A Study of Human Papillomavirus (HPV) Types in Young South African Women and HPV Variants in South African Couples

By

Xolani-Kakuhle Mndende

MNDXOL002

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN

In fulfilment of the requirements for the degree

M. Sc in Medical Virology

Faculty of Health Sciences

UNIVERSITY OF CAPE TOWN

April 2013

Supervisor: Professor Anna-Lise Williamson

Co-supervisors: Dr Z. Z. A. Mbulawa and Dr T. L. Meiring

I, Xolani-Kakuhle Mndende, hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate

Date: ……

I owe my deepest gratitude to my supervisors Anna-Lise Williamson, Zizipho Ziphozakhe Anita Mbulawa and Tracy Leonora Meiring for their guidance, support and patience from the initial to the final level enabling me to develop an understanding of the subject. I would also like to thank Bruce Allan for his assistance during my lab work.

It is a pleasure to thank our collaborators from CAPRISA who made this research possible; Vivek Naranbhai, Quarraisha Abdool Karim and Ayesha BM Kharsany.

I am heartily thankful to my funders; National Research Foundation, Poliomyelitis Research Foundation and Marion Waddell.

I am indebted to many of my friends and colleagues who supported me; Vuyolwethu Seku, Andile Nofemela, Tsungai Jongwe, Kudakwashe Mhandire, Eunice Inkuum, Alltalents Murahwa, Thabo Diphoko, everyone at eTywaka kuGatyane, eDutywa, Montclair, Fort Hare University and Mowbray Baptist Church. My family; Onesimo, Siphosihle, Andile, bhuti Daluxolo, malume uZibekile, Sisonke, Nandipha, sisi Andiswa, Yandiswa, Nosis, Thabisa, Thanduxolo, makhulu uNomzi, tamkhulu uMbulelo and everyone else whom I forgot to mention. Thank you very much for your support, encouragement and love, ndibamba ngazozibini, halala!

This thesis is dedicated to my mother (Dr Nokuzola Mndende), my grandmother (Nothembile Mndende) and to my Creator. “Trust in the LORD with all your heart, and do not lean on your own understanding. In all your ways acknowledge Him, and He will direct your paths.” Proverbs 3:5-6

Contents

List of figures

List of tables

Abbreviations

ASCUS	Atypical squamous cells of undetermined significance
bp	Base pairs
CAPRISA	Centre for the AIDS programme of research in South Africa
CC	Cervical cancer
CDC	Center for Disease Control and Prevention
CIN(1-2-3)	Cervical intraepithelial neoplasia grade 1-2-3
CVL	Cervico-vaginal lavage
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dsDNA	Double stranded DNA
ERE	Estrogen responsive element
GRE	Glucocorticoid response element 2
HIV	Human immunodeficiency virus
HPLC	High performance liquid chromatography
HPV	Human papillomavirus
ICO	Institut Català d'Oncologia
HR-HPV	High-risk HPV
HSIL	High grade squamous intra-epithelial lesion
ICC	Invasive cervical carcinoma
kb	Kilo bases
LA	Linear array
LCR	Long control region
LR-HPV	Low-risk HPV
LSIL	Low-grade squamous intraepithelial lesions

NF-1	Nuclear factor 1
OC	Oral contraceptive
Oct1	Octamer-binding factor 1
ORF	Open reading frame
Pap smear	Papanicolaou smear
PCR	Polymerase chain reaction
pRb	Retinoblastoma protein
RCA	Rolling circle amplification
SHPC	Streptavidin-horseradish peroxide conjugate
SIL	Squamous intraepithelial lesions
STI	Sexually Transmitted Infection
TEF1	Transcription enhancing factor 1
TMB	3,3',5,5'-tetramethylbenzide
URR	Upstream regulatory region
VLP	Virus-like particle
WHO	World Health Organization

A Study of Human Papillomavirus (HPV) Types in Young South African Women and HPV Variants in South African Couples

Abstract

Human Papillomavirus (HPV) is the most common sexually transmitted infection. HPV infections are important in the pathogenesis of cervical cancer, the second most common cancer in women worldwide. Epidemiological data on HPV types demonstrate that the vaccine targeted HPV 16 and 18 are the most prevalent types found in cancer globally, while HPV types 6 and 11 are found in 90% of genital warts. However, there are significant geographic variations in the prevalence of less common genotypes. Data on HPV prevalence, type distribution, persistence and clearance in young South African women is limited and will provide justification for the introduction of vaccination. Therefore it is important that accurate determination of the prevalence of different HPV types in this country are known. HPV variants differ by 2% in the L1 gene and differ by up to 5% in the non-coding long control region (LCR). Therefore, the LCR is considered the most variable region in the entire HPV genome.

For this M.Sc. thesis two studies were performed:

- Genotyping of HPV from sexually-active HIV negative young women
- Analysis of HPV variants isolated from South Africa couples.

Genotyping of HPV from sexually-active HIV negative young women

There is limited information on the prevalence of HPV and the HPV types infecting young South African women and this study aimed to determine the HPV type specific prevalence in 223 young women in a longitudinal study.

Cervico-vaginal lavage (CVL) specimens were collected at baseline and quarterly visits from sexually-active HIV negative women that were ≤30 years of age, participating in a HIV seroincidence study in KwaZulu Natal, South Africa. The Roche Linear Array HPV Genotyping assay was used to determine HPV types from 434 CVL specimens. An overall total of 434 CVL pellets were genotyped, this included 223 baseline specimens, 106 second visit specimens (defined as 2-4 months after enrolment), 69 third visit specimens (5-7 months after enrolment), 27 fourth visit specimens (8-10 months after enrolment) and 9 fifth visit specimens (11-13 months). HPV prevalence was found to be: 67.26% (150/223) at baseline, 65.09% (69/106) at second visit, 60.87% (42/69) at third visit, 59.26% (16/27) at fourth visit and 66.67% (6/9) at fifth visit. HPV prevalence is high in this population compared to other studies performed in HIV negative populations of similar age.

Analysis of HPV variants isolated from South Africa couples

The objective of the second study was to determine the concordance of infection with HPV 16, 58 and 53 variants within heterosexual couples. A total of 12 HPV 16 positive couples, eight HPV 58 positive couples and 12 HPV 53 positive couples were identified from a cohort of South African heterosexually active couples. Specimens from these couples underwent HPV type-specific LCR PCR and sequencing to identify the variant present. In the 12 couples positive for HPV 16, ten were concordant (i.e. shared the same HPV 16 variant), and two were discordant. For the eight HPV 58 positive couples; seven couples shared the same variant while partners in one couple had different variants. From the 12 HPV 53 couples; ten couples were concordant and two couples were discordant. The majority of couples with HPV genotype-specific concordant infections shared the same variant of the specific HPV types.

Chapter 1: Literature Review

1. Introduction

Human papillomavirus (HPV) is considered the most common sexually transmitted infection (STI) (Castellsague, 2008; Kjaer *et al*., 2001) and is causally associated with cervical cancer (Walboomers *et al*., 1999) as well as other cancers such as those of the anus, oropharynx, penis, vagina and vulva (reviewed by Chaturvedi, 2010). It is estimated that 75% of sexually active men and women will become HPV infected in their life time (Palefsky, 2007; Da Ros and Schmitt, 2008). In this literature review, HPV molecular biology, prevalence of HPV in young women as well as HPV variants transmission between partners in a couple will be reviewed.

1.1 Global burden of cervical cancer

In 2008 cervical cancer was responsible for about 275 000 deaths worldwide and 85.5% of these deaths were from developing countries (Globocan, 2008). Compared to women from other world regions, African women have the highest cervical cancer incidence (Figure 1.1) (Globocan, 2008). This is because women in Africa are most likely not to be screened

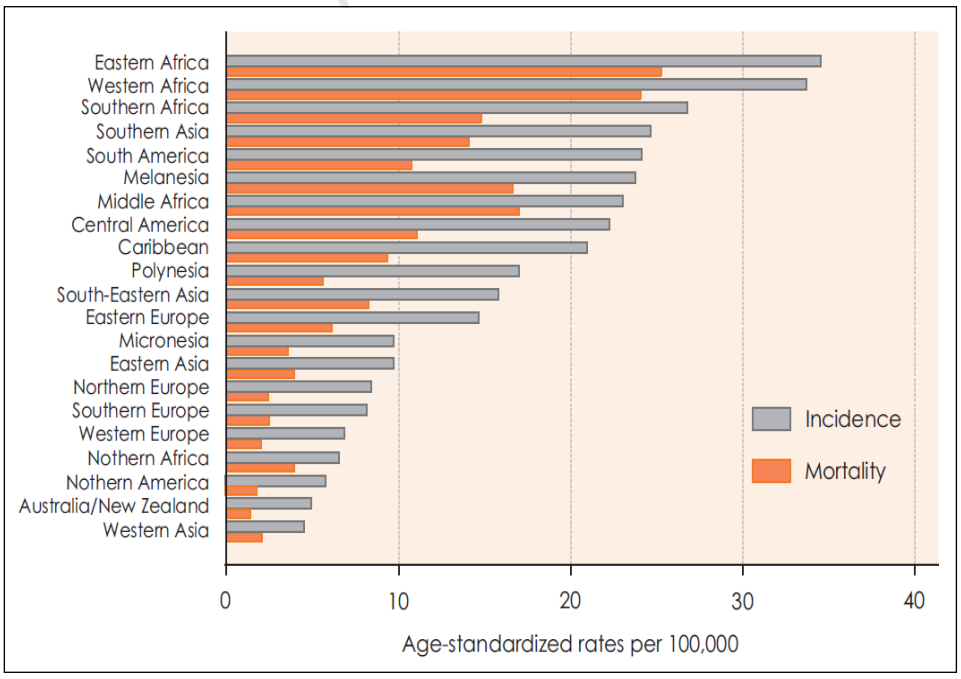


Figure 1.1 Worldwide age standardized incidence (in grey) and mortality (in orange) rates of cervical cancer per 100 000 by world region (Forman *et al*., 2012 adapted from Globocan, 2008).

for cervical cancer and have treatment for pre-cancerous lesions (de Sanjose *et al*., 2007; Kitchener *et al*., 2006). Significant reductions have been achieved in both the incidence and mortality from cervical cancer in countries with well organised screening programs (Szostek *et al*., 2008).

