DNA was measured with the assay described by Weissenborn et al. (27). HPV-45 E6 DNA was amplified in a Light Cycler PCR and detection system (Roche Molecular Systems) in a twenty-µl reaction mixture containing 1x DNA Master Hybridization Probe Mix with the Fast Start *Taq* DNA polymerase (Roche Molecular Biochemicals), 0.3 pmoles of each HPV-45 primer 45 E6-F (nucleotide position 463-486; 5'-TTAAGGACAAACGAAGATTTCACA-3') and 45 E6-R (nucleotide position 670-647; 5'-ACACAACAGGTCAACAGGATCTAA-3'), and 50 nM of fluorogenic 45 E6-TM probe (nucleotide position 491-514; FAM-5'-AGCTGGACAGTACCGAGGGCAGTG-3'-TAMRA). Cycling parameters included an activation step at 95ºC for 10 min followed by 50 cycles at 95ºC for 15 sec, 60ºC for 5 sec and 65°C for 45 sec. For each of the four genotypes analyzed, cycle thresholds obtained for each sample were compared to those of a titration curve obtained by serial ten-fold dilutions of HPV-16, 18, 31 or 45 plasmids in a fixed amount of 75 ng of human genomic DNA (Roche Diagnostics) in 10 mM Tris-HCl [pH 8.2]. Each assay consistently detected 10 HPV DNA copies (data not shown). HPV viral loads were expressed as the number of HPV DNA copies per cell.

HPV-16 integration assays. The presence of integrated HPV-16 was investigated with real-time PCR assays targeting E6 and E2, as previously described (28). Since HPV integration often results in the disruption of the E2 gene, detection of a greater quantity of HPV-16 E6 compared to HPV-16 E2 strongly suggests the presence of integrated HPV-16 DNA (16,17). Two µl of each processed sample was tested in duplicate in each HPV-16 E6 and E2 assays. Cycle thresholds were compared to those of serial ten-fold dilutions of an HPV-16 plasmid in a fixed amount of 2,000 copies of human genomic DNA (Roche Diagnostics). The presence of integrated HPV-16 DNA was suspected for specimens with ratios of HPV-6 E6 and E2 copies (HPV-6 E6/E2) at or above 2, as previously discussed for types 16 and 33 (18,28,29).

HPV-16 integration was confirmed by restriction site PCR (RS-PCR), a sequencing technique that demonstrates the presence of HPV-16 and human DNA junctions in the same amplicon (30). Briefly, 9 HPV-16-specific primers spanning the entire HPV-16 genome were combined separately with 6 restriction site oligonucleotides designed to anneal on selected restriction sites on the human genome, in a two-step hemi-nested PCR performed in a 9600 Thermal Cycler. Amplicons were migrated in a 2% agarose gel stained with ethidium bromide. When visible bands were obtained, direct double-stranded PCR-sequencing was done by a cycle-

sequencing method (BigDye terminator ready reaction kit, Perkin-Elmer) using the internal primers and a sequencing primer on 20 ng of purified amplicon.

Statistical Methods. Cross-sectional correlations were calculated using the Pearson correlation coefficient (r) among viral load measurements at entry and follow-up visits for all four HPV types. Geometric means were calculated as a function of selected characteristics reported at the index visit of first positivity for a specific HPV type. The Kaplan Meir technique was used to estimate the cumulative probability of infection clearance as a function of time (length of follow-up). Proportional hazards regression was also used to estimate the clearance of type-specific infection, stratified by tertiles (33.33%) of their viral load distributions (31). Logistic generalized estimating equations (GEE) was utilized to assess whether a viral load measured at a given visit was a predictor of persistent HPV infection at the subsequent visit (32,33). Exponential coefficients of crude and adjusted odd ratios for age and sexual partners between persistent HPV at visit (t) and viral load at visit (t-1) within specified periods of follow-ups were presented to reveal the associations.

Results

The 621 women enrolled in the McGill-Concordia cohort contributed to 2650 completed visits (mean of 4.3 visits/subject) during an average of 21.5 months of follow-up. Nearly all cervical specimens (n=2,570, 97.6%)) were suitable for HPV testing (22). The mean age of participants was 23 years (median age, 21 years; age range, 17-42 years). Nearly half of participants had had more than 4 lifetime sexual partners (22). The prevalence, incidence and mean duration of HPV infections have been reported elsewhere (22). High-grade squamous intraepithelial lesion (SIL) was shown on 4 smears while of low-grade SIL was shown on 49 smears, precluding the analysis of association between HPV loads and SIL.

Samples positive for HPV-16 (n=220 from 104 women, mean of 2.1±1.2 samples/woman), -18 (n=80 from 43 women, mean of 1.9±1.1 samples/woman), -31 (n=75 from 36 women, mean of 2.1±1.2 samples/woman), or -45 (n=33 from 19 women, mean of 1.7±1.2 samples/woman), were further tested with type-specific quantitative real-time PCR assays. Descriptive statistics of log-transformed HPV loads stratified by baseline and follow-up time points are presented in Figure 1. We observed no significant intra-patient diversity between two subsequent HPV-16 DNA loads using a Wilcoxon rank sum test between medians [p=0.831 for

visits (1,2); p=0.127 for (2,3); p=0.340 for (3,4); p=0.731 for (4,5)] . The other three HPV types exhibited some intra-patient diversity between observed time points. No particular trend was observed with viral load groups at different time period.

To investigate the consistency of HPV load measurements across the cohort, correlations of type-specific HPV load measured at baseline and at various follow-up visits were calculated (Table 1). HPV loads were significantly correlated only when considering neighboring visits. The strength of association was diluted as time progressed. Stronger correlations between consecutive visits were found in women with positive infection with HPV-16. Correlation coefficients of HPV loads were non-significant when visits 12 months apart were compared, except when all four HPV types were combined (Figure 2).

HPV loads stratified by various selected characteristics at the index visit of HPV-positivity for each given HPV type are provided in table 2. To assess if some of these factors were determinants for higher HPV DNA loads, we compared the geometric means of log-transformed HPV load values with 95% confidence intervals. HPV DNA loads were consistently slightly higher among women younger than 25 years of age, with > 2 lifetime sex partners, with sequential HPV co-infections but not those with concurrent co-infections, and current smokers. For all types except HPV-16, HPV DNA loads were also higher in past smokers. The influence of the use of contraceptives (condom or oral contraceptive) on HPV DNA load was not consistent across the four types (Table 2), but suggested that the use of one of these contraceptive methods was associated with a higher HPV DNA load. The biggest difference of HPV DNA loads for all types studied was found when comparing women with one partner and those with more than two lifetime sexual partners.

Table 3 and Figure 3 show the pattern of the duration and clearance of HPV infection according to the HPV DNA load measured for each genotype. Once infected, 30% of HPV-16, 44% of HPV-18, 63% of HPV-31 and 55% of HPV-45 infections cleared within 6 months. Infection clearance was faster among women with lower tertiles of HPV viral loads. Relative to upper tertiles of HPV load, the age-adjusted hazard ratios of lower tertiles were 2.8 (1.0-8.1), 3.5 (1.1-11.2) and 2.4 (1.7-3.5) for HPV-16, HPV-18 and all four types combined, respectively. The relation between HPV loads and infection clearance followed a clear dose-response pattern for HPV types 16, 18 and all four types combined, even after adjusting for age and lifetime number

of sexual partners. There were no clear association between HPV-31 and HPV-45 HPV loads and clearance of these infections.

The association between HPV viral load and persistence of infection was investigated for each genotype by calculating crude and adjusted (age, number of sexual partners) GEE Odds Ratios with exponential coefficients (Table 4). Odd ratios for HPV persistence were 2.7 (1.1-6.3) and 3.0 (1.3-7.0) for middle and upper tertiles of HPV-16 DNA loads, respectively, compared to lower tertiles. Similarly, odd ratios for middle and upper tertiles of HPV-18 DNA loads were 3.8 (0.8-18.6) and 39.1 (5.1-302.0), respectively. These associations were ambiguous for persistence of HPV types 31 and 45. In order to avoid arbitrary selection of tertiles, we plotted a Receiver Operating Characteristic curve (ROC) by considering measured HPV-16 load continuously (Figure 4). The area under the curve was 0.6544, also suggesting that HPV load at visit (t-1) is a predictor for HPV-16 infection at visit (t).

We then investigated if HPV-16 infections in young women resulted in HPV-16 integration into the human genome. Forty eight specimens were excluded of this analysis because they contained <15 copies of HPV-16 DNA per µl of extracted DNA. HPV-16 E6 and E2 DNA were thus quantitated on 172 HPV-16-positive samples. The mean HPV-16 E6 viral load was 57.5±324.6 DNA copies per cell (95% confidence interval, 8.6-106.4; median, 0.92, range 0.0001-4084.7 copies per cell). The mean HPV-16 E6/E2 ratio was 0.97 ±0.25 (95% confidence interval, 0.93-1.01; median, 0.97; range, 0.5-2.48). Although 2 samples had HPV-16 E6/E2 ratios $\geq$ 1.5 and <2.0 (1.54 and 1.70), samples collected at consecutive visits from these two participants yielded HPV-16 E6/E2 ratios near or below 1.0 (data not shown). One sample from one woman at her last visit generated a HPV-16 E6/E2 ratio of 2.48 with a HPV-16 E6 DNA load of 0.94 copy/cell. HPV-16 was detected only at the last of five visits attended by this participant. Normal cytology smears were obtained at the first four visits while a low-grade SIL smear was obtained at the fifth visit. RS-PCR was performed on the 18 samples that generated a HPV-16 E6/E2 ratio $\geq$ 1.2. Despite using several primer combinations, we could not demonstrate the presence of cellular and HPV junctions in any of the samples tested. A minimal amount of 35 ng of cellular DNA per test was analyzed for the only sample with a ratio > 2, unsuccessfully, which could have limited our ability to sequence HPV-human DNA junctions.

Since we detected HPV-16 integrated forms only once, we investigated if quantity of cellular DNA introduced in the quantitative assays hampered our ability to measure HPV-16 E6

and E2, as reported by another group (34). When mixtures of episomal HPV-16 and DNA extracted from SiHa cells were tested, we observed interference with quantitation of HPV-16 E2 and E6 with DNA extracted from 10^6 and from 10^4 SiHa cells, respectively (data not shown). One thousand copies of episomal HPV-16 DNA was detected without loss of signal when tested in a mixture of DNA extracted from up to 200,000 cells in the HPV-16 E2 assay and up to 40,000 cells in the HPV-16 E6 test (Table 5). Interference of HPV-16 quantitation by input DNA was not an issue in our study since all samples tested contained $\leq$ 39,200 copies of cellular DNA per test.

Discussion

In this report, HPV DNA load was measured for four high-risk types with real-time PCR on a set of samples collected prospectively in young women. These four genotypes are amongst the most frequently detected in cervical cancer. In opposite to cross-sectional studies of older women, the four HPV genotypes were detected at similar loads (1,11,35). HPV loads were not substantially different between women with single and multiple type infections, except for HPV-31 loads. It is still unclear if multiple type infections are caused by a selective immune deficit against HPV infection, by exposure to HPV before an efficient humoral response or by exposure to multiple partners or partners with multiple type infections.

In our study, the quantitative real-time PCR assays utilized to estimate HPV loads were specific and reproducible (25,28). The number of HPV DNA copies was normalized for cell content by quantitation of β-globin DNA. The HPV-16 integration assay was devised considering the genetic diversity of HPV-16 (28,36). Using type-specific quantitative assays allowed isolating the effect of HPV type loads in multiple type infections. Consistent measurements for HPV types 16, 18, and 31 were shown for the five visits. The HPV-16 viral load assay was the steadiest and most precise. The HPV-45 loads were the most tangled but few cases were infected by HPV-45. The cohort design also allows testing time-dependent correlations by distinguishing the differences within or between pair visits. Such models reveal two types of correlations: cross-sectional familial correlation between pairs from the same HPV type and serial correlation among repeated measures for the same pair. Among the measured viral load pairs, as expected, the correlations were stronger in the neighboring visits, especially for HPV-16 and 18. This is an indication of reproducibility of laboratory results. In addition, the

weak correlations over a 12 month period suggest that most of the index infections are probably transient after one year.

As reported by others, younger woman (< 25 years) harbored higher HPV loads (1). These younger women were possibly exposed to HPV while they were immunologically naïve to HPV. In the current study, we observed a greater HPV burden among current and former smokers. This finding could be due to a defective cell-mediated immunity against HPV induced by tobacco (37). Results from this cohort as well as others suggest that tobacco smoking increases the duration of high-risk HPV infection (23,38). This could be explained in part by the increased replication of HPV in women exposed to tobacco.

Although the regular use of condoms protects against most sexually transmitted infections, they are not as efficient against HPV infection (39). We found a trend with the consistent use of condoms for having higher HPV loads for all four genotypes. These results could be biased by the fact that condoms could be associated with risky sexual behavior or exposure to new sexual partners (23). Condom use has been associated in one study with regression of CIN and clearance of HPV (40). Although oral contraceptive use did not modify the duration of high-risk HPV infection in our cohort (23), HPV-18 DNA loads were markedly increased in women using oral contraceptive. Oral contraceptives may also be a proxy for a higher number of sexual partners.

HPV-16 and -18 loads were good predictors of the duration of infection in opposite to HPV-31 and -45. HPV-16 load has been reported by others to be a stronger predictor for persistence or lesions than HPV-18, 31 or 33 loads (13). There was a clear dose-response relationship between HPV load and persistence of HPV-16 and HPV-18 infections. The current investigation adds the concept that clearance rates depended largely on the level of HPV load. Viral-hosts interactions play a determinant role in the clearance of viral infections (41). HPV has developed several mechanisms to evade the host immune system (42). Functional differences between HPV-encoded proteins could also explain why some types and variants have a better viral fitness with a greater ability to persist (14,41,43). HPV loads were greater with HPV co-infections at different visits (sequential) than concurrent co-infection. In recent studies (44,45), two or more oncogenic HPV types diagnosed concurrently did not confer an additional risk of developing lesions. All but one study confirmed that sustained or increased viral loads, especially for HPV-16, were predictive of persistent infection (3,5,8,11). In a cohort of nearly

6,000 women in France, women with HPV loads above 10 pg/ml were less likely to clear the infection, irrespective of the age of participants (8). Similarly, another cohort study conducted in the Netherlands reported that women with low HPV-16 loads were five times more likely to clear HPV-16 infection (5). In a third study conducted in Brazil, there was a dose-response relationship between increasing viral loads and risk of incident abnormal smear over time (3). (3,5,8,46).

HPV-16 integration often disrupts the E2 gene, resulting in uncontrolled expression of HPV-16 oncoproteins (14,17). Quantitation of load of integrated HPV-16 forms could be a better biomarker for CIN-2,3 than HPV load, although integration of HPV-16 or load of HPV-16 integrated forms has been at best weakly associated with CIN-2,3 (15,19,20,28,47). One study investigated the rate of HPV-16 and 18 integration in women aged 15 to 19 years old (48). Disruption of the E2 gene was demonstrated in up to 25% of incident HPV-16 infections, suggesting that HPV-16 E2 disruption was a common event occurring early during the infection. The higher detection rate of HPV-16 integration compared to our study could be related to different risk factors for HPV infection in this cohort. The entire E2 gene was studied for disruption in that study while our assay focused on the hinge region of E2. Disruption of E2 could occur more frequently at the very early phases of infection in younger women while half of participants to our cohort had had more than four sexual partners at recruitment. In another study on older women followed longitudinally, nearly 50% of women with persistent or transient infections also had mixed integrated and episomal forms (11). In opposite to our study, over 50% of these participants had Pap smears anomalies. Surprisingly, HPV integration was not found to be associated with persistence. Results from other cohorts are thus awaited to assess the rate and persistence of HPV integration at early stages of infection.

We recognize that there are some limitations in our study. Few women had abnormal smears, reducing our power to test the associations between HPV load and lesion outcomes. The time interval between visits can influence the assessment of persistence and clearance. The majority of women in our cohort returned within 6 months of each visit and there was only a small proportion of women whose time interval between visits extended beyond a year (22). The association between higher viral loads and persistence would only be distorted if it had been associated differentially with time between visits. Given that the participants were unaware of their HPV and viral load status, this is an unlikely scenario. It is also unlikely that their behavior

and other risk factors will change in this short span of time. The same associations between HPV loads and persistence were obtained when HPV persistence was defined more stringently by using three 3 consecutive HPV-positive visits for the same type.

Consecutive detection of HPV DNA is due to either ongoing viral replication or reactivation of latent infections (49). The design of our study can not discriminate between these possibilities. Participants considered as having persistent infection could have been reinfected with another isolate of the same HPV type. This is however unlikely since women with persistent infection were all infected with the same HPV variant (manuscript in preparation). Both prevalent and incident HPV infections were included in our analysis of persistence, increasing the power of our analyses. Though it may not have a direct implication on HPV clearance, the exact duration of prevalent cases is unknown. We could have introduced a survival bias because a greater proportion of prevalent HPV infections represent persistent infections compared to incident infections. However, our conclusions did not change when we restricted our analyses to incident infections. We also analyzed the entire cohort to utilize the clustered binary outcome of persistence, by using the 'visit number' as panel variable. This method increases the power of our analyses by fitting logistic generalized estimate equations. The odds ratios were adjusted for age and numbers of sexual partners, and were in the same direction of hazard ratios of clearance stratified by viral load tertiles. So we assume that the bias incurred due to inclusion of prevalent cases is also minimal. We also doubt that misclassification of HPV status have significantly occurred in our cohort, since there were very few women with persistent type-specific infections who had an intervening visit with a negative HPV test result and more than 80% of the same type persistent infections occurred during consecutive visits (22).

Apart from studying type-specific viral loads, we also investigated the association between persistence and combined viral load results. Cumulative viral loads may be a marker for the presence of multiple type HPV infections which are associated by themselves with high-grade lesions (50). The results obtained by combining viral loads from various types need to be interpreted with caution as this strategy poses difficulties in terms of tertile cut-offs. Receiver Operating Characteristics (ROC) curves were used to check the sensitivity of these arbitrary tertile cut-offs.

We can not exclude the possibility that HPV-16 was disruption during integration in areas outside of the E2 hinge region, could not be excluded. This could result in the underestimation of

the prevalence HPV-16 integration in our study. Nevertheless, HPV-16 E2 hinge is the most frequently disrupted site in studies conducted in North America (17). Our assay would also be falsely negative in cases where multiple copies of HPV-16 integrated into head-to-tail tandems in which only flanking HPV-16 sequences are disrupted. Although the analytical sensitivity of real-time PCR for quantitation of HPV-16 E6 and E2 DNA is excellent, the clinical sensitivity and specificity of these assays to detect HPV-16 integration have not been thoroughly assessed. We could not confirm the presence of HPV-16 integration with a standard technique identifying the presence of HPV-human DNA junctions in the only sample with a HPV-16 E6/E2 ratio above 2.0. However, the quantity of sample that could be analyzed was limited. RS PCR is a tedious procedure sometimes difficult to interpret. A recent report using a similar technique to demonstrate the presence of HPV-human junctions did not find HPV-16 integration in specimens from women with low-grade SIL (21). Real-time PCR assays are interesting techniques to detect integrated HPV forms but further studies on a greater number of specimens using in parallel several techniques for detection of integrated HPV-16 need to be conducted. Until then, we should interpret these results obtained with these assays.

In conclusion, this study demonstrates a clear association between HPV load and persistence of HPV-16 and18 infections in young women at the early stages of their sexual life. This association will depend on the HPV type studied. More longitudinal studies are needed to clarify the onset of HPV integration and its relationship with disease progression.

Acknowledgement:

This study was supported by the Canadian Institutes for Health Research operating grants and a CIHR Team grant support.

Table 1: Between-visit correlation (r) of HPV load measurements by HPV type at entry and follow-up visits

HPV type	Visit	Correlation with visit at			
		6 months	12 months	18 months	24 months
16	Entry	0.4789* (29)	0.3781 (20)	0.2046(14)	0.4583 (6)
	6 months		0.7561* (27)	0.6764* (21)	0.4282 (10)
	12 months			0.8295* (30)	0.5810+ (17)
	18 months				0.8043* (24)
18	Entry	0.0103 (10)	0.5481 (5)	--	--
	6 months		0.3132 (11)	0.6612 (5)	--
	12 months			0.0173 (8)	--
	18 months				0.9949+ (3)
31	Entry	0.9562* (6)	0.4242 (5)	0.8583 (4)	--
	6 months		0.2427 (6)	0.8063+ (5)	0.5138 (5)
	12 months			0.1930 (11)	0.3574 (7)
	18 months				0.7289* (12)
45	Entry	--	--	--	--
	6 months		--	--	--
	12 months		--	--	--
	18 months				--

Combined HPV types (16, 18 31,45)	Entry	0.3200+ (48)	0.4208+ (36)	0.3647+ (23)	0.4753 (11)
	6 months		0.5491* (46)	0.5562* (34)	0.3505 (16)
	12 months			0.5391* (48)	0.3463+ (27)
	18 months				0.7778* (36)

HPV loads were measured with type-specific PCR assays. * P value is at 1% level of significance +P value is at 5% level of significance

Pearson correlation coefficients (r) were presented with number of subjects in the parenthesis

Table 2: Geometric means of HPV viral loads as a function of selected characteristics at the first occurrence of positivity for a given HPV type

Characteristics*	HPV 16			HPV 18			HPV 31			HPV 45		
	N	Mean	95% CI	N	Mean	95% CI	N	Mean	95% CI	N	Mean	95% CI
Age <25	72	4.41	3.98-4.89	32	4.52	3.63-5.65	31	5.22	4.69-5.83	14	4.95	4.16-5.89
Age 25+	31	3.97	3.30-4.78	11	4.01	3.34-4.82	5	4.55	3.41-6.08	5	4.59	3.24-6.48
No co-infection #	10	3.95	2.56-6.07	5	1.64	0.26-10.26	4	7.06	5.46-9.12	3	4.87	1.23-19.3
Co-inf-sequential	56	4.71	4.28-5.18	21	5.26	4.52-610	17	4.99	4.44-5.61	8	5.16	4.26-6.24
Co-inf-concurrent	38	3.77	3.17-4.47	17	4.45	3.76-5.28	15	4.96	4.14-5.95	8	4.64	4.15-5.19
No smoking	60	4.33	3.87-4.84	23	3.91	2.92-5.25	22	4.90	4.27-5.63	11	4.72	3.57-6.24
Current	25	4.61	3.84-5.51	15	4.94	4.12-5.91	8	5.11	4.25-6.14	3	5.03	3.83-6.60
Former	18	3.66	2.78-4.84	5	5.21	2.71-10.01	6	6.42	5.55-7.44	5	5.00	4.58-5.47
No Condom use	15	3.57	2.64-4.82	10	4.50	3.08-6.57	4	4.63	2.68-7.97	3	3.23	0.55-18.9
Used sometimes	44	4.60	4.02-5.26	17	4.03	2.77-5.87	13	5.75	4.90-6.75	12	4.95	4.38-6.15
Regular use	39	4.54	4.03-5.13	15	4.84	3.94-5.96	16	5.01	4.28-5.87	5	5.12	4.50-5.82
No OC Use	23	4.09	3.28-5.10	13	4.37	3.59-5.33	10	5.36	3.99-761	7	4.68	4.02-5.45

Use Sometimes	63	4.39	3.95-4.88	25	4.67	3.99-5.46	21	4.95	4.32-5.66	10	4.92	3.88-6.23
Regular use	7	3.95	3.45-4.64	4	7.42	7.21-7.64	4	6.91	0.85-36.1	2	5.13	--
One Partner	5	3.53	2.02-6.16	5	2.46	0.55-10.87	2	5.03	3.12-8.12	5	3.16	0.52-5.81
2-3 Partners	24	4.79	4.23-5.42	7	5.15	3.52-7.52	13	5.51	4.54-6.67	2	5.34	-
4+ Partners	75	4.17	3.72-4.67	31	4.68	4.09-5.34	21	4.99	4.40-5.66	12	5.39	4.12-6.71

*above variables are having updated information at each visit. There are treated like dynamic variables at the time of first occurrence of the HPV positivity.

#we defined *concurrent* co-infection as the detection of more than one HPV type in the cervical cell specimen collected at a given visit. We defined *sequential* co-infection as infections with multiple HPV types detected at different visits (Thomas 2000)

Table 3: Hazard Ratios of HPV Clearance from Cox Regression Models, Stratified by Various Levels of Viral Load

HPV type	Viral load level	No. Eligible cases/ No. of Events	HR (CI 95%)	
			Unadjusted	Adjusted
	Tertiles 3	26/6	1.0	1.0
16	Tertiles 2	28/11	1.94 (0.71-5.27)	1.95 (0.72-5.31)
	Tertiles 1	27/9	2.61 (0.23-7.38)	2.79 (0.96-8.12)
	Tertiles 3	14/4	1.0	1.0
18	Tertiles 2	12/3	2.03 (0.45-9.08)	2.21 (0.46-9.82)
	Tertiles 1	13/11	3.39 (1.08-10.71)	3.49 (1.09-11.19)
31	Tertiles 3	9/6	1.0	1.0
	Tertiles 2	11/5	0.86 (0.15-6.03)	0.89 (0.19-6.80)
	Tertiles 1	9/0	--	--
	Tertiles 3	6/4	1.0	1.0
45	Tertiles 2	6/3	0.54 (0.11-2.72)	0.55 (0.11-2.74)
	Tertiles 1	6/3	0.90 (0.30-3.50)	0.92 (0.34-3.57)
Combined HPVs	Tertiles 3	54/20	1.0*	1.0*
(16, 18, 31, 45)	Tertiles 2	58/23	1.54 (0.85-2.81)	1.59 (1.11-2.40)

Tertiles 1	55/23	2.28 (1.25-4.17)	2.42 (1.68-3.51)

*Since the observations are not independent within strata, we used sandwich robust variance estimator of variance in place of the conventional calculations for confidence interval. Adjusted hazard ratios were adjusted for age and sexual partners.

Table 4: Odd ratios for associations between persistent HPV at visit (t) and viral load at visit (t-1) within specified periods of follow-up, McGill-Concordia Cohort study (GEE model with exponential coefficients)

HPV type	Viral load	No. of Events	OR and 95% CI	
			Crude	Adjusted*
16	Tertiles 1	29	1.0	1.0
	Tertiles 2	44	2.59 (1.13-5.93)	2.66 (1.13-6.29)
	Tertiles 3	42	2.94 (1.28-6.73)	2.99 (1.28-6.97)
18	Tertiles 1	6	1.0	1.0
	Tertiles 2	11	1.77 (0.53-5.88)	3.81 (0.79-18.6)
	Tertiles 3	20	17.1 (3.67-79.7)	39.1 (5.08-302)
31	Tertiles 1	22	1.0	1.0
	Tertiles 2	18	0.48 (0.09-2.46)	0.48 (0.09-2.61)
	Tertiles 3	16	0.19 (0.04-0.90)	0.19 (0.04-0.97)
45	Tertiles 1	6	1.0	1.0
	Tertiles 2	3	0.18 (0.05-8.46)	1.62 (0.07-353)
	Tertiles 3	5	1.00 (0.04-22.8)	1.30 (0.08-174)

Events are considered as HPV outcome in the visit (t), where t=2, 3, 4, and 5. Viral load cut-offs at visit (t-1), where t-1=1, 2, 3, 4 were used as predictors of out come.

*ORs were adjusted for age and number of lifetime sexual partners

Table 5. Interference of background human DNA in quantitation of HPV-16 DNA with HPV-16 E6 and HPV-16 E2 real-time PCR assays.

No. of copies of human DNA	Quantitation of 1000 copies of episomal HPV-16 DNA with real-timer PCR assays	
	HPV-16 E2	HPV-16 E6
0	933±13	1150±39
$8x10^3$	1080±35	1114±7
$1x10^4$	1097±38	1179±148
$2x10^4$	1026±149	1109±59
$4x10^4$	1125±15	1015±46
$6x10^4$	1173±10	820±30
$8x10^4$	1151±100	726±19
$1x10^5$	1082±82	787±32
$2x10^5$	1071±38	562±20

1000 copies of HPV-16 DNA were amplified in a background of various copies of human DNA and detected in real-time PCR assays for quantitation of HPV-16 E6 and HPV-16 E2. Results are means ± 1 standard deviation of duplicate values. Repeat experiments gave similar results.

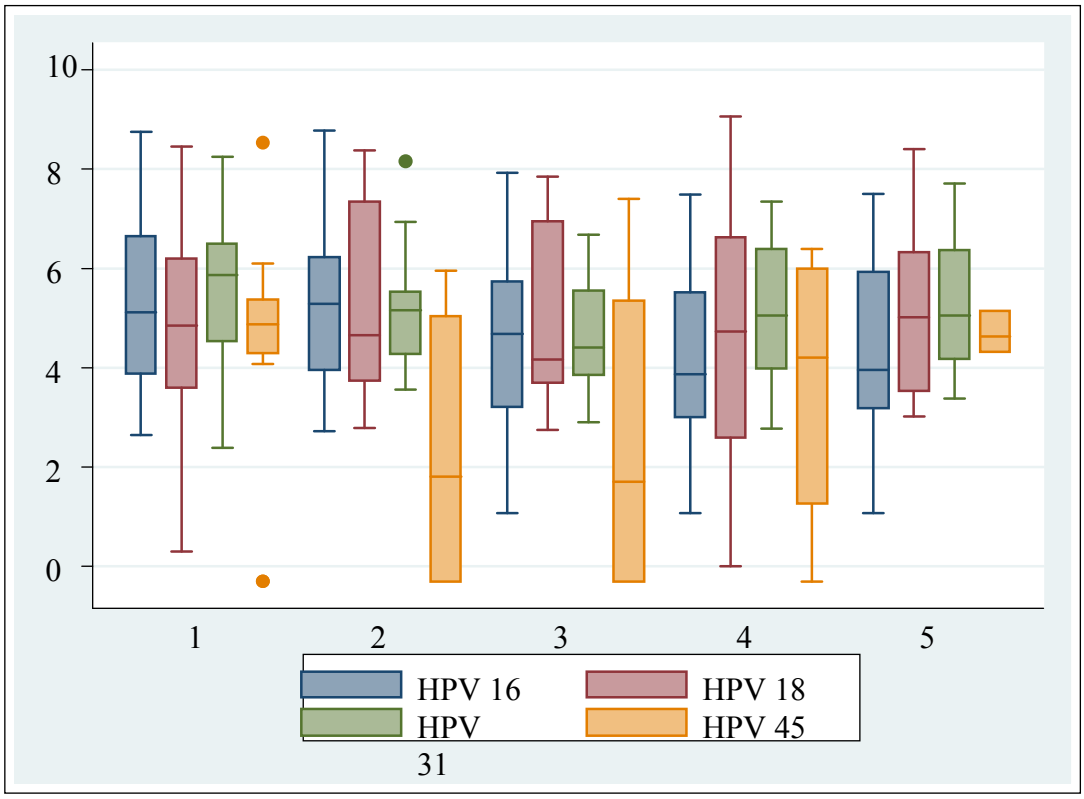


Figure-1: Log-transformed HPV loads (HPV DNA copies per cell) at recruitment and at follow-up Visits among HPV-positive women for the 4 genotypes studied. The length of each box corresponds to the interquartile range, with the top boundary of the box representing 75th and bottom boundary the 25th percentile. The horizontal line in the box indicates the median value. Outlier values are shown in circles outside the boxes.

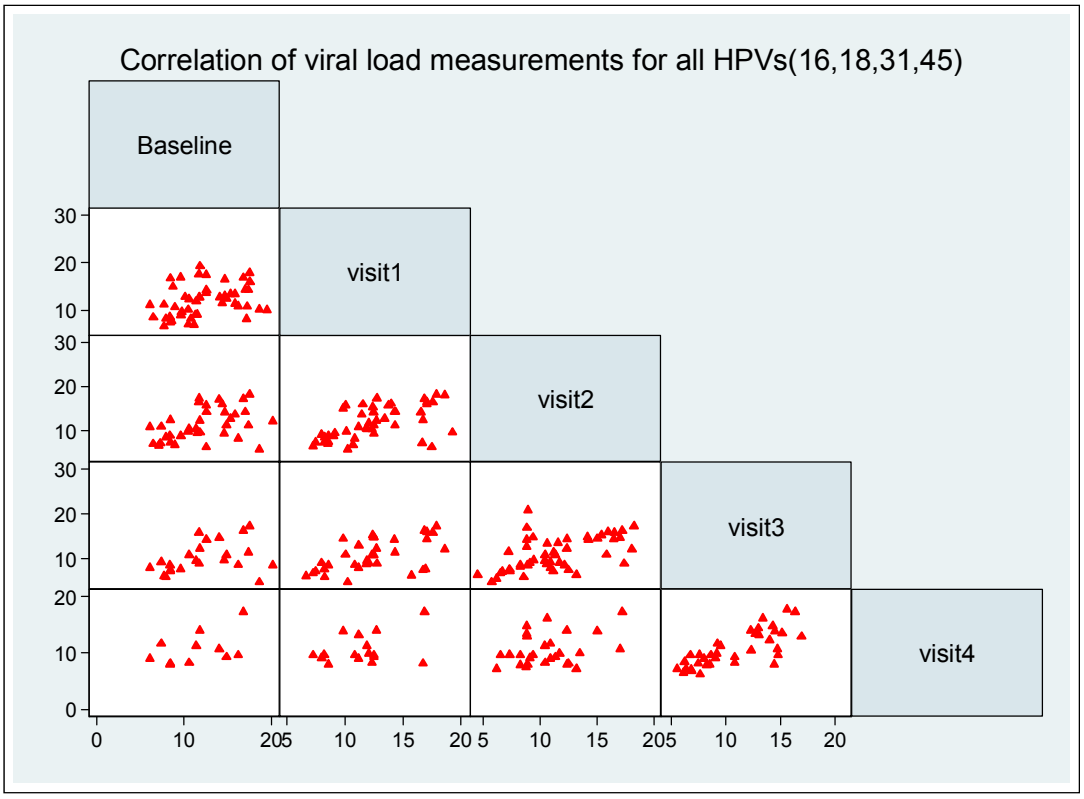


Figure-2: Correlation matrix for viral load measurement of four combined HPV types (HPV-16, 18, 31, 45) at accrual and follow-up (see material and methods).

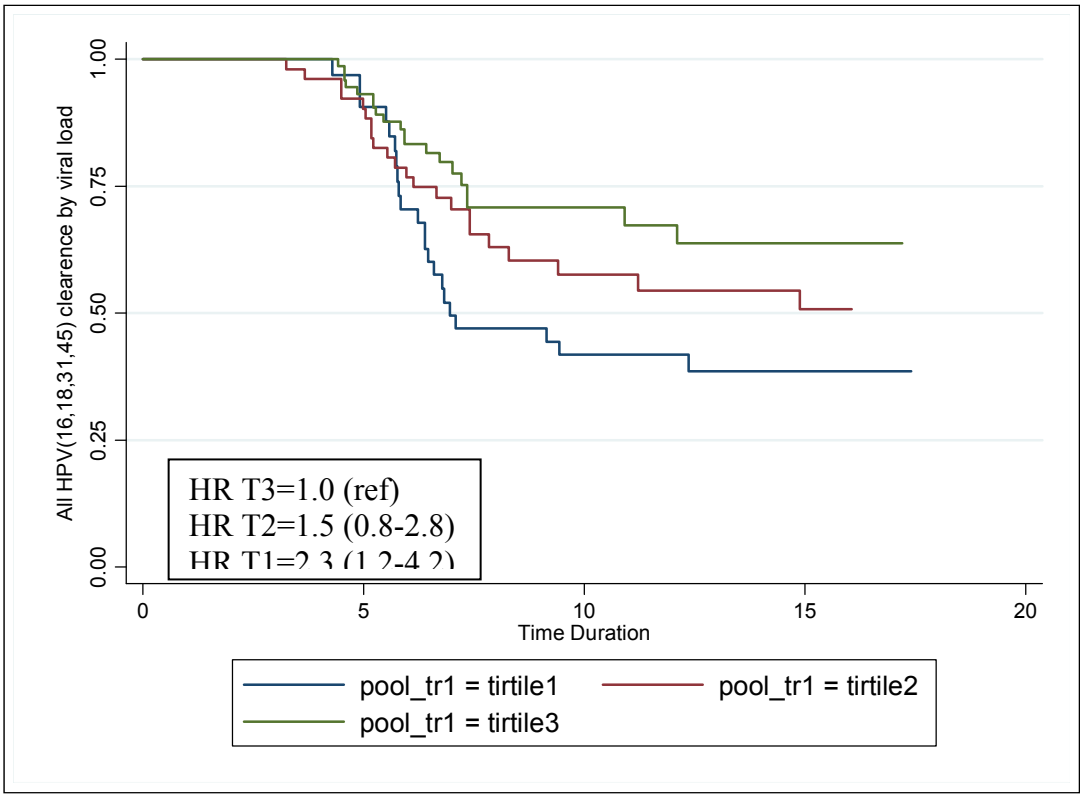


Figure 3: Combined HPV (16, 18 31 45) clearance stratified by tertiles of viral load assuming multiple events per individuals (unit of analyses is infection). Since the observations are not independent within strata, we used robust variance estimator of variance in place of the conventional calculations for confidence interval. Hazard Ratios and confidence intervals are shown in the above text box.