1.1.1 Cervical cancer in South Africa

Recent data on the prevalence of cervical cancer are limited for South African women (Denny, 2010). Cancer registry data from 1993-1995 showed that an average of 3387 new cases of cervical cancer were reported annually in that time point (Sitas *et al*., 1998). In 1998 and 1999, a total of 6061 and 5203 cases of cervical cancer were reported, respectively (Mqoqi *et al*., 2004). Fonn and co-workers (2002) determined the age specific prevalence rates of cervical cancer in 20603 South African women from across the nine provinces. A total of 92 (0.45%) women, with an average age of 51.3 years, were found to have invasive cancer (Fonn *et al*., 2002). Currently, there are about 16.84 million females (≥ 15 years) at risk for cervical cancer in South Africa (WHO/ICO, 2010). Every year, the estimated number of new cervical cancer cases diagnosed in South Africa is 5 743, with 3 027 deaths due to cervical cancer (WHO/ICO, 2010).

1.2 Other cancers caused by HPV

HPV is also potentially associated with several other cancers besides cervical cancer. Based on molecular and epidemiologic evidence, it is well known that a fairly large proportion of cancers of the oropharynx, vagina, vulva, anus and penus are etiologically related to HPV infection (Chaturvedi, 2010). Globally, HPV is associated with 100% cervical cancer cases, 90-93% of anal cancer cases, 40-64% vaginal cancer cases, 12-63% of oropharyngeal cancer cases, 40-51% cancer cases of the vulva and 36-40% of penile cancer cases (Table 1.1) (Chaturvedi, 2010).

Table 1.1 Burden of HPV related cancers (Chaturvedi, 2010).

	Cervix	Anus	Oropharynx	Penis	Vagina	Vulva
Molecular and Epidemiological evidence for a causal association with HPV	+	+	+	+	+	+
Proportion of cancers attributable to HPV infection (%)	100	90-93	12-63	36-40	40-64	40-51
Proportion of HPV-positive cancers attributable to HPV16/18 (%)	70-76	93	89-95	63-87	80-88	80-86
Annual number of cases worldwide in Men	NA	13050-13485	5100-26775	9468-10520	NA	NA
Annual number of cases worldwide in Women	492800	14310-14787	1152-6048	NA	16000-25600	
Annual number of cases worldwide in Total	492800	27360-28272	6252-32823	9468-1052	16000-25600	

(NA)= not applicable

1.3 Cofactors associated with progression to cervical cancer

High risk (HR) HPV infection is a necessary but not sufficient cause of cervical cancer; therefore, other cofactors are necessary for progression of HR-HPV infection to cervical cancer (Figure. 1.2) (Moscicki *et al*., 2006).

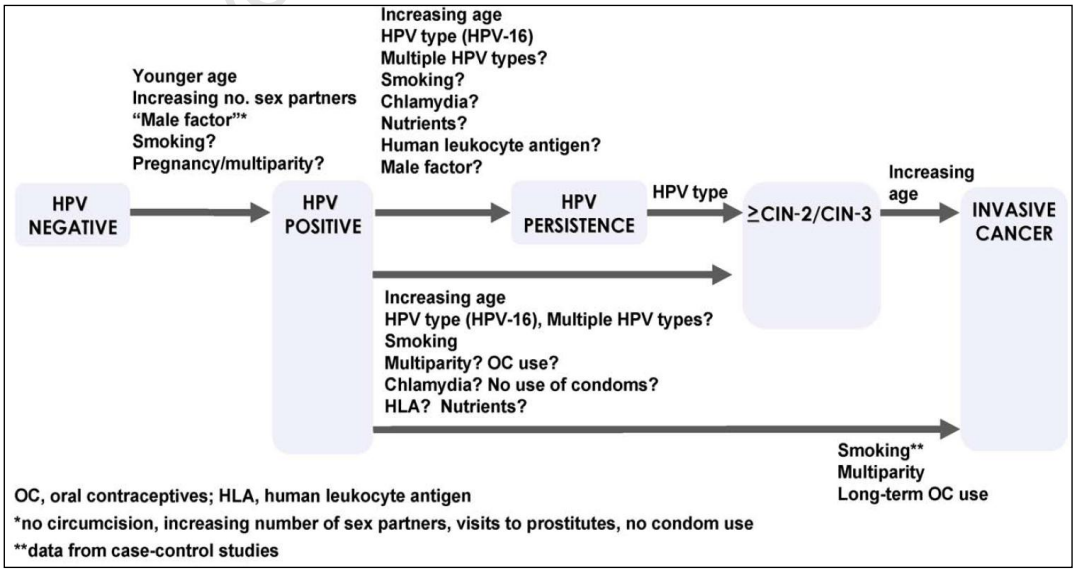


Figure 1.2 Factors associated with the natural history of HPV and progression to invasive cervical cancer (Moscicki *et al*., 2006).

Hence, the most crucial steps in cervical carcinogenesis are not only the acquisition of an HPV infection, but also the steps involving progression from a normal cervix to precancerous lesions (Moscicki *et al*., 2006).

Cervical lesions caused by HPV infection differ in severity and are cytologically classified as atypical squamous of undetermined significance (ASCUS), low-grade squamous intraepithelial lesions (LSIL) and high-grade squamous intraethelial lesions (HSIL). The LSIL is the early stage of cervical disease and is classified histologically as cervical intraepithelial neoplasia 1 (CIN1). The HSIL is the advanced stage of cervical disease and is classified histologically as cervical intraepithelial neoplasia 2/3 (Wright *et al*., 2007).

Smoking, high parity, HPV type, multiple HPV types, *Chlamydia trachomatis* (CT), micronutrient depletion, and long term use of oral contraceptives (OC) have been suggested to be important determinants in the initial stage of severe cervical lesions (Castellsague and Munoz, 2003; Deluca *et al*., 2011; Gargiulo *et al*., 2007; Gariglio *et al*., 2009; Matos *et al*., 2005; Pista *et al*., 2011; Wang *et al*., 2009; Xi *et al*., 2002). The above cofactors are thought to act by influencing the acquisition of HPV infection, increasing the risk of HPV persistence and lastly by increasing the risk of progression from HPV infection to cervical lesion eventually leading to cancer of the cervix (Castellsague and Munoz, 2003).

1.3.1 Environmental factors

- *Contraceptive use*

It has been suggested that long term use of oral contraceptive (OC) is associated with the development of cervical cancer in HPV infected individuals (Castellsague and Munoz, 2003; Moscicki *et al*., 2006). Not all studies have confirmed this, Shapiro et al (2003) reported that neither combined oestrogen/progestogen OC nor injectable progestogen increased the risk of clinically evident invasive cervical cancer (ICC) in South African women. The association between the use of OC and cervical cancer has

been thought to be due to sexual activity rather than a direct effect from OC (Xi *et al*., 2002). However a recent study investigated the direct effect of OC use and found that oral contraceptive users are 85% less likely to clear HR-HPV infections than non-users (Marks *et al*., 2011). Oestrogen/progesterone is known to increase the risk of cancer by suppressing the immune system (Marks *et al*., 2011).

Other studies infer that long term oral contraceptive use does not influence the acquisition or the persistence of HPV infection but rather enhances the probability of malignant transformation of the infection (Bertram, 2004; Vaccarella *et al*., 2006). A possible mechanism for this phenomenon is that, hormonal contraceptives may increase the expression of viral oncogenes by interacting with the estrogen responsive elements (ERE) on the long control region (LCR) (Chen *et al*., 1996a). Together the LCR with E2 are responsible for HPV transcriptional regulation. Therefore a positive interaction between estrogen and ERE on the LCR may result in the viral genes expression, replication and transformation (Lopez-Saavedra *et al*., 2009).

- *Parity*

High parity is associated with viral persistence and increased risk of cervical cancer (Castellsague and Munoz, 2003; Louie *et al*., 2009a; Moscicki *et al*., 2006). Whereas nullparous women have a lower risk of squamous cell carcinoma of the cervix compared to parous women (Munoz *et al*., 2002). The risk of developing squamous cell carcinoma of the cervix further increases among parous women with increasing number of pregnancies (Munoz *et al*., 2002). An explanation for the increased risk of cervical carcinoma among women who are multiparous because their transformation zone is maintained on the exocervix and result in direct exposure to HPV and other cofactors (Autier *et al*., 1996; Liao *et al*., 2012; Munoz *et al*., 2002).

- *Smoking as a risk factor for cervical cancer*

Smoking is one of the cofactors contributing to the progression of HPV infection to cervical cancer (Collins *et al*., 2010; Moscicki *et al*., 2006;

Schmeink *et al*., 2011). Cigarette contain carcinogens such as Benzo[*a*]pyrene, which upregulate the amplification of HPV genome which in turn may increase the chances of viral integration into the host genome causing progression to cancer of the cervix (Alam *et al*., 2008). In an analysis of HR-HPV persistence and associated risk factors in a cohort composed of young unscreened women, the number of lifetime sexual partners and smoking were associated with HR-HPV type persistence (Schmeink *et al*., 2011). Smokers were found to have a two-fold increased risk of HR-HPV persistence (Schmeink *et al*., 2011). Together cofactors such as multiple infections, multiple sexual lifetime partners and smoking were found to increase the risk of HR-HPV persistency (Schmeink *et al*., 2011). HPV 16 positive women who are current smokers are particularly at higher risk of progressing to HSIL (Wang *et al*., 2009). Current smokers were found to have a significantly increased risk of developing squamous cell cervical carcinoma compared to non-smokers (Appleby *et al*., 2006). This risk increased in relation to the number of cigarettes smoked per day and with decreased age at starting smoking. Furthermore past smokers had a lower risk of squamous cell carcinoma than current smokers (Appleby *et al*., 2006). Collins and colleagues (2010) reported that current young women who smoke and who are sexually active are twice as likely to be diagnosed with high grade lesion compared to non-smokers (Collins *et al*., 2010). However, other studies have found no difference in persistence between current smoker and current non-smokers (Maucort-Boulch *et al*., 2010).

Castellsague and Munoz (2003) concluded that smoking; long term OC use and parity are co-factors that regulate the risk of progression from HPV infection to HSIL and eventually to cervical cancer. Therefore, multiparous women that are current smokers and use OC need closer monitoring for HPV infection and cytological abnormalities then women in the general population (Castellsague and Munoz, 2003).

- *Infections with other microorganisms and increased risk of cervical cancer*

HPV is more common in HIV infected women than in HIV negative women (Denny *et al*., 2012), particularly in those women who are immunocompromised with low CD4 lymphocytes (Jamieson *et al*., 2002; Strickler *et al*., 2005). Memiah and others (2012) found a positive correlation between precancerous cervical lesions and CD4 count. Women with CD4 <200/ul were found to be 1.6 time more likely to develop precancerous cervical lesions (Memiah *et al*., 2012). In a recent South African study investigating HPV incidence and clearance in both HIV positive and negative men and women followed at six months intervals for a period of 24 months, a higher detection rate of new HPV in HIV positive men and women was reported compared to HIV negative men and women. HIV positive men and women had a significantly lower rate of HPV clearance compared to HIV negative men and women (Mbulawa *et al*., 2012). Among HIV positive women from Cape Town younger than 60 years of age, it was found that HIV positive women with ICC were on average 6 years younger than HIV negative women (Moodley *et al*., 2006). Another study investigating HIV prevalence among women with cervical cancer, mean age was 41 years from KwaZulu Natal, reported that HIV infected women were on average 13 years younger than HIV negative women (Moodley *et al*., 2005).

Chlamydia trachomatis (CT) is another well-known microorganism to play a major role in the aetiology of CIN by facilitating the infection and persistence of HR-HPV (Deluca *et al*., 2011). Association between CT and HPV is not unusual, since they share the same route of sexual transmission (Oh *et al*., 2009; Simonetti *et al*., 2009). Deluca and others (2011) demonstrated a significant association between concurrent CT and HPV infection. Oh and co-workers (2009) found a positive association with the number of lifetime sexual partners, husband's extramarital sexual relationships, and cytological abnormalities to be significant for HR-HPV types, but not for CT. According to Simonetti et al (2009) the possible association between CT and HPV infections with the incidence of ICC is due to the causal expression of pro-

inflammatory mediators, causing a change in cell to cell adhesion, and affecting cell differentiation (Simonetti *et al*., 2009). Madeleine and others (2007) speculated that the inflammatory response and metaplasia triggered by CT infections promotes cell cycle and therefore increases the number of non-dividing differentiating cells that are needed for HPV replication (Madeleine *et al*., 2007).

Herpes simplex virus type 2 (HSV-2) is another well-known microorganism to contribute to progression to CIN. It has been suggested that HSV-2 infections acting together with HPV infection may increase the risk of ICC (Smith *et al*., 2002). This is due to the fact that HSV-2 also infects cervical squamous epithelial tissues found in the squamous-columnar junction, the site where invasive cervical cancer arises (Smith *et al*., 2002).

1.4 HPV structure and genome organisation

HPVs are small non-enveloped double stranded (ds) DNA viruses almost 55nm diameter in size. They have a circular genome composed of approximately 8000 base pairs (bp) in size which is wrapped inside a protein shell (Munoz *et al*., 2006). The genome consists of three regions namely; the early (E) coding region, late (L) coding region and the long control region (LCR) (Figure 1.3) (Doorbar, 2006; Munoz *et al*., 2006).

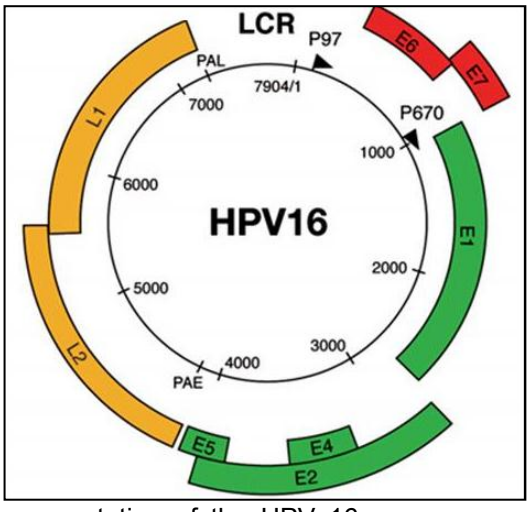


Figure 1.3 Schematic representation of the HPV 16 genome, presenting early region in green and red (E1, E2, E4 and E5 as well as the oncogenes E6 and E7). The late region structural genes are shown in orange (L1 and L2) and LCR region (Doorbar *et al*., 2012).

The LCR, which is also known as the upstream regulatory region (URR), contains the origin of replication and the P97 early promoter, but does not code for proteins (Donne *et al*., 2010; Veress *et al*., 2001; Vinokurova and von Knebel, 2011). The HPV genome encodes six early proteins termed (E1, E2, E4, E5, E6 and E7) and two late proteins which are L1 and L2 (Figure 1.3), which are essential for the replication of the viral DNA and the assembly of newly synthesed virus particles (Doorbar, 2006; Munoz *et al*., 2006). The E1 gene encodes a protein necessary for viral genomic replication (Wilson *et al*., 2002). The E2 gene product plays a role in both the transcription and the replication of HPV. Protein encoded by the E2 gene has a regulatory effect on the transcription of the transforming proteins E6/E7 (Schmidt *et al*., 2005). The E1 and E2 proteins are also involved in maintaining viral latency by minimizing the expression of viral proteins in order to evade the host cellular immune surveillance (Pittayakhajonwut and Angeletti, 2010; Schmidt *et al*., 2005; Wilson *et al*., 2002). E4 protein is present in large concentration during late HPV infection and is involved in the release of particles (Doorbar *et al*., 1986). E4 is also required to form a complex with E1 for the efficient execution of productive phase of HPV 18 (Wilson *et al*., 2007). The E5 plays a role in transformation, a study done by Maufort and others (2010) showed that the E5 protein, in conjunction with prolonged estrogen treatment, induced cervical cancer in E5 transgenic mice (Maufort *et al*., 2010). E6 and E7 proteins are expressed at low levels during the infectious process and are the primary oncoproteins (de Sanjose *et al*., 2010; Frazer *et al*., 2011; Schiffman *et al*., 2007). HPV oncogenes E6 and E7 encode proteins that interact with tumour suppressor proteins and deregulate the cell cycle's control mechanisms by creating genomic instability (Doorbar, 2006; Mazumder *et al*., 2011). Malignant progression of cervical neoplasia is associated with the expression of these oncoproteins (de Sanjose *et al*., 2010; Schiffman *et al*., 2007). E6 blocks apoptosis by inhibiting p53 and E7 inhibits retinoblastoma protein (pRB) which halts cell cycle arrest. The L1 and L2 genes encode structural proteins making up the viral capsid which are produced late in the infectious cycle of HPV and enclose the circular double stranded DNA (Lowe *et al*., 2008) (Table 1.2).

Table 1.2 Papillomavirus encoded proteins: site in infected keratinocytes and function

Protein name	Site in infected keratinocytes	Protein function
E1	Nucleus, basal cells	Necessary for viral genomic replication
E2	Nucleus, basal cells	Necessary for viral genomic replication and has regulatory effect on the transcription of the transforming proteins E6/E7
E4	Cytoplasm, spinous layer	Involved in the release of viral particles
E5	Transmembrane, spinous layer	Plays a role in transformation
E6	Nucleus, basal, spinous	Required to maintain S-phase-like state during cell cycle, blocks apoptosis by inhibiting p53
E7	Nucleus, basal, spinous	Required to maintain S-phase-like state, inhibits retinoblastoma protein (pRB) which halts cell cycle arrest
L1	Nucleus, mature squames	Major capsid protein
L2	Nucleus, mature squames	Minor capsid protein

1.5 HPV classification

The family *Papillomaviridae* contains 29 genera which are composed of 189 papillomavirus types (Figure 1.4) (Bernard *et al*., 2010).