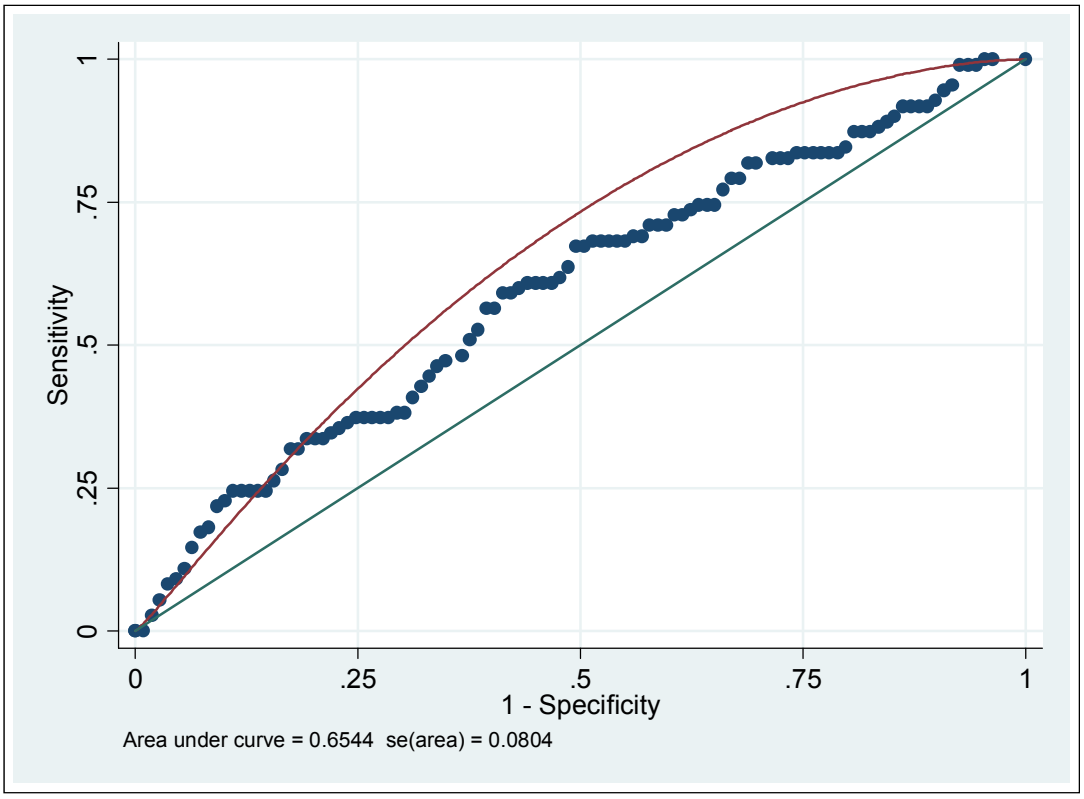


Figure-4: Predicted ROC curve between the viral load (continuous) at visit t-1 and persistent HPV 16 infection at visit (t) within specified periods of follow-up. Red line is the fitted line based on the plotted values.

Reference List

1. **Flores, R., M. Papenfuss, W. T. Klimecki, and A. R. Giuliano**. 2006. Cross-sectional analysis of oncogenic HPV viral load and cervical intraepithelial neoplasia. Int.J Cancer **118**:1187-1193.

2. **Tsai, H. T., C. H. Wu, H. L. Lai, R. N. Li, Y. C. Tung, H. Y. Chuang, T. N. Wu, L. J. Lin, C. K. Ho, H. W. Liu, and M. T. Wu**. 2005. Association between quantitative high-risk human papillomavirus DNA load and cervical intraepithelial neoplasm risk. Cancer Epidemiol.Biomarkers Prev. **14**:2544-2549.

3. **Schlecht, N. F., A. Trevisan, E. Duarte-Franco, T. E. Rohan, A. Ferenczy, Villa, LL, and E. L. Franco**. 2003. Viral load as a predictor of the risk of cervical intraepithelial neoplasia. Int J Cancer **103**:519-524.

4. **Ylitalo, N., P. Sorensen, A. M. Josefsson, P. K. E. Magnusson, P. K. Anderson, J. Ponten, H. O. Adami, U. B. Gyllensten, and M. Melbye**. 2000. Consistent high viral load of human papillomavirus 16 and risk of cervical carcinoma in situ: a nested case-control study. Lancet **355**:2194-2198.

5. **van Duin, M., P. J. F. Snidjers, H. F. J. Schrijnemakers, F. J. Voorhorst, L. Rozendaal, M. A. E. Nobbenhuis, A. J. C. van den Brule, R. H. M. Verheijen, T. J. Helmerhorst, and C. J. L. M. Meijer**. 2002. Human papillomavirus 16 load in normal and abnormal cervical scrapes: an indicator of CIN II/III and viral clearance. Int J Cancer **98**:590-595.

6. **Gravitt, P. E., M. B. Kovacic, R. Herrero, M. Schiffman, C. Bratti, A. Hildesheim, J. Morales, M. Alfaro, M. E. Sherman, S. Wacholder, A. C. Rodriguez, and R. D. Burk**. 2007. High load for most high risk human papillomavirus genotypes is associated with prevalent cervical cancer precursors but only HPV16 load predicts the development of incident disease. Int J Cancer **121**:2787-2793.

7. **Ho, G. Y. F., R. D. Burk, S. Klein, A. S. Kadish, C. J. Chang, P. Palan, J. Basu, R. Tachezy, R. Lewis, and S. Romney**. 1995. Persistant genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. J Natl Cancer Inst **87**:1365-1371.

8. **Dalstein, V., D. Riethmuller, J. L. Pretet, C. Le Bail, J. L. Sautiere, J. P. Carbillet, B. Kantelip, J. P. Schaal, and C. Mougin**. 2003. Persistence and load of high-risk HPV are predictors for development of high-grade cervical lesions: a longitudinal French cohort study. Int J Cancer **106**:396-403.

9. **Ho, G. Y., R. Bierman, L. Beardsley, C. J. Chang, and R. D. Burk**. 1998. Natural history of cervicovaginal papillomavirus infection in young women. New Engl J Med **338**:423-428.

10. **Fontaine, J., C. Hankins, D. Money, A. Rachlis, K. Pourreaux, A. Ferenczy, and F. Coutlee**. 2008. Human papillomavirus type 16 (HPV-16) viral load and persistence of HPV-16 infection in women infected or at risk for HIV. J.Clin.Virol. **43**:307-312.

11. **Kulmala, S. M., I. P. Shabalova, N. Petrovitchev, K. J. Syrjanen, U. B. Gyllensten, B. C. Johansson, and S. M. Syrjanen**. 2007. Type-specific persistence of high-risk human papillomavirus infections in the New Independent States of the former Soviet Union Cohort Study. Cancer.Epidemiol.Biomarkers.Prev. **16**:17-22.

12. **Trottier, H., S. Mahmud, J. C. Prado, J. S. Sobrinho, M. C. Costa, T. E. Rohan, L. L. Villa, and E. L. Franco**. 2008. Type-specific duration of human papillomavirus infection: implications for human papillomavirus screening and vaccination. J.Infect.Dis. **197**:1436-1447.

13. **Cuzick, J.** 1997. Viral load as surrogate for persistence in cervical human papillomavirus infection, *In* E. L. F. Franco and J. Monsonego (eds.), Development in Cervical Cancer Screening and prevention.

14. **Pett, M. and N. Coleman**. 2007. Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? J Pathol **212**:356-367.

15. **kulmala, S.-M. A., S. M. Syrjanen, U. B. Gyllensten, I. P. Shabalova, N. Petrovichev, P. T. Tosi, K. J. Syrjanen, and B. C. Johansson**. 2007. Early integration of high copy HPV-16 detectable in women with normal and low grade cervical cytology and histology. J Clin Pathol **59**:513-517.

16. **Peitsaro, P., B. Johansson, and S. Syrjanen**. 2002. Integrated human papillomavirus type 16 is frequently found in cervical cancer precursors as demonstrated by a novel quantitative real-time PCR technique. J Clin Microbiol **40**:886-891.

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

18. **Fontaine, J., C. Hankins, M. H. Mayrand, J. Lefevre, D. Money, S. Gagnon, A. Rachlis, K. Pourreaux, A. Ferenczy, and F. Coutlee**. 2005. High levels of episomal and integrated HPV-16 DNA are associated with high-grade cervical lesions in women at risk or infected with HIV. AIDS **19**:785-794.

19. **Cheung, J. L., K. W. Lo, T. H. Cheung, J. W. Tang, and P. K. Chan**. 2006. Viral load, E2 gene disruption status, and lineage of human papillomavirus type 16 infection in cervical neoplasia. J Infect Dis **194**:1706-1712.

20. **Briolat, J., V. Dalstein, M. Saunier, K. Joseph, S. Caudroy, J. L. Pretet, P. Birembaut, and C. Clavel**. 2007. HPV prevalence, viral load and physical state of HPV-16 in cervical smears of patients with different grades of CIN. Int J Cancer **121**:2198-2204.

21. **Matovina, M., I. Sabol, G. Grubisic, N. M. Gasperov, and M. Grce**. 2009. Identification of human papillomavirus type 16 integration sites in high-grade precancerous cervical lesions. Gynecol.Oncol. **113**:120-127.

22. **Richardson, H., G. Kelsall, P. Tellier, H. Voyer, M. Abrahamowicz, F. Coutlée, and E. Franco**. 2003. The natural history of type-specific HPV infections in female university students. Cancer Epidemiol Biom Prev **12**:485-490.

23. **Richardson, H., M. Abrahamowicz, P. P. Tellier, G. Kelsall, R. du-Berger, A. Ferenczy, F. Coutlee, and E. L. Franco**. 2005. Modifiable risk factors associated with clearance of type-specific cervical human papillomavirus infections in a cohort of university students. Cancer Epidemiol.Biomarkers Prev. **14**:1149-1156.

24. **Lefevre, J., C. Hankins, K. Pourreaux, The Canadian Women's HIV study Group, and F. Coutlée**. 2004. Prevalence of selective inhibition of HPV-16 DNA amplification in cervicovaginal lavages. J Med Virol **72**:132-137.

25. **Lefevre, J., C. Hankins, K. Pourreaux, H. Voyer, The Canadian Women's HIV study Group, and F. Coutlée**. 2003. Real -time PCR assays using internal Controls for quantitation of HPV-16 and β-globin DNA in cervicovaginal lavages. J Virol Methods **144**:135-144.

26. **Gravitt, P. E., C. Peyton, C. Wheeler, R. Apple, R. Higuchi, and K. V. Shah**. 2003. Reproducibility of HPV 16 and HPV 18 viral load quantitation using TaqMan real-time PCR assays. J Virol Methods **112**:23-33.

27. **Weissenborn, S. J., A. M. Funke, M. Hellmich, P. Mallmann, P. G. Fuschs, H. J. Pfister, and U. Wieland**. 2003. Oncogenic Humn papillomavirus DNA loads in human immunodeficiency virus-positive women with High-grade cervical lesions are strongly elevated. J Clin Microbiol **41**:2763-2767.

28. **Azizi, N., J. Brazete, C. Hankins, D. Money, J. Fontaine, A. Koushik, A. Rachlis, K. Pourreaux, A. Ferenczy, E. Franco, and F. Coutlee**. 2008. Influence of HPV-16 E2 polymorphism on quantitation of HPV-16 episomal and integrated DNA in cervicovaginal lavages from women with cervical intraepithelial neoplasia. J Gen Virol **89**:1716-1728.

29. **Khouadri, S., L. L. Villa, S. Gagnon, A. Koushik, H. Richardson, A. Ferenczy, G. Matlashewski, M. Roger, E. L. F. Franco, and F. Coutlée**. 2007. Viral load of episomal and integrated forms of Human papillomavirus type 33 in High-grade squamous intraepithelial lesions of the uterine cervix. Int J Cancer **121**:2674-2681.

30. **Thorland, E. C., S. L. Myers, D. H. Persing, G. Sarkar, R. M. McGovern, B. S. Gostout, and D. I. Smith**. 2000. Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites. Cancer Res. **60**:5916-5921.

31. **Cox, D. R.** 1972. Regression Models and Life Tables. Journal of the Royal Statistical Society Series **B 34**:187-220.

32. **Liang, K.-Y. and S. L. Zeger**. 1986. Generalized linear model in longitudinal data analysis. Biometrika **73**:13-22.

33. **Hardin, J. W. and J. M. Hilbe**. 2002. *In* Champmen and Hall (ed.), Generalized estimating equations. CRC, New York.

34. **Ruutu, M. P., S. M. Kulmala, P. Peitsaro, and S. M. Syrjanen**. 2008. The performance of the HPV16 real-time PCR integration assay. Clin.Biochem. **41**:423-428.

35. **Snidjers, P. J. F., C. J. Hogewoning, A. T. Hesselink, J. Berkhof, F. J. Voohorst, M. C. Bleeker, and C. J. L. M. Meijer**. 2006. Determination of viral load thresholds in cervical scrapes to rule out CIN3 in HPV 16, 18, 31 and 33-positive women with normal cytology. Int J Cancer **119**:1102-1107.

36. **Coutlee, F., M. H. Mayrand, M. Roger, and E. L. Franco**. 2009. Detection and typing of human papillomavirus nucleic acids in biological fluids. Public Health Genomics **12**:308-318.

37. **Poppe, W. A., P. S. Ide, M. P. Drijkoningen, J. M. Lauweryns, and F. A. Van Assche**. 1995. Tobacco smoking impairs the local immunosurveillance in the uterine cervix. An immunohistochemical study. Gynecol.Obstet.Invest **39**:34-38.

38. **Giuliano, A. R., R. L. Sedjo, D. J. Roe, R. Harri, S. Baldwi, M. R. Papenfuss, M. Abrahamsen, and P. Inserra**. 2002. Clearance of oncogenic human papillomavirus (HPV) infection: effect of smoking (United States). Cancer.Causes.Control **13**:839-846.

39. Burchell, A. N., Tellier, P. P., Hanley, J., Coutlee, F., and Franco, E. L. **HPV infection in young adult couples in new sexual relationships: the importance of partner's HPV status**. Sexually Transmitted Diseases . 2009.
Ref Type: In Press

40. **Bleeker, M. C., C. J. Hogewoning, F. J. Voorhorst, A. J. van den Brule, P. J. Snijders, T. M. Starink, J. Berkhof, and C. J. Meijer**. 2003. Condom use promotes regression of human papillomavirus-associated penile lesions in male sexual partners of women with cervical intraepithelial neoplasia. Int.J.Cancer **107**:804-810.

41. **Wang, S. S. and A. Hildesheim**. 2003. Viral and host factors in human papillomavirus persistence and progression. J Natl.Cancer Inst.Monogr. 35-40.

42. **Lehoux, M., C. M. D'Abramo, and J. Archambault**. 2009. Molecular mechanisms of human papillomavirus-induced carcinogenesis. Public.Health.Genomics **12**:268-280.

43. **Burk, R. D., Z. Chen, and K. Van Doorslaer**. 2009. Human papillomaviruses: genetic basis of carcinogenicity. Public.Health.Genomics **12**:281-290.

44. **Levi, J. E., S. Fernandes, A. F. Tateno, E. Motta, L. P. Lima, J. Eluf-Neto, and C. S. Pannuti**. 2004. Presence of multiple human papillomavirus types in cervical samples from HIV-infected women. Gyn Oncol **92**:225-231.

45. **An, H. J., N. H. Cho, S. Y. Lee, I. H. Kim, C. Lee, S. J. Kim, M. S. Mun, S. H. Kim, and J. K. Jeong**. 2003. Correlation of cervical carcinoma and precancerous lesions with human papillomavirus (HPV) genotypes detected with the HPV DNA chip microarray method. Cancer **97**:1672-1680.

46. **Hildesheim, A., M. H. Schiffman, P. E. Gravitt, A. G. Glass, C. E. Greer, T. Zhang, D. R. Scott, B. B. Rush, P. Lawler, M. E. Sherman, R. J. Kurman, and M. M. Manos**. 1994. Persistence of type-specific human papillomavirus infection among cytologically normal women. J Inf Dis **169**:235-240.

47. **Guo, M., N. Sneige, E. G. Silva, Y. J. Jan, D. E. Cogdell, E. Lin, R. Luthra, and W. Zhang**. 2007. Distribution and viral load of eight oncogenic types of human papillomavirus (HPV) and HPV 16 integration status in cervical intraepithelial neoplasia and carcinoma. Mod Pathol **20**:256-266.

48. **Brown, D. R., J. T. Bryan, H. Cramer, and K. H. Fife**. 1993. Analysis of human papillomavirus types in exophytic condylomata acuminata by hybrid capture and Southern blot techniques. J Clin Microbiol **31**:2667-2673.

49. **Cuzick, J., E. Beverley, L. Ho, G. Terry, H. Sapper, I. Mielzynska, A. Lorincz, W. K. Chan, T. Krausz, and P. Soutter**. 1999. HPV testing in primary screening of older women. Br J Cancer **81**:554-558.

50. **Cuschieri, K. S., H. A. Cubie, M. W. Whitley, G. Gilkison, M. J. Arends, C. Graham, and E. McGoogan**. 2005. Persistent high risk HPV infection associated with development of cervical neoplasia in a prospective population study. J Clin Pathol **58**:946-950.

DISCUSSION

Rationale for the evaluation of HPV16 Viral Load and Integration

The role of human papillomaviruses in the development of cervical cancer is well recognized [227], with HPV DNA detected in at least 99.7% of cervical cancers [124]. Of the 40 human papillomavirus that infect the anogenital area, approximately 15 are considered high-risk for the development of cervical cancer. HPV16, the most prevalent oncogenic type, is found in more than 50% of cervical cancers [3, 4].

Cervical cancer, however, only develops in a small fraction of women infected with oncogenic HPV types. The disease also develops several years to decades post infection. This indicates that besides high risk HPV infection, other factors contribute to the development of cervical cancer.

After infection with a high risk papillomavirus, the infection is transient and cleared, or it can give rise to a progressive cervical intraepithelial disease pattern: CIN 1 may either regress spontaneously or progress to CIN 2/3 precancerous lesions and CIN 2/3 may progress toward invasive carcinoma. A persistent infection is required for the lesions to develop and to progress to CIN2/3 or carcinoma [228, 229]. A high viral load has been associated with increased persistence of HPV infection, and with an increased risk of developing CIN2/3 or cancer [108, 228, 230].

HPV16 DNA is found to be integrated in the majority of invasive cervical cancers [231, 232]. Integrated HPV DNA has also been detected in cell

lines isolated fom cervical neoplasia and in immortalized human keratinocytes [20, 233]. Integration of HPV DNA into host cell DNA usually occurs in the E1 and/or E2 genes. The product of the E2 gene is a repressor of the E6 and E7 transforming viral genes. Integration or disruption of the E2 gene results in loss of E6/E7 negative feedback control, leading to deregulation of E6 and E7 expression. Overexpression of E6 and E7 oncoproteins might promote the development of neoplasia.

Therefore viral load and integration of viral DNA into the human genome are at present being investigated as important risk factors in the progression of cervical lesions to cervical cancer. Traditionally, several methods have been used to detect integration: multiple displacement amplification, Southern blot analysis, two dimensional gel electrophoresis, in situ hybridisation, and PCR amplification of the E1/E2 region of the virus.