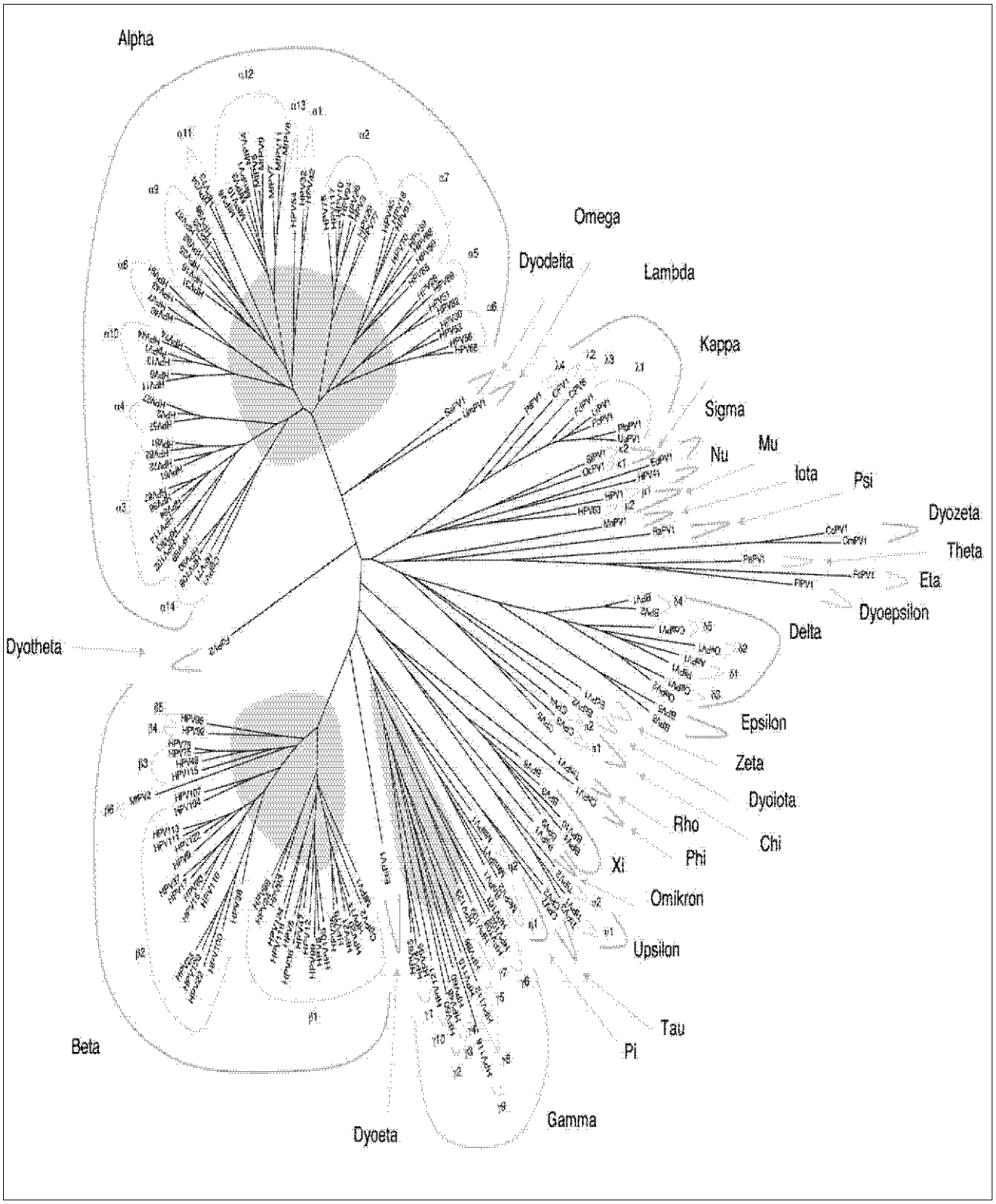


Figure 1.4 Phylogenetic tree of Papillomavirus types based on alignment of the L1 open reading frame (ORF) sequences (Bernard *et al*., 2010).

Of the 189 papillomavirus types; 120 types are from humans, 64 types are from non-human mammals, 3 types from birds and 2 types are from reptiles (Bernard *et al*., 2010). Of the 120 HPV types that have been characterized there are about 40 types which have been found in the genital tract (Bernard *et al*., 2010; Jung *et al*., 2004; Munoz *et al*., 2006).

HPV genotypes differ by 10% in DNA sequences based on the L1 open reading frame (ORF) (Bernard *et al*., 2010; de Villiers *et al*., 2004). Types that are closely related, by approximately 80-90%, are classified as members of the same species. They usually share important biological properties, such as tissue tropism and pathogenicity (Bernard *et al*., 2010; Chow *et al*., 2010). Subtypes are defined as being different by 2-10% from any HPV types (de Villiers *et al*., 2004). HPV variants differ by less than 2% in the L1 gene and by up to 5% in the LCR (de Villiers *et al*., 2004; Ho *et al*., 1993a).

HPVs can be divided into two groups (Doorbar *et al*., 2012), according to the site of infection, these are the mucosotropic group (Castellsague, 2008; Chow *et al* 2010) and the cutaneotrophic group (Castellsague, 2008; Chouhy *et al*., 2010). The mucosotropic group is found in the Alpha genera, which is the largest genus (Figure. 1.4) composed of 15 species which infects the upper aerodigestive tract, head and neck mucosa and the anogenital tract (Chow *et al*., 2010). This genus can be further subdivided into three categories depending on the frequency with which they cause cervical cancer, as either; high risk (HR), probably high risk and low risk (LR) (Table 1.3) (Doorbar, 2006; Chow *et al*., 2010; Jung *et al*., 2004).

Table 1.3 Classification of HPV types according to the frequency associated with cervical cancer (Bernard *et al*., 2010; Dobec *et al*., 2009; Munoz *et al*., 2003).

Risk associated with cervical cancer	**HPV types**
Low risk	6 ,11, 40, 42, 54, 55, 61, 62, 64, 67, 69, 70, 71, 72, 81, 83, 84, 89, and IS39
Probably high risk	26, 53 and 66
High risk	16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73 and 82

Of the HR-HPV types, HPV 16 is associated with 53.5% of cervical cancer cases worldwide; while HPV 18 is associated with 17.2% cases (Figure 1.5) (Castellsague, 2008; Clifford *et al*., 2006). Together HPVs 16, 18, 45 and 31 are responsible for 80.3% cases of global cervical cancer (Bosch *et al*., 2008; Castellsague, 2008; de Sanjose *et al*., 2010; Munoz *et al*., 2004).

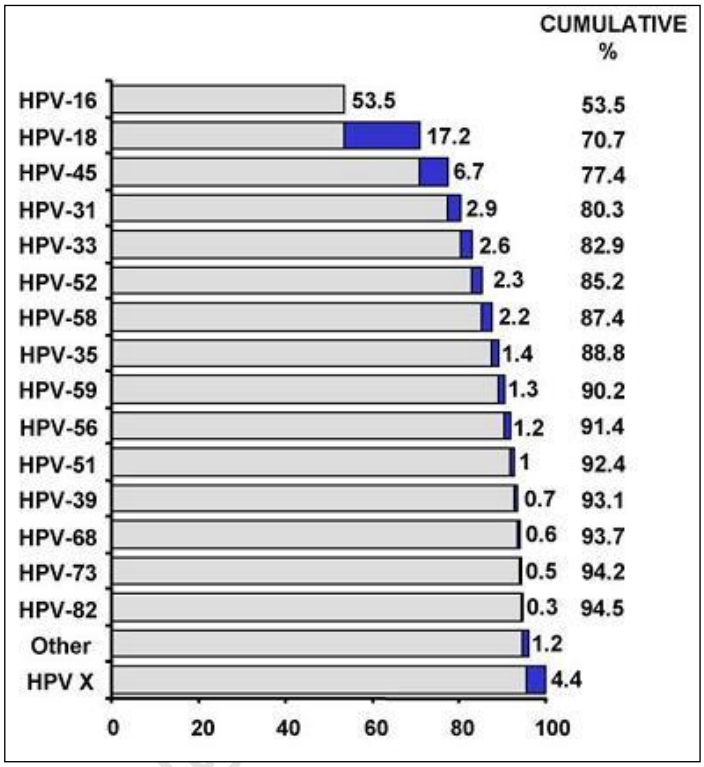


Figure 1.5 HPV types associated with cervical cancer (Clifford *et al*., 2006, adapted from Munoz *et al*., 2004).

1.6 Association of HPV types, multiple infections and viral load with abnormal cervical cytology

HPV factors involved in the progression of an HPV infected cervix to invasive cancer include HPV type, single or multiple infections and viral load (Castellsague and Munoz, 2003; Moscicki *et al*., 2006). It is known that the clearance rate of HR-HPVs is lower than for low risk (LR) HPV types, since HR-HPV types tend to have a higher viral load (Trottier *et al*., 2008). HR-HPV types with high viral loads are associated with HSIL (Hesselink *et al*., 2009). Kovacic and others (2006) investigated the associations of cytologic abnormality with viral load of type-specific HPV infection. An increased viral load of any single HR-HPV type was associated with cytologic abnormalities (P *trend*<0.0001) (Table 1.4). This effect was largely caused by Alpha (α) 9

(consisting of HPVs; 16, 31, 33, 35, 52, 58 and 67) and α11 (consisting of HPVs 34 and 73) (P *trend*<0.0001), especially HPV 16 (P *trend*<0.0001). The association between abnormal cytology and viral load was also significant for α5/α6 types (P *trend*<0.04) and slightly significant for α7 type (P *trend*<0.05). There were no significant association for noncarcinogenic HPV types from α1/α8/α10/α13 (P *trend*<0.5) and α3/α15 species (P *trend*<0.4). It was found that the percentage of cytologic abnormality varied greatly depending on the infecting carcinogenic HPV species with respect to its viral load (Kovacic *et al*., 2006). This further suggests that carcinogenic high viral load in ≥ cervical intraepithelial neoplasia 2 (CIN2) is not only uniquely associated with HPV 16 infection (Gravitt *et al*., 2007). The association between ≥ CIN2 and HR-HPV types with high viral load was significant even after excluding women with HPV 16 (Kovacic *et al*., 2006).

Table 1.4 Percentage of women with atypical squamous cell of undetermined significance (ASCUS) or worse cytologies by qualitative viral load (Kovacic *et al*., 2006).

HPV type/group/species	Low (1)		Moderate (2-3)		High (4-5)		Total N	P_{trend} †
	N	% Abnormal ‡	N	% Abnormal	N	% Abnormal		
α1/α8/α10/α13	18	11.1	38	21.1	16	18.8	72	0.5
α9/α11	48	4.2	104	23.1	210	46.7	362	**<0.0001**
α9/α11 (HPV16-)	26	7.7	62	25.8	118	39.0	206	**0.001**
α7	43	18.6	66	19.7	109	32.1	218	**0.05**
α5/α6	32	25.0	60	20.0	111	37.8	203	**0.04**
α3/α15	62	11.3	135	11.9	170	14.7	367	0.4
All carcinogenic	82	13.4	173	23.7	315	45.7	570	**<0.0001**
Carcinogenic (HPV16−)	60	18.3	131	25.2	223	41.3	414	**<0.0001**
Noncarcinogenic	121	13.2	230	13.9	301	19.6	652	0.06

PCR signal strength was expressed as; 1 indicating a low viral load, 2-3 indicates an intermediate viral load and 4-5 indicate the highest viral load.

HPV 16 is the most frequent HPV type in cervical cancer irrespective of geographical area (Trottier *et al*., 2008), and is the type most likely to be associated with progression to CIN3 (Schiffman *et al*., 2009; Winer *et al*., 2005). Since HPV 16 has a higher intratypic variation than other HR-HPVs (Schiffman *et al*., 2010; Raiol *et al*., 2009), it has been suggested that this type is not as efficiently cleared by the immune system as other HR types (Palefsky *et al*., 2006; Strickler *et al*., 2003). Longitudinal studies have shown

a correlation between HPV 16 viral load and rapid progression from infection to CIN3 (Dalstein *et al*., 2003; Santos *et al*., 2003; *Xi et al*., 2011).

Several studies have concluded that women with multiple HPV infections are at higher risk for developing HSIL compared to women infected by single HPV type (Mejlhede *et al*., 2009; Pista *et al*., 2011; Spinillo *et al*., 2009; Trottier *et al*., 2008). Other studies have found no association between multiple HPV infections with developing HSIL (Bosch and Munoz, 2002; Selva *et al*., 2009). This is due to the fact that in individuals infected with multiple HPV types, HSIL is caused by a single HPV type. It is speculated that other types are likely to be minor and have no additional risk contributing to neoplasia (Gargiulo *et al*., 2007). Therefore if the multiple infections are due to noncarcinogenic HPV types, then this will not increase the risk of developing cervical cancer even though noncarcinogenic HPV types may have a high viral load (Gravitt *et al*., 2007; Kovacic *et al*., 2006).

1.7 HPV variants

HPV genotypes differ by more than 10% from each other in DNA sequence based on the L1 open reading frame. HPV types that are closely related by approximately 80-90% are classified as members of the same species (Bernard *et al*., 2010). HPV variants differ by less than 2% in the L1 gene and by up to 5% in the LCR (Calleja-Macias *et al*., 2005; de Villiers *et al*., 2004; Ho *et al*., 1993a). Therefore, the LCR is the most variable region in the entire HPV genome (Chan *et al*., 2011; Ho *et al*., 1993a).

The LCR contains regulatory elements that are involved in viral DNA replication and transcription of various factors that either suppress or activate the p79 promoter (Cornet *et al*., 2012; Lei *et al*., 2011; Pande *et al*., 2008). The p79 promoter is located upstream of the oncogenic early genes (Cripe *et al*., 1990) and is responsible for the regulation of various HPV genes, including the E6 and E7 oncogenes (Lei *et al*., 2011; Pande *et al*., 2008). The E6/E7 oncogenes are involved in the formation of tumours (Cornet *et al*.,

2012; Mazumder *et al*., 2011). Therefore sequence variations in the LCR and E6/E7 oncoproteins may reflect different oncogenetic potential among different HR-HPV variants (Kammer *et al*., 2002; Veress *et al*., 1999). Furthermore, an assessment of HPV variants with the severity of lesion should be determined in order to determine the risk of different HPV variant in cervical carcinogesis (Cento *et al*., 2011; Chan *et al*., 2002; Chang *et al*., 2011; Pande *et al*., 2008).

The LCR is divided into three functionally distinct regions which are termed; 5' region, central region and 3' region (Figure 1.6). The 5' region is approximately 300 base pairs (bp) in size, and is located between the L1 stop codon and the E2 binding site (O'Connor *et al*., 1995). This region contains nuclear matrix attachment regions and signals for transcriptional termination (Stunkel and Bernard, 1999). The central region is located between two E2 binding sites namely E21 and E22 (Figure 1.6), and is approximately 400 bp and contains epithelial cell-specific enhancer (O'Connor *et al*., 1995; Stunkel and Bernard, 1999). The 3' region is 140 bp and contains the origin of replication (ori) (Kammer *et al*., 2000; O'Connor *et al*., 1995; Stunkel and Bernard, 1999).

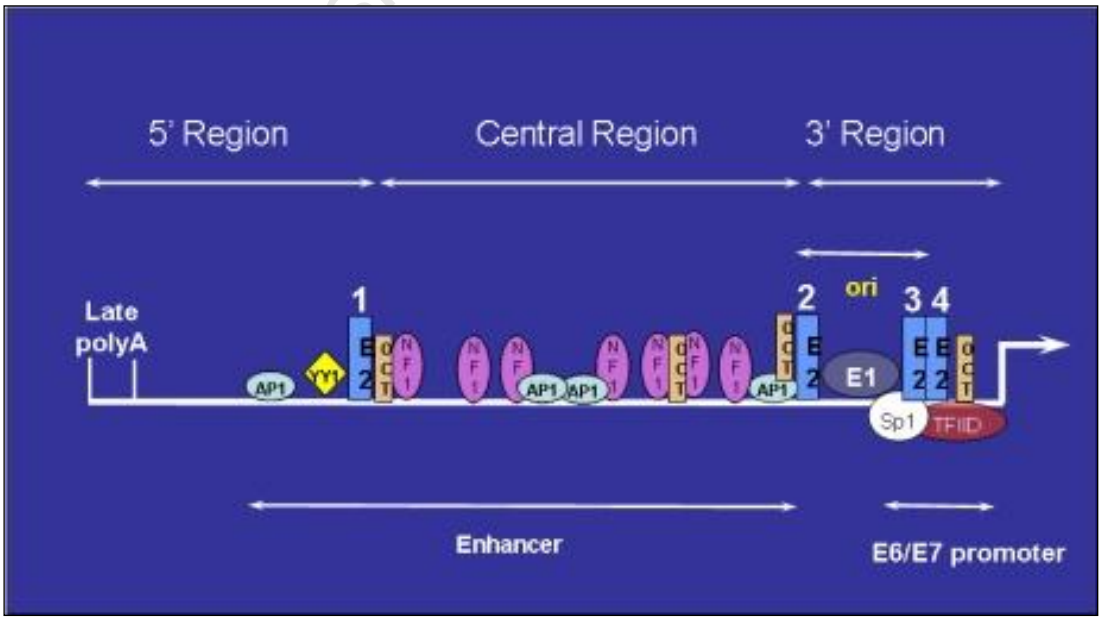


Figure 1.6 Organization of HPV 16 LCR. The LCR is divided into three main regions termed; 5' region, central region and 3' region to which various transcription factors bind (htt://wenliang.myweb.uga.edu).

Studies have shown that variants of specific HPV types are often grouped according to the specific geographic area or human ethnic group from which they were isolated (Burk *et al*., 2009; Ho *et al*., 1993b; Lizano *et al*., 2009; Ong *et al*., 1993; Tanzi *et al*., 2009). HPV 16 variants are grouped into five distinct phylogenetic branches based on the LCR sequences (Figure 1.7) (Ho *et al* 1993a; Mendoza *et al*., 2013; Tanzi *et al*., 2009; Yamada *et al*., 1997).