Most researchers, have however, opted for the PCR technique, and have studied HPV DNA integration by either quantitative real-time PCR or qualitative PCR targeting the E1 and E2 gene, or simply the E2 gene.

Some investigators report viral integration almost exclusively in high grade lesions and invasive carcinoma [9, 209, 212]. There are, however, reports of early integration of HPV DNA in low-grade cervical lesions and in asymptomatic infections [15, 28-30], and these results have been obtained with real-time quantitative PCR or qualitative PCR.

Real-time Quantitative PCR

While real-time quantitative PCR has been recently developed [28], it has since been widely applied. The technique has the advantage of measuring both viral load and integration status. The E2 and E6 ORFs are quantified with two sets of primer/probe specific for E2 and E6 respectively. Measurement of absolute copy number of E2 and E6 genes allow for determination of HPV 16 viral load and ratio of E2/E6 determines integration status.

Using quantitative real time PCR technique, one research group found integration in over 50% of CIN I [28].

Fontaine et al[213], from our team, demonstrated for the first time, in an extensive report, and this contrary to previous reports which detected mostly episomal HPV DNA in normal or LSIL smears [209, 219], that HPV integration occurs even in normal women before the presence of HPV-induced lesions.

Early integration findings are further supported by Kulmala et al who report normal morphological Pap smears to contain episomal and integrated forms. He also reported the mixed form to be the most frequent physical state of HPV 16 in low-grade cervical lesions [15].

Evaluating with the same method, and so as to maximize detection of integration, Cheung et al applied three sets of quantitative real-time PCR directed against the amino-terminal, hinge and carboxyl terminal of the E2

gene [234]. The E2 gene was found to be disrupted in 67.9% of normal/CIN I lesions.

Detection of integrated sequences in preneoplastic lesions in several other studies has consolidated these findings [30, 213, 235-237]. Kulmala et al. found purely episomal HPV 16 DNA to be associated with normal and LSIL smears only, although normal and LSIL smears also revealed the presence of mixed forms (episomal and integrated HPV) and pure integrated forms. In agreement with these results, Huang et al observed that 83.3% of CIN I contained integrated DNA, with the mixed pattern being the most prevalent physical form of HPV 16. According to Cricca et al, the most prevalent form of HPV DNA in CIN 1 is episomal, 27.8% of samples had mixed forms and integrated forms only being undetected in CIN1. Saunier et al. also reported that 29 % of normal smears had integrated HPV 16 genomes. Quantitative PCR therefore determines the episomal frequency in CIN I to be 15.4%, according to Kulmala, 17% by Huang, 36% by Saunier and 72% by Cricca. In this study, however, we found only 1 out of 89 women to have have an E6/E2 ratio greater than 2, thus resulting in an integration frequency of 1.1%.

Qualitative PCR

Other investigators study integration by evaluating disruption of the E2 gene qualitatively. The integrity of the E2 gene is evaluated by using sets of specific primers to amplify regions of full length HPV 16 E2. PCR products are identified on a 2% agarose ethidium bromide stained gel. The E2 is considered intact when all HPV 16 E2 primers produce a band on gel electrophoresis. Otherwise, integration is concluded when there is no amplification of at least one of overlapping or nonoverlapping E2 amplicons, or when all the E2 primers are negative for amplification. This assay thus detects pure integrated forms when episomal forms are absent.

To investigate the physical status of HPV 16 in low grade squamous intraepithelial lesions, Gallo et al amplified 13 of these lesions by qualitative PCR. As the E2 gene PCR product was not detected in 7 of the 13 LSIL HPV 16 samples analysed, 50% viral integration into the cellular genome was concluded [29].

In another study [238], Li et al evaluated integration in all grades of cervical lesions by the ampification of 3 overlapping fragments of the E2 gene. Multiplex PCR followed by densitometry electrophoresis image analysis of gel discriminated between episomal and mixed forms. HPV-16 was considered integrated if one or more fragments of the three amplimers were absent. The authors report HPV 16 integration in 16.7% of CIN I.

Consistent with these findings, Collins SI et al 2009, assessed the integrity of the E2 gene with 5 sets of primers spanning overlapping regions of the

full length of E2 gene, and demonstrated that 26% of incident HPV-16 infection had a disrupted E2, leading the authors to conclude HPV16 integration [239]. Yet in another study, Cricca M et al studied integration in precancerous lesions of different grade by evaluating the combined physical status of E1 and E2 [240]. Among the low grade intraepithelial lesions, 15.4% had one of the E1 fragments absent, while all of the HSIL samples had at least one of E1 or E2 disrupted/deleted fragments. The authors concluded that these samples contained HPV-16 integrated forms without episomal forms. The frequency of HPV integration in low grade cervical neoplasia therefore ranges from 0% [9, 212] to more than 50% by either quantitative [15, 28, 237] or qualitative PCR [29]. One suggested possible explanation for the discrepancy is the sensitivity of real time PCR technique, its ability to detect low copy numbers in small intraepithelial lesions, as opposed to Southern blot or two-dimensional electrophoresis which require a large amount of DNA to detect physical status.

Another plausible explanation, however, is that variation in reported integration frequency could be attributed, at least in part, to the detection of integration in qualitative and quantitative real time PCR being defined by the lack of E2 amplification. Absence of E2 amplification could be due to causes other than HPV-16 integration.

We have recently shown in our laboratory that negative amplification of E2 can be due to point mutations within the primer binding region, primer selection in conserved regions of E1/E2 ORFs being a prerequisite for a valid HPV integration assay [241]. However, in natural infections, some

HPV 16 variants may carry nucleotide variations not yet characterized. A study determining the effect of HPV 16 intratypic variations on the quantitative real-time PCR evaluation of E2/E6 ratio, demonstrated that incorporation of degenerate bases into the primers and probe can efficiently compensate for the reduced amplification efficiency due to mismatched nucleotides. It was shown in this study that mismatches between primers and probes and their binding sites, unequally affect the amplification efficiency of E2 and/or E6 ORF, giving rise to a decreased E2/E6 ratio in the range from 0 to less than 1, thus being misinterpreted as integration [242].

Lack of amplification could also, however, result from inhibition [218], some of the E1/ E2 primer pairs (or E2/E6 in quantitative PCR), being more sensitive than others to PCR inhibiting substances, as for instance, organic solvents. While amplification of a housekeeping gene such as B-globin or B-actin determines the adequacy of the samples for PCR analysis in terms of nucleic acid quantity and sample degradation, Lefebre et al suggest amplification of internal controls to screen for the presence of inhibitors in real-time PCR assays and advise sample dilution to eliminate 90% of inhibitory activity in cervical specimens [243].

Besides reduced inter and intra assay variability, quantitative real-time PCR requires a high and comparable amplification efficiency for E2 and E6 (E7) primer pairs. Sensitivity and specificity of primers are also important parameters to be considered, some primer pairs being more sensitive to excess background DNA [244]. The sensitivity of methods to measure E6 and E2 should be similar.

In this study, the physical status of HPV-16 DNA in 220 cytological samples was analysed by real-time PCR amplification of E2 and E6. Equivalent copy numbers of E2 and E6 indicates the presence of episomal forms only, whereas a ratio of E6/E2 greater than 1.2 was interpreted as mixed forms. Utilizing these criteria, quantitative real-time PCR allowed the selection of 9 clinical samples with integrated and episomal forms.

Real time PCR amplification of the E2 and E6 regions and comparison of E2 to E6 ratio of the DNA amplification products, however, provides only indirect evidence of viral integration into the host genome [28]. Therefore, RS-PCR was performed on the nine samples with an E6/E2 ratio greater than 1.2 to confirm integration and to determine the viral human DNA junction.

Separation of RS-PCR Amplicons by Agarose Gel Electrophoresis

Nested RS-PCR was performed on the nine clinical specimens with an E6/E2 ratio greater than 1.2. DNA from HPV 16, human genomic DNA and SiHa DNA served as controls. The RS-PCR products were resolved by agarose gel electrophoresis on 1.5-2% gel. Figure 16 below shows a typical nested PCR gel.

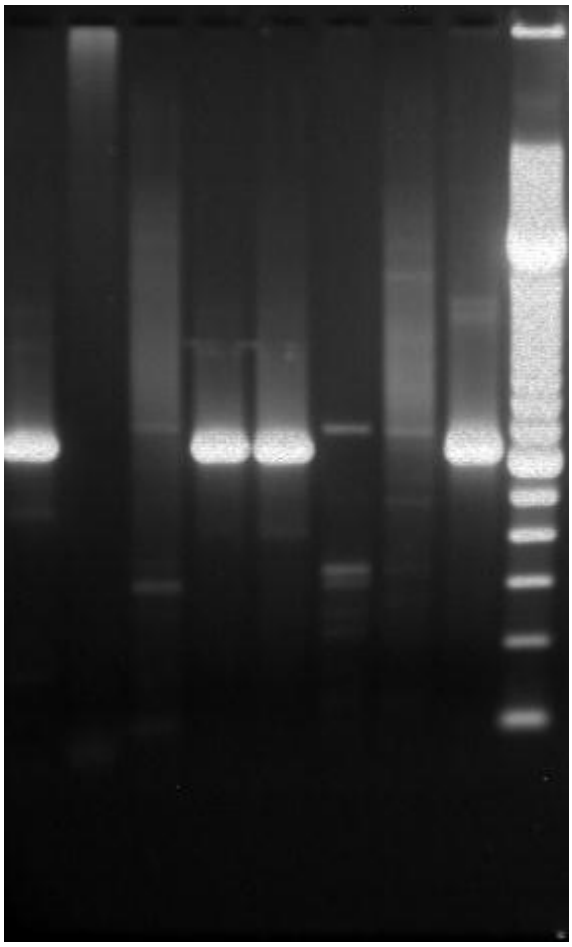


Figure 16: Representative gel electrophoresis of RS-PCR amplification products. Lane 1-6, cervical exfoliated specimens; Lane 7, human genomic DNA;
Lane 8, HPV16; Lane 9, 100 bp ladder.

As RS-PCR products are electrophored many DNA bands are resolved due to: non-specific DNA amplification, episomal DNA amplification, human-viral integrants, genomic DNA amplification. In order to minimize the number of DNA bands to be excised under UV light in the dark room the following analysis was applied to select the potential integrant amplicons.

In case of viral integration, RS-PCR amplicons include the site of viral disruption and the human genomic integration site. However, in case of episomal status, RS-PCR products display amplicon of the same size as plasmid HPV 16, the negative control for HPV integration or may display bands as in human genomic DNA lane. RS-PCR amplicons of interest, that is, not aligning with human genomic DNA (lane 7), or with pBR322 HPV16 (lane8) amplification products, are excised from the gel and sequenced to allow confirmation of integrants. Despite DNA bands being quite faint, a common occurrence, prolonged electrophoresis was nevertheless applied, in some cases, for adequate separation of the potential integrants because faint DNA amplicons could possibly represent a rare integrative event (at this stage indistinguishable from non-specific DNA amplification). This technique may thus prove to be insensitive due to the difficulty of identifying amplicons containing integrated HPV DNA.

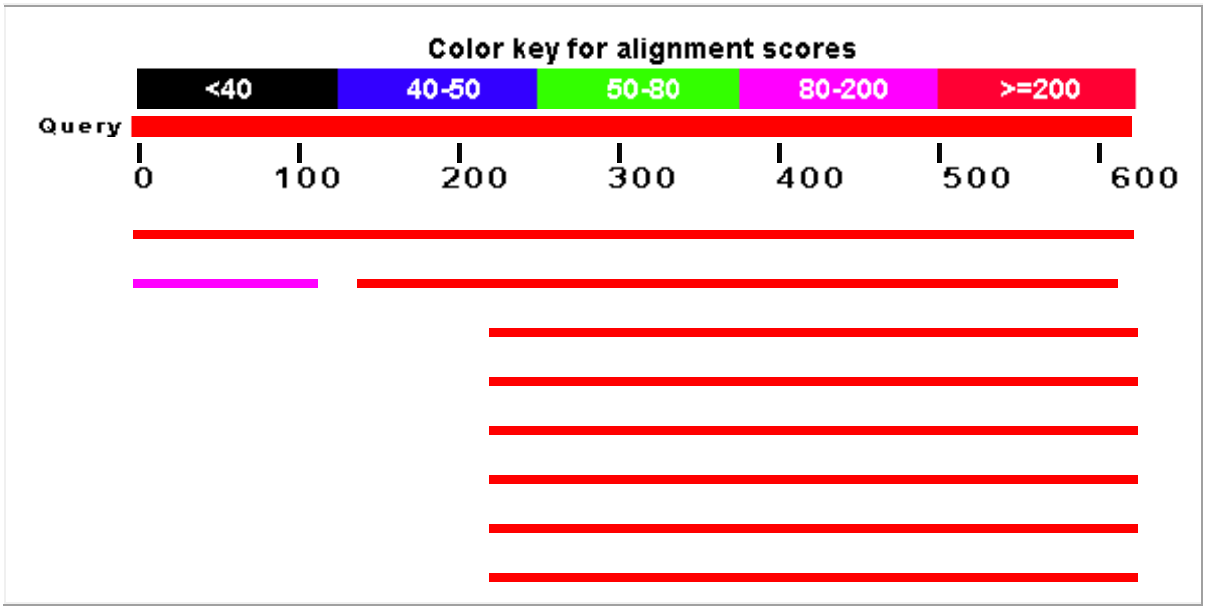
RS-PCR technique detects integration in SiHa cells

To establish the feasibility of RS-PCR in detecting integration, SiHa, an established cell line obtained from a human cervical carcinoma, served as an

integration positive control. SiHa cells harbour 1-2 copies of integrated HPV16 genome into chromosome 13 [245]. Integration in SiHa cells occurs in the E2 gene open reading frame at nucleotides 3134-3384, the E6 and E7 ORFs remain intact. Consequently, the SiHa cells serve as an excellent control of HPV 16 integration, in quantitative real time PCR and restriction site PCR.

The RS-PCR technique was performed on SiHa genomic DNA. Subsequently, amplified products were sequenced, and a search for viral host cell sequences using the National Center for Biotechnology Information BLAST database revealed HPV 16 DNA integrated into the human genome. Figure 17 below depicts nucleotide blast alignment of SiHa cell line DNA, the RS-PCR integration positive control, confirming HPV 16 DNA (red arrow) and human DNA (in yellow) identical to previously published junctional sequence [20].

SiHa 75



gb|AF001599.1|AF001599 Human papillomavirus type 16 integrated SiHa HPV-16 variant, replication protein gene, complete cds, and flanking cellular sequence
Length=3531
Score = 1166 bits (1292), Expect = 0.0
Identities = 660/668 (98%), Gaps = 3/668 (0%)
Strand=Plus/Minus

HUMAN DNA

```
Query  1     AATTCATAAAATATTCCAANNTTGTGATAAGTGTGAATCCGGTGGAATTACAATGAGAAT  60
             ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Sbjct  3370  AATTCATAAAATATTCCATAGCTGTGATAAGTGTGAATCCGGTGGAATTACAATGAGAAT  3311

Query  61    TTAGATATATA--AGTATAGTAGACATACTGGGTTATCTGAGGTGTCCTTCATTGGGATT  118
             ||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3310  TTAGATATATATAAGTATAGTAGACATACTGGGTTATCTGAGGTGTCCTTCATTGGGATT  3251

Query  119   CCCTTATCCCCACCACTAGTCACAGAAGCAGGTTCCTAACAGTCATCTTTAGACATTGTG  178
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3250  CCCTTATCCCCACCACTAGTCACAGAAGCAGGTTCCTAACAGTCATCTTTAGACATTGTG  3191
                                                                  HPV16
Query  179   ACAATGCCCAATAGGCTCCAGTCTTATCTGTCCAGGGTTAGTCAGCTAAGCTGAGCAGCA  238
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3190  ACAATGCCCAATAGGCTCCAGTCTTATCTGTCCAGGGTTAGTCAGCTAAGCTGAGCAGCA  3131

Query  239   TATGTCTCCATCAAACTGCACTTCCACTGTATATCCATGtttttttATACATCCTGTTGG  298
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3130  TATGTCTCCATCAAACTGCACTTCCACTGTATATCCATGTTTTTTTATACATCCTGTTGG  3071

Query  299   TGTAGTTAAATACACTTCAAGGCTAACGTCTTGTAATGTCCACTTTTCATTACTATATTG  358
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3070  TGTAGTTAAATACACTTCAAGGCTAACGTCTTGTAATGTCCACTTTTCATTACTATATTG  3011

Query  359   TGAGTTATATATTGTTTCTAACGTTAGTTGCAGTTCAATTGCTTGTAATGCTTTATTCTT  418
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3010  TGAGTTATATATTGTTTCTAACGTTAGTTGCAGTTCAATTGCTTGTAATGCTTTATTCTT  2951

Query  419   TGATACAGCCAGTGTTGGCACCACCTGGTGGTTAATATGTTTAAATCCCATTTCTCTGGC  478
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2950  TGATACAGCCAGTGTTGGCACCACCTGGTGGTTAATATGTTTAAATCCCATTTCTCTGGC  2891

Query  479   CTTGTAATAAATAGCACATTCTAGGCGCATGTGTTTCCAATAGTCTATATGGTCACGTAG  538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2890  CTTGTAATAAATAGCACATTCTAGGCGCATGTGTTTCCAATAGTCTATATGGTCACGTAG  2831

Query  539   GTCTGTACTATCATTTTCATAATGTGTTAGTATTTTGTCCTGACACACATTTAAACGTTG  598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2830  GTCTGTACTATCATTTTCATAATGTGTTAGTATTTTGTCCTGACACACATTTAAACGTTG  2771

Query  599   GCAAAGAGTCTCCATCGNTTTTCCTTGNCCTCGTCCTCGTGCAAACTTAATCTGGACCAC  658
             |||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||
Sbjct  2770  GCAAAGAGTCTCCATCG-TTTTCCTTGTCCTCGTCCTCGTGCAAACTTAATCTGGACCAC  2712
```

Figure 17 : Blast Nucleotide Alignment of SiHa no. 75 with integration occurring (as published) at nucleotides 3134 of E2 sequence

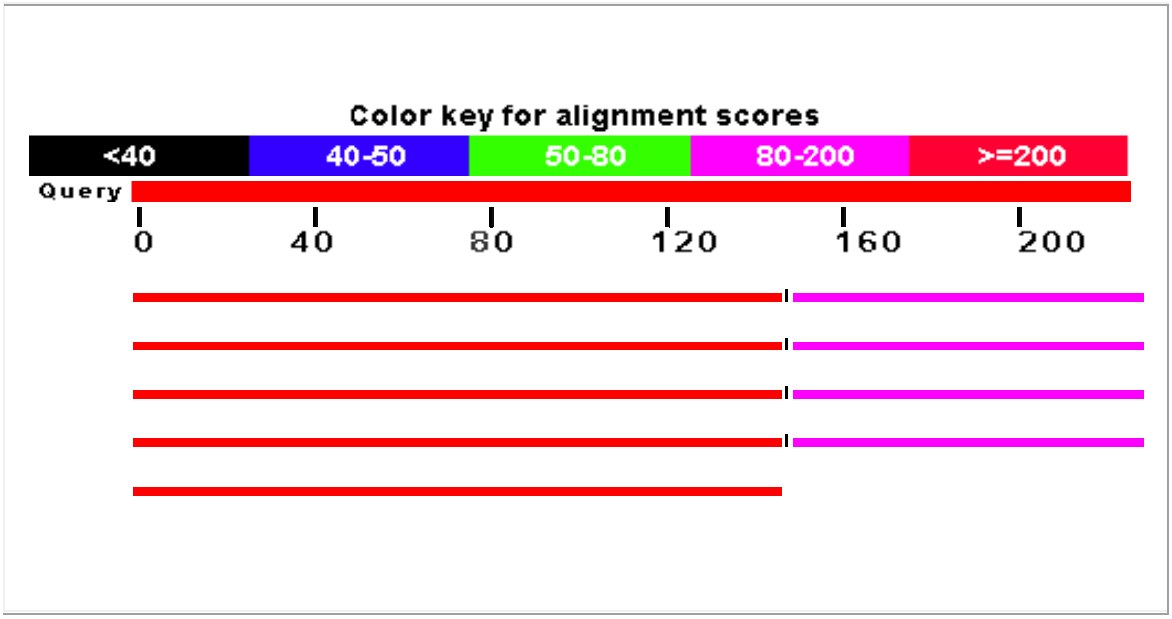
The RS-PCR technique is thus functional in the laboratory, and able to detect viral genomic sequences in the integration positive control, the SiHa cell line.

Integration negative control detects PCR artefacts

In the analysis of potential HPV integrants by nucleotide blast, suitably matched controls were also included. The HPV16 genome, cloned in pBr322, was the integration negative control throughout the procedure. The episomal DNA was particularly important to distinguish the true integrants in the clinical samples from non-specific amplification inherent to the primers/assay system employed.

Amplified RS-PCR products were sequenced in both directions with a T7 primer (complementary to the the T7 phage promoter in RSOs) and with the HPV specific nested primer used in 2nd round of RS-PCR. However, use of T7 phage promoter primer and of same nested primer as in RS-PCR reaction, rather than sequencing primers internal to nested primer, results in nonspecific blast alignments which at first analysis, promised to be potential HPV/ human junctions.

Further, by incorporating HPV16 as a negative control into RS-PCR and blast database analysis, amplification of various artefacts was detected. The characteristic pattern of the various artefacts as seen in the cervical specimen F268 is presented in figure 18A and 18B below. Nested PCR thus gave rise to unexpected HPV 16 short PCR products with internal deletions.



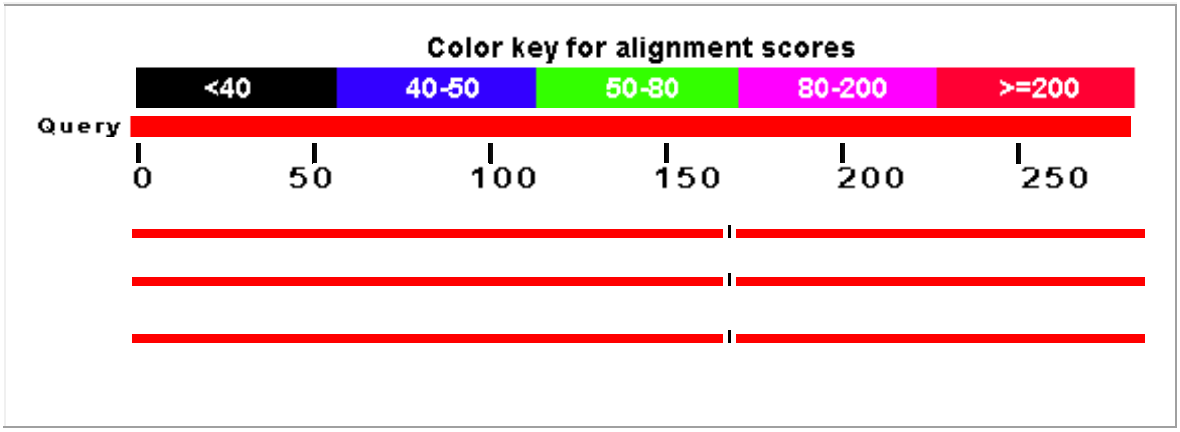
gb|FJ610151.1| Human papillomavirus type 16 strain CU6, complete genome
Length=7906
Score = 228 bits (252), Expect = 9e-57
Identities = 148/155 (95%), Gaps = 6/155 (3%)
Strand=Plus/Plus

```
Query  1     CCCCATGGTACCAATGTTGCNGTAAATACCAGGGTGAATTGTCCACCATTAGAGTTTAAT  60
             ||||||| |||||||||||| |||||| ||||| ||| ||||||||||||||||||| |||
Sbjct  6156  CCCCATG-TACCAATGTTGCAGTAAAT-CCAGG-TGA-TTGTCCACCATTAGAGTT-AAT  6210

Query  61    AAACACAGATTATTCAGGATGGTGATATGGTTGATACTGGCTTTGGTGCTATGGACTTTA  120
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  6211  AAACACAG-TTATTCAGGATGGTGATATGGTTGATACTGGCTTTGGTGCTATGGACTTTA  6269
                                                 <= L1
Query  121   CTACATTACAGGCTAACAAAAGTGAAGTTCCACTG  155
             ||||||||||||||||||||||||||||||||||
Sbjct  6270  CTACATTACAGGCTAACAAAAGTGAAGTTCCACTG  6304
```

Score = 152 bits (168), Expect = 6e-34
Identities = 84/84 (100%), Gaps = 0/84 (0%)
Strand=Plus/Minus

```
                  E2 =>
Query  156   TAACGTCTTGTAATGTCCACTTTTCATTACTATATTGTGAGTTATATATTGTTTCTAACG  215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3047  TAACGTCTTGTAATGTCCACTTTTCATTACTATATTGTGAGTTATATATTGTTTCTAACG  2988

Query  216   TTAGTTGCAGTTCAATTGCTTGTA  239
             ||||||||||||||||||||||||
Sbjct  2987  TTAGTTGCAGTTCAATTGCTTGTA  2964
```

Figure 18A: PCR Amplification of L1- E2 Deletion Artefact in specimen F268 amplified with NP4 and RSO1 primers- T7 12.2.



gb|FJ610148.1| Human papillomavirus type 16 strain CU3, complete genome

Length=7906 Score = 304 bits (336), Expect = 2e-79

Identities = 173/175 (98%), Gaps = 1/175 (0%) Strand=Plus/Plus

```
Query  1     AGACCCTGCTTTTATAACCACTCCCACTAAACTTATTACATATGATAATCCTGCATATGA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4950  AGACCCTGCTTTTATAACCACTCCCACTAAACTTATTACATATGATAATCCTGCATATGA  5009

Query  61    AGGTATAGATGTGGATAATACATTATATTTTTCTAGTAATGATAATAGTATTAATATAGC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  5010  AGGTATAGATGTGGATAATACATTATATTTTTCTAGTAATGATAATAGTATTAATATAGC  5069
                                                                  <= NP5
Query  121   TCCAGATCCTGACTTTTTGGATATAGTTGCTTTACATAGGCCGGCCATTAACCTC  175
             |||||||||||||||||||||||||||||||||||| | ||||||||||
Sbjct  5070  TCCAGATCCTGACTTTTTGGATATAGTTGCTTTACATAGGCCAG-CATTAACCTC  5123
```

Score = 215 bits (238), Expect = 8e-53 Identities = 123/124 (99%), Gaps = 1/124 (0%) Strand=Plus/Plus

```
Query  175   CAGGTAA-TCATTATTTGGTATGAGTTTAATGAAATTTCTGCAAGGGTCTGTAATATGTT  233
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2306  CAGGTAAATCATTATTTGGTATGAGTTTAATGAAATTTCTGCAAGGGTCTGTAATATGTT  2365

Query  234   TTGTAAATTCTAAAAGCCATTTTTGGTTACAACCATTAGCAGATGCCAAAATAGGTATGT  293
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2366  TTGTAAATTCTAAAAGCCATTTTTGGTTACAACCATTAGCAGATGCCAAAATAGGTATGT  2425

Query  294   TAGA  297
             ||||
Sbjct  2426  TAGA  2429
```

Figure 18 B: PCR Amplification of L1- E2 Deletion Artefact in specimen F268 amplified with NP5 and RSO3 primers- NP8

The specificity of a 297 bp sequence obtained from nested RS PCR in F268 cervical sample was confirmed by designing primers internal to the junctional sequence of the amplicon, followed by PCR amplification. A product of expected size and sequence was obtained. In addition, amplification of HPV 16, the integration negative control, was also performed with these internal primers. Unfortunately, the 297 bp sequence was also amplified in HPV 16 integration negative control, confirming the HPV 16 deletion amplicon to be a PCR artefact.

The truncated PCR products possibly result from within a difficult to amplify region, as that of AT-rich region. In some of the short PCR amplicons, the PCR primers are oriented in opposite directions within the region deleted from the nested PCR product. Both of the PCR sequences flanking the primers are correct, however the PCR product is an artefact.

Although RS-PCR is an established method to obtain unknown sequence information adjacent to a known sequence, the data presented demonstrates that caution should be exercised when analysing junction fragments obtained by PCR based data and emphasizes how important it is to utilize, in conjunction with RS-PCR, a non-PCR based strategy (such as Southern blot) to confirm the viral-cellular junction sequence.

During DNA elongation, genomes containing numerous interspersed repeat elements, or inverted repeats, as that of sequence aggt in HPV 16, may give rise to PCR artefacts.

It is noteworthy to mention that Matovina et al [246], utilizing another PCR technique to confirm integration, namely the DIPS PCR method, also reported in 10 of the 176 HPV16 positive precancerous cervical lesion samples analysed, fragments of different size than expected. In that study, the PCR amplicons were sequenced to verify integration, only to reveal HPV DNA, suggesting the presence of episomal HPV 16 only. The authors proposed presence of HPV 16 variants with internal rearrangements or gain / loss of restriction site.

Explanation of RS-PCR Integration Negative Results

In this study, the integration status of HPV-16 in normal/LSIL cervical samples was first assessed with real-time PCR-based quantification of the E2 and E6 gene ratio. This technique assumes that HPV-16 E2 and E6 genes are present in equivalent amounts within each episomal HPV genome and that a deleted or absent E2 signal indicates integrated HPV forms. Therefore as the E2 gene is deleted or absent upon HPV integration, the1:1 theoretical ratio of E6/E2 (indicating equal amounts of E6 and E2 genes), is expected to shift towards the E6 gene. Otherwise, an E6/E2 ratio greater than 1 is suggestive of the presence of integrated forms.

As it had been previously determined, however, in our laboratory that an E6/E2 ratio for the integration-negative low-risk HPV-6 type in the range from 0.4 to 2.0 was due to assay variability [247], it was concluded that integrated forms were to be considered only in samples with an E6/E2 ratio above 2.0. Given that in this study there was only 1 sample with an E6/E2

ratio above 2.0, and so as to ensure that our definition of integration did not miss certain types of integration, RS-PCR analysis were also conducted in 9 samples with an E6/E2 ratio above 1.2. Based on the RS-PCR technique, there was no evidence of integration in the low/normal grade lesions, in agreement with other studies [212, 248].

A combined approach study of E1/E2 polymerase chain reaction amplification and DIPS PCR also revealed no integration of HPV DNA in LSIL, although mixed forms were present in 1 sample of unknown status and integration was determined in only 1 ASCUS cervical specimen [246]. However, ASCUS can translate into any one of the possible histological diagnoses, ranging from a normal cervix to carcinoma.

While the RS-PCR data is negative for integration, the results do not exclude the possibility of undetected HPV integration occurring in low grade lesions and this for various reasons:

Sampling strategy

One probable explanation for this is the sampling strategy used, ie, cervical scrapes vs. biopsies. Although analyses of exfoliated cells enable an easy access to clinical material and are ideal in a diagnostic setting, the results are not representative for the overall quantity of the nucleic acids or for the diversity of viral mRNA species in the underlying lesions [118]. In other words, the exfoliated cells are not representative of cells in deeper layers of the cervix.

In a recent analysis, Steinau et al. [249] have compared primary cervical tissue with normal exfoliated cells from 7 donors to evaluate the differences between the two tissue sources. Approximately 50% of the genes present in the primary tissue could also be identified in exfoliated cells. This indicates that exfoliated cells are only a partial representation of the lesion perhaps due to the under-representation of basal cells in exfoliated tissue.

Sensitivity of RS-PCR

The exact sensitivity of RS-PCR is also unknown, although attempt was made to define it. Given that SiHa cells were utilized as integration positive control throughout the procedure, and assuming 1 copy of HPV 16 integrated in SiHa cultured cells, SiHa DNA representing 50, 10, 5, 2 and 1 copies of integrated HPV-16 DNA was RS-PCR amplified with 3 different sets of primers (3HPV specific primers and 3 RSOs). Amplification products were separated on gel, sequenced and NCBI Blast nucleotide alignment was performed. HPV-16 integration in SiHa cells was detected with at least 10 copies of HPV-16. More sensitive detection is limited at least in part by the selection and excision of faint DNA bands under ultraviolet light (Gel photograph not included).

Episomal Forms May Mask Integrants

Identification of a small number of integrated forms in a background of mainly episomal forms may be difficult. Exfoliated cells present a diversely HPV- infected cell population [250], and in low-grade lesions in particular, an abundant episomal signal from HPV-infected cells may mask the few integrated forms.

Arias-Pulido et al [251], conducted reconstitution experiments to assess the sensitivity of the quantitative real time PCR assay, and as a result concluded that the method allows distinction of integrated forms, only when these are in 100-fold excess of episomal forms. Similar results were reported suggesting the assay's lack of sensitivity [214]. More recently, Ruutu MP et al [244], have investigated the performance of quantitative real time PCR assay, so widely adopted by researchers, and have reported to be unable to detect integration when episomal forms are in 10 fold excess of integrated forms.

Given that quantitative real-time PCR was applied based on the assumption that the E6/E2 ratio would allow selection of HPV16 specimens with integrated forms, it is possible that the rare integrants, in a background of 10- 100 fold excess of episomal, were missed by quantitative real time PCR.

In other words, quantitative PCR assays are not sensitive in specimens containing mixed forms.

Primers and Probe Locations

-Integration defined by E2 Hinge Region

The quantitative real-time PCR assay assumes integration based on a decreased level of E2 amplification with respect to the E6 gene copy number. However, in order to maximize sensitivity, primers are often limited to amplify a segment less than 150 base pairs. Therefore, disruptions occurring in other areas of the 1097 base pair E2 gene will consequently be undetected by quantitative PCR.

The hinge region has been identified as the most common site of disruption [15, 17, 30, 251, 252], the E2 primers and probe locations in this study were selected to recognize E2 hinge region.

In this study, amplification of an 82 bp fragment located between nucleotides 3362-3443 of the E2 hinge region assumes episomal status. Conversely, no amplification assumes integration. Although this is the same region that is deleted in SiHa cells and in the majority of cervical cancer samples, integration could have occurred in E2 but outside the area defined by the E2 hinge region primer pairs employed, going unnoticed by quantitative real-time PCR.

An additional upstream site position 2962-3138 has also been shown to be frequently interrupted in HPV 16 positive carcinomas [15, 17]. To account for the possibility that integration might involve other parts of the E2 gene

some investigators identify integration by examining overlapping fragments within the entire E2 ORF [15].

In this study, integration occurring outside E2 hinge region might have been missed as quantitative PCR selected potential integrants based on E6/E2 ratio, with E2 delineated by 82 base pairs in the hinge region.

To circumvent this limitation, however, some studies design and validate several small contiguous primer pairs to cover the entire E2 ORF.

-Integration occurs outside the E2 ORF

The E2 open reading frame has been identified as the preferential site of integration because it has been found to be disrupted or deleted more frequently than other sites. However, integration could result in disruption outside the E2 gene, such as within the E1 as disruption usually involves E2 and E1 ORFs [17, 18, 219, 251, 253], or even at other sites, as for instance in E4, E5 and L2 ORF.

Arias-Pulido et al demonstrated deletions in both E1 and E2 ORFs in cervical carcinoma, with 78.4% observed within HPV 16 E2 hinge region [251].

Therefore it is possible that integration has been missed with these techniques, and future analysis should include the E1 ORF.

-CaSki type integration

HPV integration may not always result in disruption of the E1 /E2 ORF. Some other type of integration might occur as exemplified in the cancer cell line CaSki.

Integration in this cancer cell results in multiple copies in a head to tail tandem repeat. In this type of situation, only the viral copy flanking cellular DNA is interrupted in E1 or E2 region, this leaves the internal copies with intact E1 and E2. It is possible that preselection of potential integrants by an E6/E2 ratio of 1.2 or greater missed this type of viral integration into the human genome.

-Random Viral Integration in Human Genome

Viral integration in the human genome is random [246, 254], with a preference in common fragile sites[222].

The human junction sequences can be distributed throughout any of the common fragile sites of the whole human genome. As every integration site is unique in each specimen, the primer used to amplify the human sequences will vary from one lesion to the other. As the human primer sequences will vary, it is not possible to optimize the length of amplicon to be amplified, reaction conditions for the primer pairs, etc. It is possible that restriction site oligonucleotides (RSO's) and /or nested primers, as well as the reaction

conditions used were not optimal to amplify viral-human junctional sequences in all specimens.

In addition, the HPV16 specific primers used for the nested RS-PCR were designed based on the fact that HPV integration rarely occurs within E6 and E7 ORFs, if otherwise, however, integration may have been undetected.

-Sampling Numbers

Real time quantitative PCR directed towards E2 hinge region selected, based on E6/E2 ratio of 1.2 or greater, nine cervical specimens to be amplified by RS-PCR.

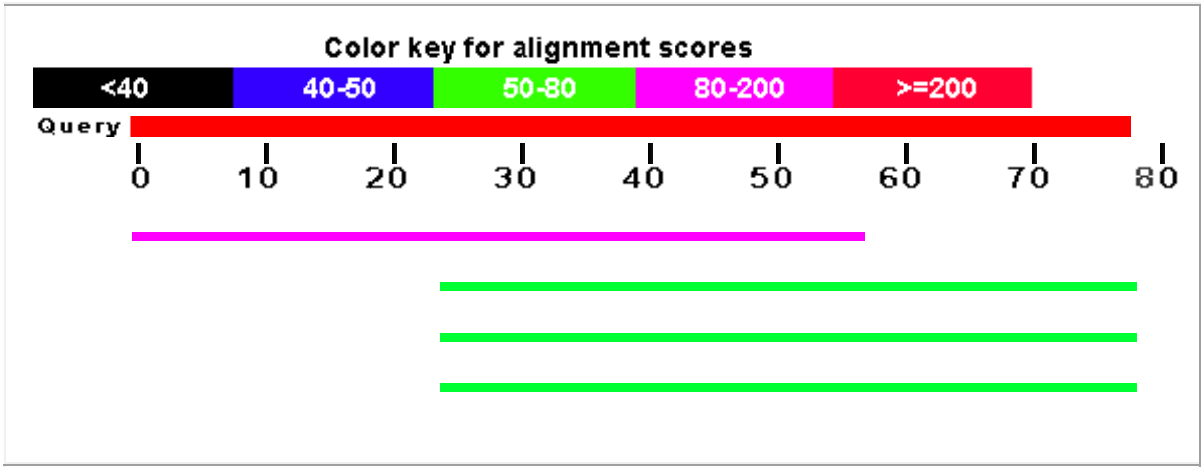
Negative integration results obtained by RS-PCR may thus be attributable in part to the small numbers of samples undergoing RS-PCR. Integration cannot be excluded until a larger set of normal and LSIL samples is explored.