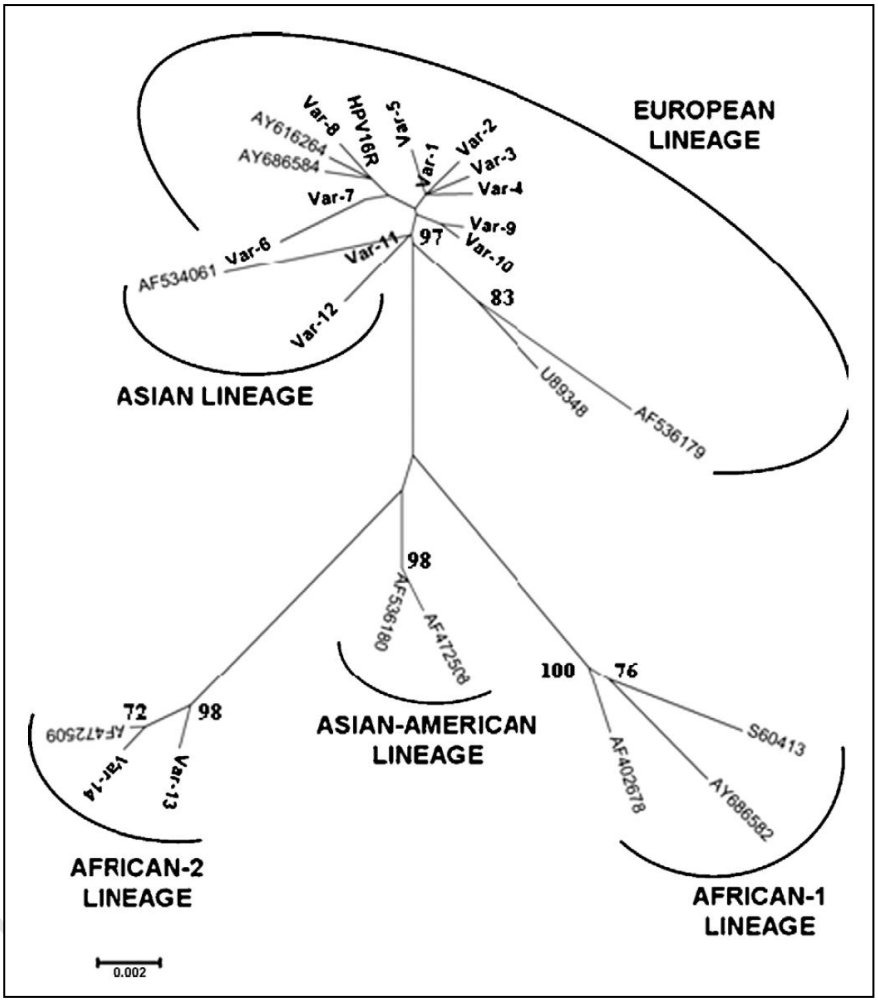


Figure 1.7 HPV 16 phylogenetic tree based on LCR sequences (Tanzi *et al*., 2009)

These separate branches were found to correspond to specific ethnic groups or geographic locations; Asian (As), Asian American (AA), African 1 (Af1), African 2 (Af2) and European (E) (Ho *et al*., 1993b; Tanzi *et al*., 2009). The inclusion of the E6 gene, gives greater phylogenetic resolution resulting in the addition of an additional branch termed the North American variant (NA) (Yamada *et al*., 1995; Yamada *et al*., 1997). Cornet and colleagues (2012)

grouped the E branch with the As branch to form the European-Asian (EAS) branch and grouped the NA branch with AA branch to form the Asian-American/ North-American (AA/NA) branch. Therefore variants of HPV 16 are currently grouped into four major branches based on E6 and LCR which are; European-Asian (EAS), African 1 (Afr1), African 2 (Afr2) and North American/Asian American (NA/AA) (Cornet *et al*., 2012).

There are two major variants in Africa termed African 1 and African 2 variants. Both the African 1 and African 2 lineages consist of sublineages namely; Afri1a, Afr1b, Afr2a and Afr2b. The Afr1a sublineage is found mostly in sub-Saharan Africa while the Afr1b lineage is found mostly in North Africa. The Afr2a sublineage is almost evenly distributed between sub-Saharan and North Africa while Afr2b is more prevalent in Sub-Saharan Africa than North Africa. The As sublineage is prevalent in Eastern Asia and the Eur sublineage is distributed across the globe. The NA sublineage is found frequently in North Africa while the Asian American 1 (AA1) and Asian American 2 (AA2) sublineages of the AA branch are both commonly found in South/Central America, with the AA1 sublineage most prominent in Asia (Cornet *et al*., 2012). Although the distribution of variants show ethnogeographic predilection some variants do exist in other regions; however not with the same prevalence as in their original geographic location (Tu *et al*., 2006).

There is limited knowledge of HPV 58 sequence variation (Canche *et al*., 2010; Chan *et al*., 2011; Raiol *et al*., 2009). The phylogenetic tree of HPV 58 based on the LCR nucleotide sequences suggests that variants of this type have evolved into 4 lineages, termed A, B, C and D (Figure 1.8) (Chan *et al*., 2011; Chen *et al*., 2011a; Liu *et al*., 2012). The ethnogeographic separation of HPV 58 lineages is not as conspicuous as that for HPV 16 variants. Nonetheless the distribution of HPV 58 variants does show some association with ethnogeographic origin (Chan *et al*., 2011). Lineage A is the most prevalent lineage, occurring worldwide. It can be divided into three sublineages; A1, A2 and A3 with sublineage A2 being the most prevalent.

Lineage B is more common in America than Africa and is subdivided into sublineages B1 and B2. Lineages C and D are more prevalent in Africa than America and Europe. Lineage C has not been subdivided into sublineages, while lineage D can be subdivided into sublineages D1 and D2 (Chan *et al*., 2011; Godinez *et al*., 2013).

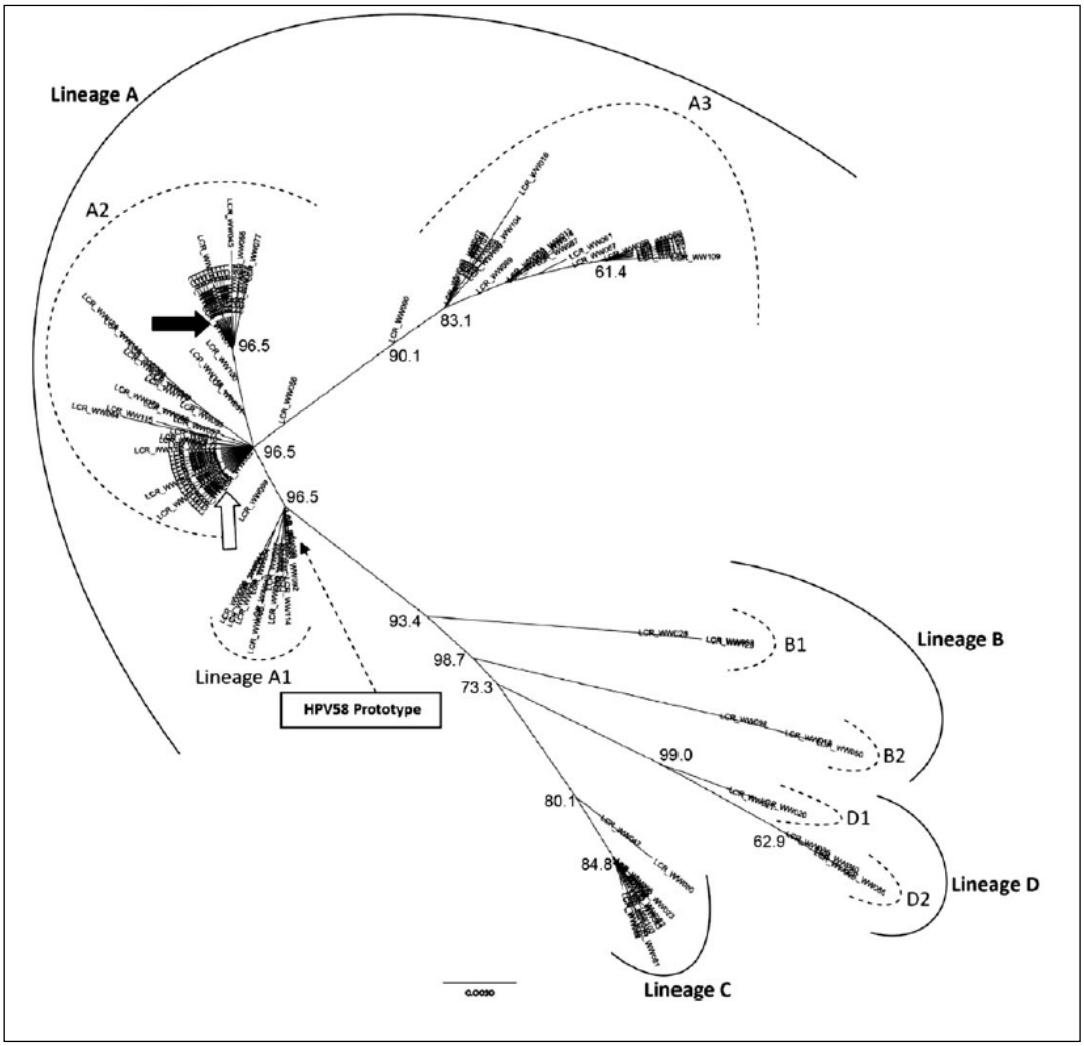


Figure 1.8 Phylogenetic tree of HPV 58 LCR sequences showing four lineages: A, B, C and D. Black and white arrows show the two most prevalent variant sequences (Chan *et al*., 2011)

HPV 53 variants do not reveal any ethnogeographic clustering (Oliveira *et al*., 2012; Wyant *et al*., 2011), but instead split into two lineages consisting of prototype-like (P-L) and non prototype-like (non P-L) variants (Cento *et al*., 2012; Kocjan *et al*., 2007; Oliveira *et al*., 2012; Prado *et al*., 2005; Wyant *et al*., 2011). Based on the LCR both lineages further form star-like phylogenetic clusters (Figure 1.9) (Kocjan *et al*., 2007), which could possibly have different

oncogenic potential of variants belonging to a particular cluster (Kocjan *et al*., 2007; Wyant *et al*., 2011). Alterations within the LCR may affect the binding affinity of cellular and viral transcriptional factors (Wyant et al., 2011), such as Yin Yan 1 (YY1) factors (Park et al., 1999). However further studies are needed to clarify such assumption (Kocjan *et al*., 2007).