-Interpretation of RS-PCR Results

The RS-PCR method is laborious, time consuming, and the results are not always clear. Unlike the DIPS PCR technique where two enzymes with a known HPV genome restriction enzyme cutting position allow for selection of PCR amplicons deviating from expected size as the potential integrants,

with RS-PCR it is impossible to determine the size of amplicons produced from HPV 16 episomes. Therefore apart from bands aligning with HPV16 plasmid negative control and those bands non-specifically human DNA generated, all other bands on agarose gel might signal integration, are selected and sequenced.

Interpretation of the results also remains difficult. Difficulty in reading electrophoregrams is exemplified by figure 19, the nucleotide sequence in red being common to human and HPV 16 DNA, leading one to suppose partial NP 5 primer (in yellow) homology with human DNA has resulted in this short amplification product. Otherwise, this could have been interpreted as integration in the specimen W2411.



gb|AC108024.3| Homo sapiens BAC clone RP11-17J21 from 4, complete sequence

Length=119882

Score = 80.6 bits (88), Expect = 7e-13

Identities = 55/60 (91%), Gaps = 3/60 (5%)

Strand=Plus/Minus

```
Query1         CCGANCTCTGGATATGTTGGAGGTTATATAGTCCAATACTACTTTGTTATTTAGTTGCTT  60
               ||||  |||||  ||||||||   ||||||||||||||||||||||||||||||||||
Sbjct  22431   CCGACCTCTG-ATATGTTGA--GTTATATAGTCCAATACTACTTTGTTATTTAGTTGCTT  22375
```

gb|FJ610152.1| Human papillomavirus type 16 strain CU7, complete genome

Length=7905

Score = 59.0 bits (64), Expect = 2e-06

Identities = 52/60 (86%), Gaps = 4/60 (6%)

Strand=Plus/Plus

```
Query  26     ATATAG-TCCA-ATACT-ACTTTGTT-ATTTAGTTGCTTTACATAGGCCAGCATTAACCT  81
              ||||||  ||||  ||  ||  |||||  |   ||  ||||||||||||||||||||||||||
Sbjct  5062   ATATAGCTCCAGATCCTGACTTTTTGGATATAGTTGCTTTACATAGGCCAGCATTAACCT  5121

Query  82     C   82  ⇦ NP5 primer
              |
Sbjct  5122   C   5122
```

Figure 19: Short sequence homologous to HPV 16 L1 ORF and to Human DNA, depicting difficulty in interpretation of NCBI BLASTresults

Recent studies suggest that viral load is associated with persistence of infection [228, 255]. High viral loads in cervical intraepithelial neoplasia lesions have been associated with a 60-fold increase in the risk of malignant progression [108]. Other investigators have demonstrated an association with increasing viral load of HPV16 and increasing severity of cervical lesions [11, 256, 257]. In contrast, some researchers have found viral loads to decrease with increasing abnormality [14] and increasing again in cancer[15].

A total of 220 HPV16 positive normal/ LSIL cervical samples were quantified for E6 and E2 DNA by quantitative real-time PCR assays. Forty eight specimens however gave a very low reading for E2 and/or E6 indicating a low HPV concentration in samples. In an attempt to better quantify and detect any possible difference between E2 and E6 ORFs perhaps undefined at near 0 values, the DNA of these specimens was concentrated by master pure. Although E6/E2 ratio was slightly altered, concentrating the DNA in 5 or 10 times its initial volume did not indicate integration.

In our specimens, high variability was found in the amount of HPV viral load, in agreement with data reported by other authors [28, 204, 212, 258]. Women younger than 25 years of age have more elevated HPV loads, also in concordance with the literature. The mean HPV 16 viral load was 57.5±324.6 copies per cell.

Some samples have a significantly low mean copy number and an E2 gene value greater than E6. Using the same technique, an E2/E6 ratio above 1.5 has been reported in six cervical lesions [259], leading the authors to define unusual integration forms. The elevated E2/E6 ratio could result from E2 duplication, similar to the CaSki cell line, in which the virus integrates in a head-to- tail tandem repeat [20], or from integration of dimeric viral DNA with an E2 portion at both ends[260], or to variability of viral quantitation resulting in uninterpretable ratios from 0.5 to 2.0.

Another plausible explanation for a relatively increased E2 to E6 gene copy number (given that E2 and E6 standard curves had comparable efficiencies

of amplification), is nucleotide alterations in the region targeted by the prototypic HPV 16 primers-probe set [261]. Quantification of E2 and E6 gene copy numbers in specimens relies on a standard curve generated by a primers-probe set complementary to the prototypic HPV 16. The primer–probe set is chosen in a conserved region of prototype HPV 16 to account for intratypic diversity, however in natural infections, HPV 16 variants may carry more nucleotide alterations than those identified. Primer-probe mismatch to the target region reduces the real-time PCR amplification efficiency, and impacts E6 differently from E2, thereby introducing significant error into the determination of E2/E6 copy number resulting in erroneous quantitative viral load and misinterpretation of integration. The E6 gene of all the HPV-16 positive samples has however been sequenced with PCR. The isolates did not present mutations at either primer or probe binding sites of E6 gene. It is possible, however, that sequence alterations in the E2 gene was responsible for an E2 value greater than E6 [242]. Incorporation of degenerate bases into the primers and probe compensates for the reduced amplification efficiency due to target nucleotide mismatches.

CONCLUSION

Further examination of a larger sample size, all grades of cervical disease with a normal control group, followed prospectively, and at all levels of experimental procedure the use of suitably matched controls, will allow to assess the association between viral load, integration and cervical disease status. It would be interesting to determine whether the detection of integrant derived transcripts by the APOT assay in cervical material has a more useful predictive value of cervical disease progression than integration at the DNA level.

At the moment, the variability of methods to detect the development of lesions, different HPV detection methods, different protocols, material of different background, different primers and probes, insufficient follow-up time, and results expressed in different units and with variable cut-off values for E2/E6 ratio defining integration, affect the comparability of results and thus do not allow for a consensus to be reached on viral load and viral integration.

Standardization, however, with prospective data, as from the HPV-PathogenISS study will allow for the comparison of viral load and integration results from different studies and laboratories. The SUCCEED study by the NCI, collects biological material from more than 1500 women with transient HPV infection at different grades of cervical dysplasia and cervical cancer, will also provide comparable data to assess viral load and integration at different steps in the progression to cervical cancer.

Cervical carcinomas develop from pre-invasive stages. With appropriate high risk HPV detection and a validated early marker, as for instance viral integration, for early diagnosis and treatment of cervical cancer precursors, cervical cancer will become a preventable disease in the near future.

REFERENCES

1. Bernard, H.U., *The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses.* J Clin Virol, 2005. **32 Suppl 1**: p. S1-6.
2. zur Hausen, H., *Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis.* J Natl Cancer Inst, 2000. **92**(9): p. 690-8.
3. Bosch, F.X., et al., *Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group.* J Natl Cancer Inst, 1995. **87**(11): p. 796-802.
4. Munoz, N., et al., *Epidemiologic classification of human papillomavirus types associated with cervical cancer.* N Engl J Med, 2003. **348**(6): p. 518-27.
5. Melkert, P.W., et al., *Prevalence of HPV in cytomorphologically normal cervical smears, as determined by the polymerase chain reaction, is age-dependent.* Int J Cancer, 1993. **53**(6): p. 919-23.
6. Jacobs, M.V., et al., *Distribution of 37 mucosotropic HPV types in women with cytologically normal cervical smears: the age-related patterns for high-risk and low-risk types.* Int J Cancer, 2000. **87**(2): p. 221-7.
7. Moscicki, A.B., et al., *Regression of low-grade squamous intra-epithelial lesions in young women.* Lancet, 2004. **364**(9446): p. 1678-83.
8. Monnier-Benoit, S., et al., *Dynamics of HPV16 DNA load reflect the natural history of cervical HPV-associated lesions.* J Clin Virol, 2006. **35**(3): p. 270-7.
9. Klaes, R., et al., *Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes.* Cancer Res, 1999. **59**(24): p. 6132-6.
10. Cuzick, J., et al., *Type-specific human papillomavirus DNA in abnormal smears as a predictor of high-grade cervical intraepithelial neoplasia.* Br J Cancer, 1994. **69**(1): p. 167-71.
11. Ylitalo, N., et al., *Consistent high viral load of human papillomavirus 16 and risk of cervical carcinoma in situ: a nested case-control study.* Lancet, 2000. **355**(9222): p. 2194-8.
12. Carcopino, X., et al., *Determination of HPV type 16 and 18 viral load in cervical smears of women referred to colposcopy.* J Med Virol, 2006. **78**(8): p. 1131-40.
13. Abba, M.C., et al., *Association of human papillomavirus viral load with HPV16 and high-grade intraepithelial lesion.* Int J Gynecol Cancer, 2003. **13**(2): p. 154-8.
14. Flores, R., et al., *Cross-sectional analysis of oncogenic HPV viral load and cervical intraepithelial neoplasia.* Int J Cancer, 2006. **118**(5): p. 1187-93.
15. Kulmala, S.M., et al., *Early integration of high copy HPV16 detectable in women with normal and low grade cervical cytology and histology.* J Clin Pathol, 2006. **59**(5): p. 513-7.
16. Jeon, S., B.L. Allen-Hoffmann, and P.F. Lambert, *Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells.* J Virol, 1995. **69**(5): p. 2989-97.

17. Kalantari, M., et al., *Disruption of the E1 and E2 reading frames of HPV 16 in cervical carcinoma is associated with poor prognosis.* Int J Gynecol Pathol, 1998. **17**(2): p. 146-53.
18. Romanczuk, H. and P.M. Howley, *Disruption of either the E1 or the E2 regulatory gene of human papillomavirus type 16 increases viral immortalization capacity.* Proc Natl Acad Sci U S A, 1992. **89**(7): p. 3159-63.
19. Choo, K.B., C.C. Pan, and S.H. Han, *Integration of human papillomavirus type 16 into cellular DNA of cervical carcinoma: preferential deletion of the E2 gene and invariable retention of the long control region and the E6/E7 open reading frames.* Virology, 1987. **161**(1): p. 259-61.
20. Baker, C.C., et al., *Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines.* J Virol, 1987. **61**(4): p. 962-71.
21. Jeon, S. and P.F. Lambert, *Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis.* Proc Natl Acad Sci U S A, 1995. **92**(5): p. 1654-8.
22. Boshart, M., et al., *A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer.* EMBO J, 1984. **3**(5): p. 1151-7.
23. Cullen, A.P., et al., *Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm.* J Virol, 1991. **65**(2): p. 606-12.
24. Das, B.C., et al., *Analysis by polymerase chain reaction of the physical state of human papillomavirus type 16 DNA in cervical preneoplastic and neoplastic lesions.* J Gen Virol, 1992. **73 (Pt 9)**: p. 2327-36.
25. Durst, M., et al., *The physical state of human papillomavirus type 16 DNA in benign and malignant genital tumours.* J Gen Virol, 1985. **66 (Pt 7)**: p. 1515-22.
26. Yee, C., et al., *Presence and expression of human papillomavirus sequences in human cervical carcinoma cell lines.* Am J Pathol, 1985. **119**(3): p. 361-6.
27. Shirasawa, H., et al., *Detection of human papillomavirus type 16 DNA and evidence for integration into the cell DNA in cervical dysplasia.* J Gen Virol, 1986. **67 (Pt 9)**: p. 2011-5.
28. Peitsaro, P., B. Johansson, and S. Syrjanen, *Integrated human papillomavirus type 16 is frequently found in cervical cancer precursors as demonstrated by a novel quantitative real-time PCR technique.* J Clin Microbiol, 2002. **40**(3): p. 886-91.
29. Gallo, G., et al., *Study of viral integration of HPV-16 in young patients with LSIL.* J Clin Pathol, 2003. **56**(7): p. 532-6.
30. Andersson, S., et al., *Type distribution, viral load and integration status of high-risk human papillomaviruses in pre-stages of cervical cancer (CIN).* Br J Cancer, 2005. **92**(12): p. 2195-200.
31. de Villiers, E., et al., *Classification of papillomaviruses.* Virology, 2004. **324**: p. 17-27.
32. Doorbar, J., *The papillomavirus life cycle.* J Clin Virol 2005. **32**(Suppl 1): p. S7-S15.
33. Ciuffo, G., *Imnfesto positivo con filtrato di verruca volgare.* Giorn Ital Mal Venereol, 1907. **48**: p. 12-17.

34. R Shope, E.H., *Infectious papillomatosis of rabbits: with a note on the histopathology.* J Exp Med, 1933. **58**: p. 607-624.
35. Koss, L.G. and G.R. Durfee, *Unusual patterns of squamous epithelium of the uterine cervix: cytologic and pathologic study of koilocytotic atypia.* Ann N Y Acad Sci, 1956. **63**(6): p. 1245-61.
36. zur Hausen, H., *Condylomata acuminata and human genital cancer.* Cancer Res, 1976. **36**(2 pt 2): p. 794.
37. zur Hausen, H., *Human papillomaviruses and their possible role in squamous cell carcinomas.* Curr Top Microbiol Immunol, 1977. **78**: p. 1-30.
38. zur Hausen, J., et al., *Attempts to detect virus-specific DNA in human tumors. II. Nucleic acid hybridizations with complementary RNA of human herpes group viruses.* Int J Cancer, 1974. **13**(5): p. 657-64.
39. Lowry, D., *The Papillomaviruses*, ed. e. Springer. 2007, New York.
40. Chen, E.Y., et al., *The primary structure and genetic organization of the bovine papillomavirus type 1 genome.* Nature, 1982. **299**(5883): p. 529-34.
41. Danos, O., et al., *Comparative analysis of the human type 1a and bovine type 1 papillomavirus genomes.* J Virol, 1983. **46**(2): p. 557-66.
42. Farmer, A.D., et al., *The Human Papillomavirus Database.* J Biomed Sci, 1995. **2**(2): p. 90-104.
43. de Villiers, E.M., *Human pathogenic papillomavirus types: an update.* Curr Top Microbiol Immunol, 1994. **186**: p. 1-12.
44. Chan, S.Y., et al., *Analysis of genomic sequences of 95 papillomavirus types: uniting typing, phylogeny, and taxonomy.* J Virol, 1995. **69**(5): p. 3074-83.
45. Forslund, O., et al., *Nucleotide sequence and phylogenetic classification of candidate human papilloma virus type 92.* Virology, 2003. **312**(2): p. 255-60.
46. Modis, Y., B.L. Trus, and S.C. Harrison, *Atomic model of the papillomavirus capsid.* EMBO J, 2002. **21**(18): p. 4754-62.
47. Trus, B.L., et al., *Novel structural features of bovine papillomavirus capsid revealed by a three-dimensional reconstruction to 9 A resolution.* Nat Struct Biol, 1997. **4**(5): p. 413-20.
48. Baker, C.C. and P.M. Howley, *Differential promoter utilization by the bovine papillomavirus in transformed cells and productively infected wart tissues.* EMBO J, 1987. **6**(4): p. 1027-35.
49. Egawa, K., *Do human papillomaviruses target epidermal stem cells?* Dermatology, 2003. **207**(3): p. 251-4.
50. Schmitt, A., et al., *The primary target cells of the high-risk cottontail rabbit papillomavirus colocalize with hair follicle stem cells.* J Virol, 1996. **70**(3): p. 1912-22.
51. Giroglou, T., et al., *Human papillomavirus infection requires cell surface heparan sulfate.* J Virol, 2001. **75**(3): p. 1565-70.
52. Joyce, J.G., et al., *The L1 major capsid protein of human papillomavirus type 11 recombinant virus-like particles interacts with heparin and cell-surface glycosaminoglycans on human keratinocytes.* J Biol Chem, 1999. **274**(9): p. 5810-22.
53. Culp, T.D. and N.D. Christensen, *Kinetics of in vitro adsorption and entry of papillomavirus virions.* Virology, 2004. **319**(1): p. 152-61.

54. Day, P.M., D.R. Lowy, and J.T. Schiller, *Papillomaviruses infect cells via a clathrin-dependent pathway.* Virology, 2003. **307**(1): p. 1-11.
55. Selinka, H.C., T. Giroglou, and M. Sapp, *Analysis of the infectious entry pathway of human papillomavirus type 33 pseudovirions.* Virology, 2002. **299**(2): p. 279-287.
56. Madison, K.C., *Barrier function of the skin: "la raison d'etre" of the epidermis.* J Invest Dermatol, 2003. **121**(2): p. 231-41.
57. Sherman, L., et al., *Inhibition of serum- and calcium-induced differentiation of human keratinocytes by HPV16 E6 oncoprotein: role of p53 inactivation.* Virology, 1997. **237**(2): p. 296-306.
58. Stanley, M.A., *Human papillomavirus and cervical carcinogenesis.* Best Pract Res Clin Obstet Gynaecol, 2001. **15**(5): p. 663-76.
59. Bonds, L., et al., *Immunohistochemical localization of cdc6 in squamous and glandular neoplasia of the uterine cervix.* Arch Pathol Lab Med, 2002. **126**(10): p. 1164-8.
60. Dehn, D., K.C. Torkko, and K.R. Shroyer, *Human papillomavirus testing and molecular markers of cervical dysplasia and carcinoma.* Cancer, 2007. **111**(1): p. 1-14.
61. You, J., et al., *Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes.* Cell, 2004. **117**(3): p. 349-60.
62. McPhillips, M.G., K. Ozato, and A.A. McBride, *Interaction of bovine papillomavirus E2 protein with Brd4 stabilizes its association with chromatin.* J Virol, 2005. **79**(14): p. 8920-32.
63. Hines, C.S., et al., *DNA structure and flexibility in the sequence-specific binding of papillomavirus E2 proteins.* J Mol Biol, 1998. **276**(4): p. 809-18.
64. Demeret, C., et al., *The human papillomavirus type 18 (HPV18) replication protein E1 is a transcriptional activator when interacting with HPV18 E2.* Virology, 1998. **242**(2): p. 378-86.
65. Steger, G. and S. Corbach, *Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein.* J Virol, 1997. **71**(1): p. 50-8.
66. Goodwin, E.C., et al., *Transactivation-competent bovine papillomavirus E2 protein is specifically required for efficient repression of human papillomavirus oncogene expression and for acute growth inhibition of cervical carcinoma cell lines.* J Virol, 1998. **72**(5): p. 3925-34.
67. Nishimura, A., et al., *Mechanisms of human papillomavirus E2-mediated repression of viral oncogene expression and cervical cancer cell growth inhibition.* J Virol, 2000. **74**(8): p. 3752-60.
68. Christensen, N.D., et al., *Rabbit oral papillomavirus complete genome sequence and immunity following genital infection.* Virology, 2000. **269**(2): p. 451-61.
69. Nicholls, P.K., et al., *Naturally occurring, nonregressing canine oral papillomavirus infection: host immunity, virus characterization, and experimental infection.* Virology, 1999. **265**(2): p. 365-74.
70. Dyson, N., et al., *The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product.* Science, 1989. **243**(4893): p. 934-7.