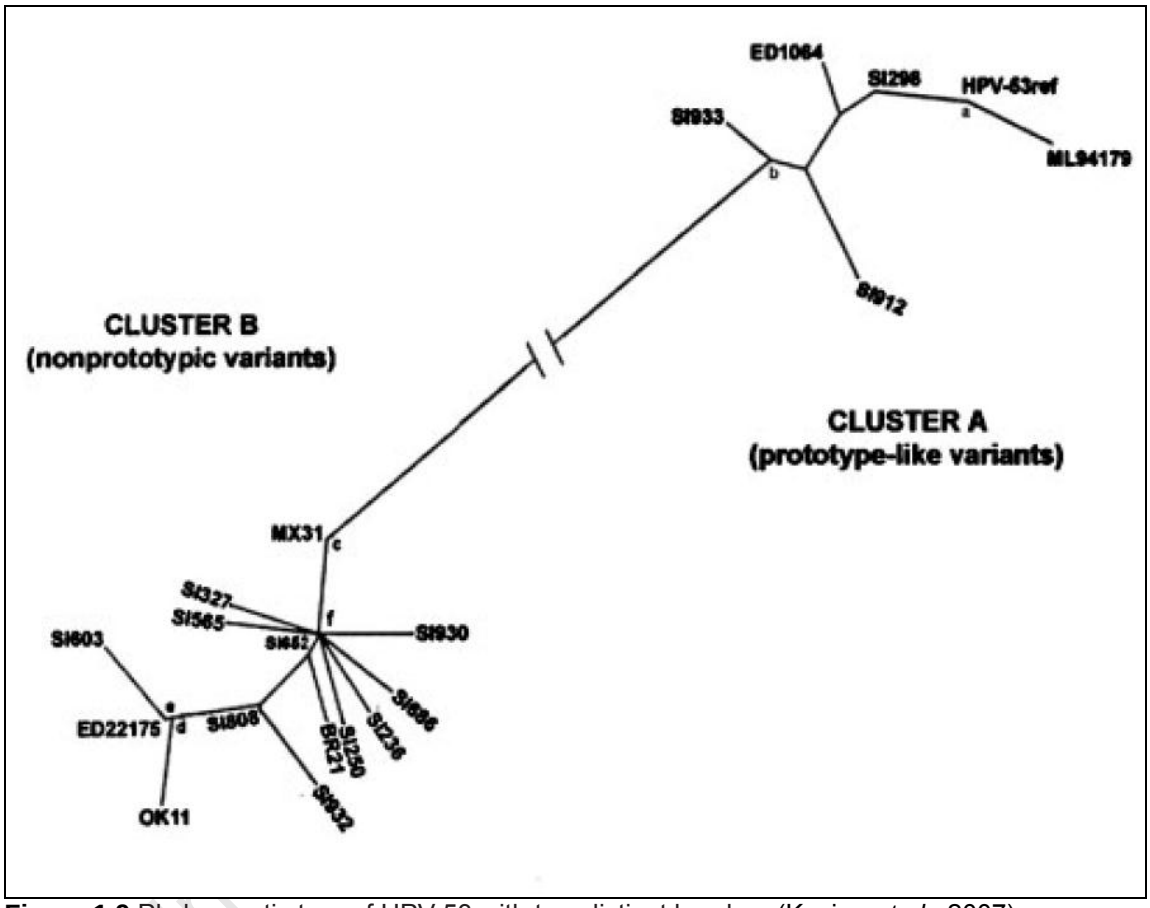


Figure 1.9 Phylogenetic tree of HPV 53 with two distinct braches (Kocjan *et al*., 2007).

1.7.1 HPV variants and associations with cervical diseases

All HPV 16 variants are carcinogenic (Schiffman *et al*., 2010) but some variants may have higher oncogenic potential compared to others (Bokal *et al*., 2010; Hildesheim *et al*., 2001; Lizano *et al*., 2006; Schiffman *et al*., 2010; Sichero *et al*., 2012). These differences in carcinogenicity may likely be explained by differences in the biological, chemical and pathogenic properties of the variants, as well as difference in the host's immunological response. Sequence variation may affect viral gene expression and viral protein properties, in turn affecting viral persistence and cell transformation.

The difference in carcinogenicity of HPV 16 variants may be due to mutation at the various transcriptional factor binding sites. The LCR contains transcription factor binding sites for the activation of transcription of the HPV oncogenes, E6 and E7. Mutations occurring in this region could therefore interfere with the activities of transcription factors and may potentially have significant role in tumorigenesis (Stephen *et al*., 2000). For example, YY1 factors are known to play a role in the maintenance of low levels of HPV transcription (Park *et al*., 1999) by repressing the p79 promoter (Stephen *et al*., 2000). YY1 also inhibits the LCR activity responsible for the expression of the E6/E7 oncogenes (Park *et al*., 1999; Stephen *et al*., 2000). Therefore mutations of YY1 motif in the LCR could lead to increase levels of HPV gene expression (Park *et al*., 1999), resulting in viral persistence and cell growth changes (Lace *et al*., 2009; Park *et al*., 1999).

Hildesheim and others (2001) investigated the risk of developing cervical cancer in HPV 16 positive women infected with specific HPV 16 variants. Results showed that the prevalence of non-European variants was 15.3% in high grade squamous intraepithelial lesions (HSIL) and 43.7% of cervical cancer. For European like variants, results showed that the prevalence was 10.7% in HSIL and 0% cancers. These results provide evidence that different HPV 16 variants differ in oncogenicity (Hildesheim *et al*., 2001). Non-Eur variants are reported to have an increased oncogenic potential compared to Eur variants (Berumen *et al*., 2001; Hildesheim *et al*., 2001; Kammer *et al*., 2000; Lei *et al*., 2011; Ordonez *et al*., 2004; Pista *et al*., 2007; Sichero *et al*., 2007; Tornesello *et al*., 2004). Among Portuguese women non-Eur variants showed a significant statistical difference in frequency from 7.3% in normal cytology to 59.3% in invasive cervical cancer. African variants were associated with significant increased risk and were found in 44.4% cases of invasive cervical cancer (Pista *et al*., 2007). Among Italian women, non-Eur AA variant was found increasing with severity of lesion from 5.9% in CIN1 to 19.4% in ICC (Tornesello *et al*., 2004). However among Slovenian women, 95% of cervical cancer cases were associated with European variants while

non-Eur variants were only found in 5% cases (Bokal, *et al*., 2010). Similar finding were observed in a South African study, European variants were responsible for 79% of ICC, 14% were associated with African variants and only 7% were associated with AA variants (Tu *et al*., 2006). Therefore the increase virulence observed with non-Eur variants in various populations may not be caused by inherent differences in virulence of different HPV 16 variants but reveal the ability of a given population to effectively mount a rigorous immunologic response against a specific HPV 16 variant prevalent (Tu *et al*., 2006).

In a recent study that attempted to identify sequence variation of HPV 58 that could possibly carry a higher oncogenic risk, isolates carrying mutations G63S and T20I in the E7 gene were reported to have an independent risk for CIN3 and invasive cervical cancer. These high risk signature mutations were found in 33% of isolates from Asia, 10% in America, 3% in Europe but not found in Africa (Chan *et al*., 2012).

HPV 53 variants may have different oncogenic potential however there is little knowledge regarding HPV 53 variants (Oliveira *et al*., 2012; Wyant *et al*., 2011), further studies are need to confirm such speculations (Kocjan *et al*., 2007).

Mutations in LCR of HPV contribute to the alternative mechanism involved in the oncogenicity of HPV and demonstrate a correlation between HPV infection and progression to cervical cancer (Chan *et al*., 2012; Pientong *et al*., 2013). Therefore since HPV variants differ by up to 5% in the LCR (Calleja-Macias *et al*., 2005; de Villiers *et al*., 2004; Ho *et al*., 1993a), it renders LCR the best appropriate surrogate for understanding oncogenetic potential of different HPV variants. This is due to the fact that the LCR contains regulatory elements that are involved in viral DNA replication and transcription of various factors that either suppress or activate the p79 promoter (Lei *et al*., 2011; Pande *et al*., 2008; Stephen *et al*., 2000). The

p79 promoter is located upstream of the oncogenic early genes (Cripe *et al*., 1990) and is responsible for the regulation of various HPV genes, including the E6 and E7 oncogenes (Lei *et al*., 2011; Pande *et al*., 2008). The E6/E7 oncogenes are involved in the formation of tumours (Mazumder *et al*., 2011).

Due to the above reasons, we therefore focused on studies that have investigated various regions of the LCR, since this region has been deemed the most appropriate surrogate for understanding HPV variants oncogenicity. Results from various studies investigating HPV 16, 58 and 53 LCRs are summarised below in Table 1.5. Please note that the South African Study done by Tu et al (2006) is not included in the table below due to the fact that they analysed the E6 region only but not the LCR.