71. Helt, A.M. and D.A. Galloway, *Mechanisms by which DNA tumor virus oncoproteins target the Rb family of pocket proteins.* Carcinogenesis, 2003. **24**(2): p. 159-69.
72. Boyer, S.N., D.E. Wazer, and V. Band, *E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway.* Cancer Res, 1996. **56**(20): p. 4620-4.
73. Munger, K., et al., *Biological activities and molecular targets of the human papillomavirus E7 oncoprotein.* Oncogene, 2001. **20**(54): p. 7888-98.
74. Brehm, A., et al., *The E7 oncoprotein associates with Mi2 and histone deacetylase activity to promote cell growth.* EMBO J, 1999. **18**(9): p. 2449-58.
75. Antinore, M.J., et al., *The human papillomavirus type 16 E7 gene product interacts with and trans-activates the AP1 family of transcription factors.* EMBO J, 1996. **15**(8): p. 1950-60.
76. Funk, J.O., et al., *Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein.* Genes Dev, 1997. **11**(16): p. 2090-100.
77. Sherr, C.J., *Cancer cell cycles.* Science, 1996. **274**(5293): p. 1672-7.
78. Scheffner, M., et al., *The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53.* Cell, 1993. **75**(3): p. 495-505.
79. Thomas, M. and L. Banks, *Inhibition of Bak-induced apoptosis by HPV-18 E6.* Oncogene, 1998. **17**(23): p. 2943-54.
80. Li, B. and Q.P. Dou, *Bax degradation by the ubiquitin/proteasome-dependent pathway: involvement in tumor survival and progression.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 3850-5.
81. Hwang, E.S., T. Nottoli, and D. Dimaio, *The HPV16 E5 protein: expression, detection, and stable complex formation with transmembrane proteins in COS cells.* Virology, 1995. **211**(1): p. 227-33.
82. Fehrmann, F., D.J. Klumpp, and L.A. Laimins, *Human papillomavirus type 31 E5 protein supports cell cycle progression and activates late viral functions upon epithelial differentiation.* J Virol, 2003. **77**(5): p. 2819-31.
83. Stoler, M.H., *Human papillomaviruses and cervical neoplasia: a model for carcinogenesis.* Int J Gynecol Pathol, 2000. **19**(1): p. 16-28.
84. Zhang, B., D.F. Spandau, and A. Roman, *E5 protein of human papillomavirus type 16 protects human foreskin keratinocytes from UV B-irradiation-induced apoptosis.* J Virol, 2002. **76**(1): p. 220-31.
85. Ashrafi, G.H., et al., *E5 protein of human papillomavirus type 16 selectively downregulates surface HLA class I.* Int J Cancer, 2005. **113**(2): p. 276-83.
86. Doorbar, J., et al., *Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network.* Nature, 1991. **352**(6338): p. 824-7.
87. Wang, Q., et al., *Functional analysis of the human papillomavirus type 16 E1=E4 protein provides a mechanism for in vivo and in vitro keratin filament reorganization.* J Virol, 2004. **78**(2): p. 821-33.
88. Peh, W.L., et al., *Life cycle heterogeneity in animal models of human papillomavirus-associated disease.* J Virol, 2002. **76**(20): p. 10401-16.

89. Munger, K. and P.M. Howley, *Human papillomavirus immortalization and transformation functions.* Virus Res, 2002. **89**(2): p. 213-28.
90. Kjaer, S.K., et al., *High-risk human papillomavirus is sexually transmitted: evidence from a follow-up study of virgins starting sexual activity (intercourse).* Cancer Epidemiol Biomarkers Prev, 2001. **10**(2): p. 101-6.
91. Frega, A., et al., *Human papillomavirus in virgins and behaviour at risk.* Cancer Lett, 2003. **194**(1): p. 21-4.
92. Burchell, A.N., et al., *Modeling the sexual transmissibility of human papillomavirus infection using stochastic computer simulation and empirical data from a cohort study of young women in Montreal, Canada.* Am J Epidemiol, 2006. **163**(6): p. 534-43.
93. Rousseau, M.C., et al., *A cumulative case-control study of risk factor profiles for oncogenic and nononcogenic cervical human papillomavirus infections.* Cancer Epidemiol Biomarkers Prev, 2000. **9**(5): p. 469-76.
94. Schiffman, M. and S.K. Kjaer, *Chapter 2: Natural history of anogenital human papillomavirus infection and neoplasia.* J Natl Cancer Inst Monogr, 2003(31): p. 14-9.
95. Wheeler, C.M., et al., *Determinants of genital human papillomavirus infection among cytologically normal women attending the University of New Mexico student health center.* Sex Transm Dis, 1993. **20**(5): p. 286-9.
96. Rylander, E., et al., *The absence of vaginal human papillomavirus 16 DNA in women who have not experienced sexual intercourse.* Obstet Gynecol, 1994. **83**(5 Pt 1): p. 735-7.
97. Cates, W., Jr., *Estimates of the incidence and prevalence of sexually transmitted diseases in the United States. American Social Health Association Panel.* Sex Transm Dis, 1999. **26**(4 Suppl): p. S2-7.
98. Baseman, J.G. and L.A. Koutsky, *The epidemiology of human papillomavirus infections.* J Clin Virol, 2005. **32 Suppl 1**: p. S16-24.
99. Bauer, H.M., et al., *Genital human papillomavirus infection in female university students as determined by a PCR-based method.* JAMA, 1991. **265**(4): p. 472-7.
100. Burk, R.D., et al., *Sexual behavior and partner characteristics are the predominant risk factors for genital human papillomavirus infection in young women.* J Infect Dis, 1996. **174**(4): p. 679-89.
101. Herrero, R., et al., *Population-based study of human papillomavirus infection and cervical neoplasia in rural Costa Rica.* J Natl Cancer Inst, 2000. **92**(6): p. 464-74.
102. Schiffman, M.H., *Recent progress in defining the epidemiology of human papillomavirus infection and cervical neoplasia.* J Natl Cancer Inst, 1992. **84**(6): p. 394-8.
103. Richardson, H., et al., *The natural history of type-specific human papillomavirus infections in female university students.* Cancer Epidemiol Biomarkers Prev, 2003. **12**(6): p. 485-90.
104. Nobbenhuis, M.A., et al., *Cytological regression and clearance of high-risk human papillomavirus in women with an abnormal cervical smear.* Lancet, 2001. **358**(9295): p. 1782-3.
105. Munoz, N., et al., *Incidence, duration, and determinants of cervical human papillomavirus infection in a cohort of Colombian women with normal cytological results.* J Infect Dis, 2004. **190**(12): p. 2077-87.

106. Rozendaal, L., et al., *PCR based high risk HPV testing is superior to neural network based screening for predicting incident CIN III in women with normal cytology and borderline changes.* J Clin Pathol, 2000. **53**(8): p. 606-11.
107. Molano, M., et al., *Determinants of clearance of human papillomavirus infections in Colombian women with normal cytology: a population-based, 5-year follow-up study.* Am J Epidemiol, 2003. **158**(5): p. 486-94.
108. Josefsson, A.M., et al., *Viral load of human papilloma virus 16 as a determinant for development of cervical carcinoma in situ: a nested case-control study.* Lancet, 2000. **355**(9222): p. 2189-93.
109. Liaw, K.L., et al., *Detection of human papillomavirus DNA in cytologically normal women and subsequent cervical squamous intraepithelial lesions.* J Natl Cancer Inst, 1999. **91**(11): p. 954-60.
110. Winer, R.L., et al., *Development and duration of human papillomavirus lesions, after initial infection.* J Infect Dis, 2005. **191**(5): p. 731-8.
111. Cox, J.T., M. Schiffman, and D. Solomon, *Prospective follow-up suggests similar risk of subsequent cervical intraepithelial neoplasia grade 2 or 3 among women with cervical intraepithelial neoplasia grade 1 or negative colposcopy and directed biopsy.* Am J Obstet Gynecol, 2003. **188**(6): p. 1406-12.
112. Schiffman, M., et al., *The carcinogenicity of human papillomavirus types reflects viral evolution.* Virology, 2005. **337**(1): p. 76-84.
113. Hildesheim, A. and S.S. Wang, *Host and viral genetics and risk of cervical cancer: a review.* Virus Res, 2002. **89**(2): p. 229-40.
114. Solomon, D., et al., *The 2001 Bethesda System: terminology for reporting results of cervical cytology.* JAMA, 2002. **287**(16): p. 2114-9.
115. Lowy, D.R. and J.T. Schiller, *Prophylactic human papillomavirus vaccines.* J Clin Invest, 2006. **116**(5): p. 1167-73.
116. Peto, J., et al., *The cervical cancer epidemic that screening has prevented in the UK.* Lancet, 2004. **364**(9430): p. 249-56.
117. von Knebel Doeberitz, M., *New markers for cervical dysplasia to visualise the genomic chaos created by aberrant oncogenic papillomavirus infections.* Eur J Cancer, 2002. **38**(17): p. 2229-42.
118. Middleton, K., et al., *Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers.* J Virol, 2003. **77**(19): p. 10186-201.
119. Di Lonardo, A.D., et al., *Egg yolk antibodies against the E7 oncogenic protein of human papillomavirus type 16.* Arch Virol, 2001. **146**(1): p. 117-25.
120. Fiedler, M., et al., *High level HPV-16 E7 oncoprotein expression correlates with reduced pRb-levels in cervical biopsies.* FASEB J, 2004. **18**(10): p. 1120-2.
121. Fiedler, M., et al., *Expression of the high-risk human papillomavirus type 18 and 45 E7 oncoproteins in cervical carcinoma biopsies.* J Gen Virol, 2005. **86**(Pt 12): p. 3235-41.
122. Meisels, A. and R. Fortin, *Condylomatous lesions of the cervix and vagina. I. Cytologic patterns.* Acta Cytol, 1976. **20**(6): p. 505-9.
123. Durst, M., et al., *A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions.* Proc Natl Acad Sci U S A, 1983. **80**(12): p. 3812-5.

124. Walboomers, J.M., et al., *Human papillomavirus is a necessary cause of invasive cervical cancer worldwide.* J Pathol, 1999. **189**(1): p. 12-9.
125. Schiffman, M.H., et al., *Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia.* J Natl Cancer Inst, 1993. **85**(12): p. 958-64.
126. Bosch, F.X., et al., *The causal relation between human papillomavirus and cervical cancer.* J Clin Pathol, 2002. **55**(4): p. 244-65.
127. Parkin, D.M., et al., *Global cancer statistics, 2002.* CA Cancer J Clin, 2005. **55**(2): p. 74-108.
128. Sanders, G.D. and A.V. Taira, *Cost-effectiveness of a potential vaccine for human papillomavirus.* Emerg Infect Dis, 2003. **9**(1): p. 37-48.
129. Howley, P., *Fields Virology*. 3rd ed. Vol. 2. 1996, Philadelphia: Lippincott-Raven.
130. Altekruse, S.F., et al., *Comparison of human papillomavirus genotypes, sexual, and reproductive risk factors of cervical adenocarcinoma and squamous cell carcinoma: Northeastern United States.* Am J Obstet Gynecol, 2003. **188**(3): p. 657-63.
131. Castellsague, X., et al., *Worldwide human papillomavirus etiology of cervical adenocarcinoma and its cofactors: implications for screening and prevention.* J Natl Cancer Inst, 2006. **98**(5): p. 303-15.
132. zur Hausen, H., *Oncogenic DNA viruses.* Oncogene, 2001. **20**(54): p. 7820-3.
133. Morrison, C., et al., *Highly differentiated keratinizing squamous cell cancer of the cervix: a rare, locally aggressive tumor not associated with human papillomavirus or squamous intraepithelial lesions.* Am J Surg Pathol, 2001. **25**(10): p. 1310-5.
134. Clifford, G.M., et al., *Human papillomavirus types in invasive cervical cancer worldwide: a meta-analysis.* Br J Cancer, 2003. **88**(1): p. 63-73.
135. Smith, J.S., et al., *Human papillomavirus type distribution in invasive cervical cancer and high-grade cervical lesions: a meta-analysis update.* Int J Cancer, 2007. **121**(3): p. 621-32.
136. Schlecht, N.F., et al., *Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia.* JAMA, 2001. **286**(24): p. 3106-14.
137. Smith, J.S., et al., *Herpes simplex virus-2 as a human papillomavirus cofactor in the etiology of invasive cervical cancer.* J Natl Cancer Inst, 2002. **94**(21): p. 1604-13.
138. Dhanwada, K.R., et al., *Characterization of human keratinocytes transformed by high risk human papillomavirus types 16 or 18 and herpes simplex virus type 2.* J Gen Virol, 1993. **74 (Pt 6)**: p. 955-63.
139. Tran-Thanh, D., et al., *Herpes simplex virus type II is not a cofactor to human papillomavirus in cancer of the uterine cervix.* Am J Obstet Gynecol, 2003. **188**(1): p. 129-34.
140. Fang, L., et al., *Suppression of human papillomavirus gene expression in vitro and in vivo by herpes simplex virus type 2 infection.* Virology, 2003. **314**(1): p. 147-60.
141. Markowska, J., et al., *The role of Chlamydia trachomatis infection in cervical cancer development.* Eur J Gynaecol Oncol, 1999. **20**(2): p. 144-6.

142. Kiviat, N.B., et al., *Localization of Chlamydia trachomatis infection by direct immunofluorescence and culture in pelvic inflammatory disease.* Am J Obstet Gynecol, 1986. **154**(4): p. 865-73.
143. Castle, P.E., et al., *An association of cervical inflammation with high-grade cervical neoplasia in women infected with oncogenic human papillomavirus (HPV).* Cancer Epidemiol Biomarkers Prev, 2001. **10**(10): p. 1021-7.
144. Smith, J.S., et al., *Evidence for Chlamydia trachomatis as a human papillomavirus cofactor in the etiology of invasive cervical cancer in Brazil and the Philippines.* J Infect Dis, 2002. **185**(3): p. 324-31.
145. Mbulaiteye, S.M., et al., *Immune deficiency and risk for malignancy among persons with AIDS.* J Acquir Immune Defic Syndr, 2003. **32**(5): p. 527-33.
146. Vernon, S.D., et al., *The HIV-1 tat protein enhances E2-dependent human papillomavirus 16 transcription.* Virus Res, 1993. **27**(2): p. 133-45.
147. Yeaman, G.R., et al., *Human immunodeficiency virus receptor and coreceptor expression on human uterine epithelial cells: regulation of expression during the menstrual cycle and implications for human immunodeficiency virus infection.* Immunology, 2003. **109**(1): p. 137-46.
148. Perfettini, J.L., et al., *Effect of Chlamydia trachomatis infection and subsequent tumor necrosis factor alpha secretion on apoptosis in the murine genital tract.* Infect Immun, 2000. **68**(4): p. 2237-44.
149. Castle, P.E. and A.R. Giuliano, *Chapter 4: Genital tract infections, cervical inflammation, and antioxidant nutrients--assessing their roles as human papillomavirus cofactors.* J Natl Cancer Inst Monogr, 2003(31): p. 29-34.
150. Moodley, M., et al., *The role of steroid contraceptive hormones in the pathogenesis of invasive cervical cancer: a review.* Int J Gynecol Cancer, 2003. **13**(2): p. 103-10.
151. de Villiers, E.M., *Relationship between steroid hormone contraceptives and HPV, cervical intraepithelial neoplasia and cervical carcinoma.* Int J Cancer, 2003. **103**(6): p. 705-8.
152. Webster, K., A. Taylor, and K. Gaston, *Oestrogen and progesterone increase the levels of apoptosis induced by the human papillomavirus type 16 E2 and E7 proteins.* J Gen Virol, 2001. **82**(Pt 1): p. 201-13.
153. Riley, R.R., et al., *Dissection of human papillomavirus E6 and E7 function in transgenic mouse models of cervical carcinogenesis.* Cancer Res, 2003. **63**(16): p. 4862-71.
154. Brake, T. and P.F. Lambert, *Estrogen contributes to the onset, persistence, and malignant progression of cervical cancer in a human papillomavirus-transgenic mouse model.* Proc Natl Acad Sci U S A, 2005. **102**(7): p. 2490-5.
155. Brewer, J., H. Benghuzzi, and M. Tucci, *The role of route of estrogen adminstration on the proliferation of SiHa cells in culture.* Biomed Sci Instrum, 2005. **41**: p. 68-73.
156. Wang, J.T., et al., *[Analysis on synergistic action between estrogen, progesterone and human papillomaviruses in cervical cancer].* Zhonghua Liu Xing Bing Xue Za Zhi, 2005. **26**(5): p. 370-3.
157. Roche, J.K. and C.P. Crum, *Local immunity and the uterine cervix: implications for cancer-associated viruses.* Cancer Immunol Immunother, 1991. **33**(4): p. 203-9.

158. Hellberg, D. and U. Stendahl, *The biological role of smoking, oral contraceptive use and endogenous sexual steroid hormones in invasive squamous epithelial cervical cancer.* Anticancer Res, 2005. **25**(4): p. 3041-6.
159. Haverkos, H.W., *Viruses, chemicals and co-carcinogenesis.* Oncogene, 2004. **23**(38): p. 6492-9.
160. Barton, S.E., et al., *Effect of cigarette smoking on cervical epithelial immunity: a mechanism for neoplastic change?* Lancet, 1988. **2**(8612): p. 652-4.
161. Poppe, W.A., et al., *Cervical cotinine and macrophage-Langerhans cell density in the normal human uterine cervix.* Gynecol Obstet Invest, 1996. **41**(4): p. 253-9.
162. Melikian, A.A., et al., *Comparative response of normal and of human papillomavirus-16 immortalized human epithelial cervical cells to benzo[a]pyrene.* Oncol Rep, 1999. **6**(6): p. 1371-6.
163. Lea, J.S., et al., *Aberrant p16 methylation is a biomarker for tobacco exposure in cervical squamous cell carcinogenesis.* Am J Obstet Gynecol, 2004. **190**(3): p. 674-9.
164. Hemminki, K., C. Dong, and P. Vaittinen, *Familial risks in cervical cancer: is there a hereditary component?* Int J Cancer, 1999. **82**(6): p. 775-81.
165. Stern, P.L., et al., *Natural HPV immunity and vaccination strategies.* J Clin Virol, 2000. **19**(1-2): p. 57-66.
166. Rudlinger, R., et al., *Human papillomavirus infections in a group of renal transplant recipients.* Br J Dermatol, 1986. **115**(6): p. 681-92.
167. Bouwes Bavinck, J.N. and R.J. Berkhout, *HPV infections and immunosuppression.* Clin Dermatol, 1997. **15**(3): p. 427-37.
168. Tindle, R.W. and I.H. Frazer, *Immune response to human papillomaviruses and the prospects for human papillomavirus-specific immunisation.* Curr Top Microbiol Immunol, 1994. **186**: p. 217-53.
169. Magnusson, P.K., P. Sparen, and U.B. Gyllensten, *Genetic link to cervical tumours.* Nature, 1999. **400**(6739): p. 29-30.
170. Levine, A.J., *p53, the cellular gatekeeper for growth and division.* Cell, 1997. **88**(3): p. 323-31.
171. Matlashewski, G.J., et al., *Primary structure polymorphism at amino acid residue 72 of human p53.* Mol Cell Biol, 1987. **7**(2): p. 961-3.
172. Beckman, G., et al., *Is p53 polymorphism maintained by natural selection?* Hum Hered, 1994. **44**(5): p. 266-70.
173. Storey, A., et al., *Role of a p53 polymorphism in the development of human papillomavirus-associated cancer.* Nature, 1998. **393**(6682): p. 229-34.
174. Jee, S.H., et al., *Polymorphism p53 codon-72 and invasive cervical cancer: a meta-analysis.* Int J Gynaecol Obstet, 2004. **85**(3): p. 301-8.
175. Munoz, N., et al., *Against which human papillomavirus types shall we vaccinate and screen? The international perspective.* Int J Cancer, 2004. **111**(2): p. 278-85.
176. Clifford, G.M., et al., *Human papillomavirus genotype distribution in low-grade cervical lesions: comparison by geographic region and with cervical cancer.* Cancer Epidemiol Biomarkers Prev, 2005. **14**(5): p. 1157-64.
177. Durst, M., et al., *Molecular and cytogenetic analysis of immortalized human primary keratinocytes obtained after transfection with human papillomavirus type 16 DNA.* Oncogene, 1987. **1**(3): p. 251-6.

178. Kanda, T., S. Watanabe, and K. Yoshiike, *Immortalization of primary rat cells by human papillomavirus type 16 subgenomic DNA fragments controlled by the SV40 promoter.* Virology, 1988. **165**(1): p. 321-5.
179. Matlashewski, G., et al., *Human papillomavirus type 16 DNA cooperates with activated ras in transforming primary cells.* EMBO J, 1987. **6**(6): p. 1741-6.
180. Phelps, W.C., et al., *The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A.* Cell, 1988. **53**(4): p. 539-47.
181. Pirisi, L., et al., *Transformation of human fibroblasts and keratinocytes with human papillomavirus type 16 DNA.* J Virol, 1987. **61**(4): p. 1061-6.
182. Schlegel, R., et al., *Quantitative keratinocyte assay detects two biological activities of human papillomavirus DNA and identifies viral types associated with cervical carcinoma.* EMBO J, 1988. **7**(10): p. 3181-7.
183. Watanabe, S., T. Kanda, and K. Yoshiike, *Human papillomavirus type 16 transformation of primary human embryonic fibroblasts requires expression of open reading frames E6 and E7.* J Virol, 1989. **63**(2): p. 965-9.
184. Storey, A., et al., *Comparison of the in vitro transforming activities of human papillomavirus types.* EMBO J, 1988. **7**(6): p. 1815-20.
185. Gage, J.R., C. Meyers, and F.O. Wettstein, *The E7 proteins of the nononcogenic human papillomavirus type 6b (HPV-6b) and of the oncogenic HPV-16 differ in retinoblastoma protein binding and other properties.* J Virol, 1990. **64**(2): p. 723-30.
186. Munger, K., et al., *The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes.* J Virol, 1989. **63**(10): p. 4417-21.
187. Berezutskaya, E., et al., *Differential regulation of the pocket domains of the retinoblastoma family proteins by the HPV16 E7 oncoprotein.* Cell Growth Differ, 1997. **8**(12): p. 1277-86.
188. Duensing, S. and K. Munger, *Mechanisms of genomic instability in human cancer: insights from studies with human papillomavirus oncoproteins.* Int J Cancer, 2004. **109**(2): p. 157-62.
189. Duensing, S., et al., *Centrosome abnormalities and genomic instability by episomal expression of human papillomavirus type 16 in raft cultures of human keratinocytes.* J Virol, 2001. **75**(16): p. 7712-6.
190. Schaeffer, A.J., et al., *E6 and E7 oncoproteins induce distinct patterns of chromosomal aneuploidy in skin tumors from transgenic mice.* Cancer Res, 2004. **64**(2): p. 538-46.
191. Huibregtse, J.M., M. Scheffner, and P.M. Howley, *Localization of the E6-AP regions that direct human papillomavirus E6 binding, association with p53, and ubiquitination of associated proteins.* Mol Cell Biol, 1993. **13**(8): p. 4918-27.
192. Patel, D., et al., *The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300.* EMBO J, 1999. **18**(18): p. 5061-72.
193. Zimmermann, H., et al., *The human papillomavirus type 16 E6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator CBP/p300.* J Virol, 1999. **73**(8): p. 6209-19.
194. Zeitler, J., et al., *Domains controlling cell polarity and proliferation in the Drosophila tumor suppressor Scribble.* J Cell Biol, 2004. **167**(6): p. 1137-46.

195. Kiyono, T., et al., *Binding of high-risk human papillomavirus E6 oncoproteins to the human homologue of the Drosophila discs large tumor suppressor protein.* Proc Natl Acad Sci U S A, 1997. **94**(21): p. 11612-6.
196. Nguyen, M.M., et al., *Requirement of PDZ-containing proteins for cell cycle regulation and differentiation in the mouse lens epithelium.* Mol Cell Biol, 2003. **23**(24): p. 8970-81.
197. Klingelhutz, A.J., S.A. Foster, and J.K. McDougall, *Telomerase activation by the E6 gene product of human papillomavirus type 16.* Nature, 1996. **380**(6569): p. 79-82.
198. Kiyono, T., et al., *Both Rb/p16INK4a inactivation and telomerase activity are required to immortalize human epithelial cells.* Nature, 1998. **396**(6706): p. 84-8.
199. Koutsky, L.A., *Genital human papillomavirus*. Sexually transmitted diseases. 1999, New York: McGraw-Hill.
200. Hildesheim, A., et al., *Human papillomavirus type 16 variants and risk of cervical cancer.* J Natl Cancer Inst, 2001. **93**(4): p. 315-8.
201. Villa, L.L., et al., *Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia.* J Gen Virol, 2000. **81**(Pt 12): p. 2959-68.
202. Cuzick, J., et al., *HPV testing in primary screening of older women.* Br J Cancer, 1999. **81**(3): p. 554-8.
203. Ho, G.Y., et al., *Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia.* J Natl Cancer Inst, 1995. **87**(18): p. 1365-71.
204. van Duin, M., et al., *Human papillomavirus 16 load in normal and abnormal cervical scrapes: an indicator of CIN II/III and viral clearance.* Int J Cancer, 2002. **98**(4): p. 590-5.
205. Schlecht, N.F., et al., *Viral load as a predictor of the risk of cervical intraepithelial neoplasia.* Int J Cancer, 2003. **103**(4): p. 519-24.
206. Lorincz, A.T., et al., *Viral load of human papillomavirus and risk of CIN3 or cervical cancer.* Lancet, 2002. **360**(9328): p. 228-9.
207. Schwarz, E., et al., *Structure and transcription of human papillomavirus sequences in cervical carcinoma cells.* Nature, 1985. **314**(6006): p. 111-4.
208. Park, J.S., et al., *Physical status and expression of HPV genes in cervical cancers.* Gynecol Oncol, 1997. **65**(1): p. 121-9.
209. Tonon, S.A., et al., *Physical status of the E2 human papilloma virus 16 viral gene in cervical preneoplastic and neoplastic lesions.* J Clin Virol, 2001. **21**(2): p. 129-34.
210. Badaracco, G., et al., *HPV16 and HPV18 in genital tumors: Significantly different levels of viral integration and correlation to tumor invasiveness.* J Med Virol, 2002. **67**(4): p. 574-82.
211. Hopman, A.H., et al., *Transition of high-grade cervical intraepithelial neoplasia to micro-invasive carcinoma is characterized by integration of HPV 16/18 and numerical chromosome abnormalities.* J Pathol, 2004. **202**(1): p. 23-33.
212. Hudelist, G., et al., *Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade.* Gynecol Oncol, 2004. **92**(3): p. 873-80.

213. Fontaine, J., et al., *High levels of HPV-16 DNA are associated with high-grade cervical lesions in women at risk or infected with HIV.* AIDS, 2005. **19**(8): p. 785-94.
214. Nagao, S., et al., *Rapid and sensitive detection of physical status of human papillomavirus type 16 DNA by quantitative real-time PCR.* J Clin Microbiol, 2002. **40**(3): p. 863-7.
215. Reidy, P.M., et al., *Integration of human papillomavirus type 11 in recurrent respiratory papilloma-associated cancer.* Laryngoscope, 2004. **114**(11): p. 1906-9.
216. Tarkowski, T.A., et al., *Improved detection of viral RNA isolated from liquid-based cytology samples.* Mol Diagn, 2001. **6**(2): p. 125-30.
217. Cuschieri, K.S., et al., *Assessment of human papillomavirus mRNA detection over time in cervical specimens collected in liquid based cytology medium.* J Virol Methods, 2005. **124**(1-2): p. 211-5.
218. Lefevre, J., et al., *Prevalence of selective inhibition of HPV-16 DNA amplification in cervicovaginal lavages.* J Med Virol, 2004. **72**(1): p. 132-7.
219. Luft, F., et al., *Detection of integrated papillomavirus sequences by ligation-mediated PCR (DIPS-PCR) and molecular characterization in cervical cancer cells.* Int J Cancer, 2001. **92**(1): p. 9-17.
220. Habis, A.H., et al., *Molecular quality of exfoliated cervical cells: implications for molecular epidemiology and biomarker discovery.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(3): p. 492-6.
221. Sarkar, G., R.T. Turner, and M.E. Bolander, *Restriction-site PCR: a direct method of unknown sequence retrieval adjacent to a known locus by using universal primers.* PCR Methods Appl, 1993. **2**(4): p. 318-22.
222. Thorland, E.C., et al., *Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites.* Cancer Res, 2000. **60**(21): p. 5916-21.
223. Ferber, M.J., et al., *Integrations of the hepatitis B virus (HBV) and human papillomavirus (HPV) into the human telomerase reverse transcriptase (hTERT) gene in liver and cervical cancers.* Oncogene, 2003. **22**(24): p. 3813-20.
224. Ragin, C.C., S.C. Reshmi, and S.M. Gollin, *Mapping and analysis of HPV16 integration sites in a head and neck cancer cell line.* Int J Cancer, 2004. **110**(5): p. 701-9.
225. Pett, M.R., et al., *Acquisition of high-level chromosomal instability is associated with integration of human papillomavirus type 16 in cervical keratinocytes.* Cancer Res, 2004. **64**(4): p. 1359-68.
226. Dall, K.L., et al., *Characterization of naturally occurring HPV16 integration sites isolated from cervical keratinocytes under noncompetitive conditions.* Cancer Res, 2008. **68**(20): p. 8249-59.
227. zur Hausen, H., *Immortalization of human cells and their malignant conversion by high risk human papillomavirus genotypes.* Semin Cancer Biol, 1999. **9**(6): p. 405-11.
228. Dalstein, V., et al., *Persistence and load of high-risk HPV are predictors for development of high-grade cervical lesions: a longitudinal French cohort study.* Int J Cancer, 2003. **106**(3): p. 396-403.

229. Wallin, K.L., et al., *Type-specific persistence of human papillomavirus DNA before the development of invasive cervical cancer.* N Engl J Med, 1999. **341**(22): p. 1633-8.
230. Lillo, F.B., et al., *Determination of human papillomavirus (HPV) load and type in high-grade cervical lesions surgically resected from HIV-infected women during follow-up of HPV infection.* Clin Infect Dis, 2005. **40**(3): p. 451-7.
231. Pirami, L., V. Giache, and A. Becciolini, *Analysis of HPV16, 18, 31, and 35 DNA in pre-invasive and invasive lesions of the uterine cervix.* J Clin Pathol, 1997. **50**(7): p. 600-4.
232. Melsheimer, P., et al., *DNA aneuploidy and integration of human papillomavirus type 16 e6/e7 oncogenes in intraepithelial neoplasia and invasive squamous cell carcinoma of the cervix uteri.* Clin Cancer Res, 2004. **10**(9): p. 3059-63.
233. Smith, P.P., et al., *Viral integration and fragile sites in human papillomavirus-immortalized human keratinocyte cell lines.* Genes Chromosomes Cancer, 1992. **5**(2): p. 150-7.
234. Cheung, J.L., et al., *Viral load, E2 gene disruption status, and lineage of human papillomavirus type 16 infection in cervical neoplasia.* J Infect Dis, 2006. **194**(12): p. 1706-12.
235. Saunier, M., et al., *Analysis of human papillomavirus type 16 (HPV16) DNA load and physical state for identification of HPV16-infected women with high-grade lesions or cervical carcinoma.* J Clin Microbiol, 2008. **46**(11): p. 3678-85.
236. Briolat, J., et al., *HPV prevalence, viral load and physical state of HPV-16 in cervical smears of patients with different grades of CIN.* Int J Cancer, 2007. **121**(10): p. 2198-204.
237. Huang, L.W., S.L. Chao, and B.H. Lee, *Integration of human papillomavirus type-16 and type-18 is a very early event in cervical carcinogenesis.* J Clin Pathol, 2008. **61**(5): p. 627-31.
238. Li, W., et al., *The physical state of HPV16 infection and its clinical significance in cancer precursor lesion and cervical carcinoma.* J Cancer Res Clin Oncol, 2008. **134**(12): p. 1355-61.
239. Collins, S.I., et al., *Disruption of the E2 gene is a common and early event in the natural history of cervical human papillomavirus infection: a longitudinal cohort study.* Cancer Res, 2009. **69**(9): p. 3828-32.
240. Cricca, M., et al., *Disruption of HPV 16 E1 and E2 genes in precancerous cervical lesions.* J Virol Methods, 2009. **158**(1-2): p. 180-3.
241. Azizi, N., et al., *Influence of human papillomavirus type 16 (HPV-16) E2 polymorphism on quantification of HPV-16 episomal and integrated DNA in cervicovaginal lavages from women with cervical intraepithelial neoplasia.* J Gen Virol, 2008. **89**(Pt 7): p. 1716-28.
242. Jiang, M., et al., *Sequence variation of human papillomavirus type 16 and measurement of viral integration by quantitative PCR.* J Clin Microbiol, 2009. **47**(3): p. 521-6.
243. Lefevre, J., et al., *Real-time PCR assays using internal controls for quantitation of HPV-16 and beta-globin DNA in cervicovaginal lavages.* J Virol Methods, 2003. **114**(2): p. 135-44.
244. Ruutu, M.P., et al., *The performance of the HPV16 real-time PCR integration assay.* Clin Biochem, 2008. **41**(6): p. 423-8.

245. el Awady, M.K., et al., *Molecular analysis of integrated human papillomavirus 16 sequences in the cervical cancer cell line SiHa.* Virology, 1987. **159**(2): p. 389-98.
246. Matovina, M., et al., *Identification of human papillomavirus type 16 integration sites in high-grade precancerous cervical lesions.* Gynecol Oncol, 2009. **113**(1): p. 120-7.
247. Khouadri, S., et al., *Viral load of episomal and integrated forms of human papillomavirus type 33 in high-grade squamous intraepithelial lesions of the uterine cervix.* Int J Cancer, 2007. **121**(12): p. 2674-81.
248. Manavi, M., et al., *Human papillomavirus DNA integration and messenger RNA transcription in cervical low- and high-risk squamous intraepithelial lesions in Austrian women.* Int J Gynecol Cancer, 2008. **18**(2): p. 285-94.
249. Steinau, M., et al., *Gene expression profile of cervical tissue compared to exfoliated cells: impact on biomarker discovery.* BMC Genomics, 2005. **6**(1): p. 64.
250. Kalantari, M., et al., *Laser capture microdissection of cervical human papillomavirus infections: copy number of the virus in cancerous and normal tissue and heterogeneous DNA methylation.* Virology, 2009. **390**(2): p. 261-7.
251. Arias-Pulido, H., et al., *Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer.* J Clin Microbiol, 2006. **44**(5): p. 1755-62.
252. Lukaszuk, K., et al., *Human papillomavirus type 16 status in cervical carcinoma cell DNA assayed by multiplex PCR.* J Clin Microbiol, 2003. **41**(2): p. 608-12.
253. Melnikow, J., et al., *Natural history of cervical squamous intraepithelial lesions: a meta-analysis.* Obstet Gynecol, 1998. **92**(4 Pt 2): p. 727-35.
254. Wentzensen, N., S. Vinokurova, and M. von Knebel Doeberitz, *Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract.* Cancer Res, 2004. **64**(11): p. 3878-84.
255. Ho, G.Y., et al., *Natural history of cervicovaginal papillomavirus infection in young women.* N Engl J Med, 1998. **338**(7): p. 423-8.
256. Guo, M., et al., *Distribution and viral load of eight oncogenic types of human papillomavirus (HPV) and HPV 16 integration status in cervical intraepithelial neoplasia and carcinoma.* Mod Pathol, 2007. **20**(2): p. 256-66.
257. Cricca, M., et al., *Viral DNA load, physical status and E2/E6 ratio as markers to grade HPV16 positive women for high-grade cervical lesions.* Gynecol Oncol, 2007. **106**(3): p. 549-57.
258. Swan, D.C., et al., *Human papillomavirus (HPV) DNA copy number is dependent on grade of cervical disease and HPV type.* J Clin Microbiol, 1999. **37**(4): p. 1030-4.
259. Yoshida, T., et al., *Quantitative real-time polymerase chain reaction analysis of the type distribution, viral load, and physical status of human papillomavirus in liquid-based cytology samples from cervical lesions.* Int J Gynecol Cancer, 2008. **18**(1): p. 121-7.
260. Wagatsuma, M., K. Hashimoto, and T. Matsukura, *Analysis of integrated human papillomavirus type 16 DNA in cervical cancers: amplification of viral sequences together with cellular flanking sequences.* J Virol, 1990. **64**(2): p. 813-21.

261. Bao, C.G., et al., *[Effects of topographic correction on remote sensing estimation of forest biomass]*. Ying Yong Sheng Tai Xue Bao, 2009. **20**(11): p. 2750-6.