Table 1.5 Review of studies describing HPV 16, 58 and 53 variants based on LCR

HPV 16 variants

Country	Authors	Genome region	Region in the LCR covered	No of isolates Analyzed	Method	Results
India	Pande *et al*., 2008	E6, E7, L1 and LCR	7521-13	60 cervical isolates	PCR and DNA sequencing	LCR variants consisted of 38 Eur, eight AA, seven Afr1, two Afr2 and five new variants. A total of 17 nucleotide substitutions were observed in the region (7521-13), the most common nucleotide substitution was a transition G7521A.
India	BhattacharjeeandSengupta, 2006	E2 and LCR	7394-7868	61 cervical isolates	PCR and DNA sequencing	LCR variant analysis revealed that 55 isolated belonged to the Eur lineage and 6 isolates belonged to the AA lineage
Mexico	Lugo-Trampe *et al*., 2012	LCR	7433-83	240 cervical isolates	PCR-restriction fragment length polymorphism (RFLP) and DNA sequencing	LCR variant analysis showed that 176 isolates cluster with the Eur lineage, 53 isolates were from the AA lineage and 11 isolates clustered with the Afr2 lineage. A total of 31 nucleotide substitutions were observed in the region (7433-83), the most common nucleotide substitution was a transition G7518A.
Slovenia	Bokal *et al*., 2010	E6, E7 and LCR	7161-7947	40 cervical isolates	PCR and DNA sequencing	The most common HPV 16 lineages found in cervical isolates among Slovenian women in descending order were; Eur lineages (95%), Afr (2.5%) lineages and the (2.5%) AA lineages. A total of 44 nucleotide substitutions were observed in the region (7161-7947), the most common nucleotide substitution was a transition G7519A.
Poland	Schmidt *et al*., 2001	LCR	7434-7845	60 cervical isolates consisting of European variants	PCR and DNA sequencing	Cervical cancer isolates mostly contain changes in sequences of YY1 binding sites whereas isolates from asymptomatic carriers harbour nucleotide changes within or close to transcription binding sites. Cervical cancers isolates mostly carried the transition mutation G7519A.
Italy	Tanzi *et al*., 2009	LCR	7315-31	58 genital isolates from men and women	PCR and DNA sequencing	A total of 14 different HPV 16 LCR variants were found, consisting of ten Eur lineages, two Asiatic lineages and two Afr2 lineages. A total of 26 nucleotide substitutions were observed in the region (7315-31), the most common nucleotide substitution was a transition G7520A.

Country	Authors	Genome region	Region in the LCR covered	No of isolates Analyzed	Method	Results
Finland	Kurvinen *et al.,* 2000	LCR	7139-145	37 cervical isolates	PCR, Cloning and sequencing	A total of 14 HPV 16 LCR variants were observed, consisting of 13 Eur lineages and one As lineage. Eur variants did not show enhanced pathogenicity. A total of 26 nucleotide substitutions were observed in the region (7139-145), the most common nucleotide substitutions were a transversion G7193T and a transition G7521A.
Uganda	Buonaguro *et al*., 2000	E6, E7, L1 and LCR	7419-84	7 Biopsies of penile squamous-cell carcinoma, 2 Cervical dysplasia and 3 Cervical carcinoma	PCR and DNA sequencing	HPV 16 Afr1 variant was identified in 100% tumours, 6.25% cervical scrapes and was associated with cancer progression. A total of 20 nucleotide substitutions were observed in the region (7419-84), the most common nucleotide substitutions were the transitions G7419A, G7488A, T7763C, and T7785C.
Zambia	Lei *et al*., 2011	LCR	7405-83	11 vaginal lavage isolates	PCR and DNA sequencing	A total of 11 HPV 16 positive isolates were analysed, consisting of five Eur wild-type and six novel HPV 16 variants. From the six novel HPV 16 variants, five were identified as Eur variants and the remaining variant fell into the Afr lineage. A total of 12 nucleotide substitutions and six insertions were observed in the region (7405-83), the most common nucleotide substitutions was a transitions G7797A.
Thailand	Chansaenroj *et al*., 2012	E2, E6, E7,L1 and LCR	7155-7886	8 HPV 16 cervical isolates	PCR and DNA sequencing	All specimens harbouring single HPV 16 infections had Eur variants whereas all HPV 16 multiple infected specimens had non-Eur variants. Furthermore single HPV 16 infections caused by the Eur variant may significantly increase the risk for developing cervical cancer whereas multiple HPV 16 infection caused by non-Eur variants may require many variations to enhance the risk of cervical cancer development. A total of 23 nucleotide substitutions were observed in the region (7175-7886), the most common nucleotide substitutions were a transversion G7193T and a transition G7521A.

Country	Authors	Genome region	Region in the LCR covered	No of isolates Analyzed	Method	Results
Worldwide	Yamada *et al.,* 1997	E6, L1 and LCR	7485-7842	408 cervical isolates	Dot-blot hybridization and sequencing	HPV 16 variants may be useful for investigation the risk of cervical neoplasia and could be informative for the design of HPV 16 vaccine strategies. A total of 26 nucleotide substitutions were observed in the region (7485-7842), the most common nucleotide substitution was a transition G7521A.
USA	Burk *et al.,* 2003	E6 and LCR	7450-7876	68 cervical isolates	PCR and DNA sequencing	HPV 16 non-Eur variants were frequently observed in adenocarcinomas. A total of 32 nucleotide substitutions were observed in the region (7450-7876), the most common nucleotide substitution was a transition G7489A.
China	Shang *et al*., 2011	E6, E7, L1 LCR	7168-48	52 cervical isolates	PCR and DNA sequencing	The most prevalent HPV 16 variants were from the Eur lineage (67.13%) followed by the As lineage (32.69%). A total of 32 nucleotide substitutions were observed in the region (7168-48), the most common nucleotide substitution was a transition G7521A.
Italy	Tornesello *et al.,* 2004	E6, E7, L1 and LCR	7432-83	90 cervical isolates	PCR and DNA sequencing	Non-Eur variants were found to be more oncogenic compared to Eur variants. A total of 17 nucleotide substitutions were observed in the region (7432-83), the most common nucleotide substitution was a transition G7521A.
Italy	Tornesello *et al*., 2008	E6, E7 and LCR	7485-83	18 penile specimens	PCR and DNA sequencing	HPV 16 A A variants were found to be more oncogenic compared to Eur variants. A total of 14 nucleotide substitutions were observed in the region (7585-83), the most common nucleotide substitution was a transition G7521A.
Uganda	Tornesello *et al*., 2000	LCR	7289-93	5 penile specimens	PCR and DNA sequencing	Rearrangements within the HPV 16 LCR were found in two out of five penile carcinomas suggesting that natural variants may play a significant role in the pathogenesis of genital carcinoma. A total of seven nucleotide substitutions were observed in the region (7289-93), the most common nucleotide substitutions were the transitions the G7489A and G7521A.
Uganda	Tornesello *et al*., 1997	E6, E7, L1 and LCR	7420–83	5 penile specimens	PCR and DNA sequencing	HPV 16 Afr1 variant were found integrated in four out of five penile carcinoma isolates.A total of 20 nucleotide substitutions were observed in the region (7420-83), the most common nucleotide substitution was a transition G7519A.
Portugal	Pista *et al* 2007	E6, L1 and LCR	7485-7834	187 cervical isolates	PCR and DNA sequencing	HPV 16 Afr variants were associated with increased risk for cervical cancer. A total of 13 nucleotide substitutions were observed in the region (7485-7834), the most common mutation was G7521A.

Country	Authors	Genome region	Region in the LCR covered	No of isolates Analyzed	Method	Results
Argentina	Picconi *et al.,* 2003	L1, E6 and LCR	7485-7842	20 cervical isolates	PCR-dot blot hybridization and sequencing	HPV 16 LCR phylogeny revealed 69% of variants were from the Eur lineage, 19% were from the A A lineage and 12% belonged to the As lineage. A total of 13 nucleotide substitutions were observed in the region (7485-7842), the most common nucleotide substitution was a transition G7521A.
Brazil	Junes-Gill *et al.,* 2008	E6 and LCR	7485-7839	81 cervical isolates	RFLP and sequencing	HPV 16 phylogenies did not show correlation between ethnicity and HPV 16 variant in a Brazialn population consisting of mixed races A total of 16 nucleotide substitutions were observed in the region (7485-7839), the most common nucleotide substitution was a transition G7521A.
China	Stephen *et al*., 2000	E6, E7 and LCR	7521-24	21 Cervical isolates	PCR and DNA sequencing	Mutations in the LCR may alter the oncogenicity of HPV in the following two ways, either by changing amino acid sequences of the oncoprotiens E6 and E7 or through alterations to transcriptional binding sites. A total of 25 nucleotide substitutions were observed in the region (7521-24), the most common nucleotide substitution was a transition A7767G and G7769A.
Korea	Park *et al*., 1999	LCR	7484-24	27 cervical isolates	PCR, cloning and DNA sequencing	Nucleotide substations at the YY1 binding site are associated to the development of cervical neoplasia. A total of 18 nucleotide substitutions were observed in the region (7484-24), the most common nucleotide substitution was a transition G7520A.
Brazil	Alencar *et al*., 2007	LCR	7233-7886	22 cervical isolates	PCR and DNA sequencing	A total of 12 Eur, Eight AA, one Afr1 and one Afr2 HPV 16 LCR variants were characterised. A total of 29 nucleotide substitutions were observed in the region (7233-7886), the most common nucleotide substitution was a transition G7521A.
Australia and New Caledonia	Watts *et al.,* 2002	E2, E4, E6 and LCR	7172-73	34 cervical isolates	PCR and DNA sequencing	A majority (82%) of the variants belonged to the Eur lineage, 12% belonged to the As branch and 2% belonged to the AA lineage. A total of 41 nucleotide substitutions were observed in the region (7172-73), the most common nucleotide substitutions were a transversionG7193T & a transition G7521A.
Colombia	Moreno-Acosta *et al*., 2008	E6 and LCR	7436-7786	15 cervical isolates	PCR-single strand conformation polymorphism (SSCP) and sequencing	HPV 16 phylogeny showed that a (88.2%), of variants belonged to the Eur lineage 8.8% of samples had AA lineages and in 2.9% of samples the HPV 16 lineage could not be identified. A total of nine nucleotide substitutions were observed in the region (7436-7786, the most common nucleotide substitution was a transition G7521A.

Country	Authors	Genome region	Region in the LCR covered	No of isolates Analyzed	Method	Results
Canada	Mayrand *et al*., 2000	LCR	7172-7495	Cervical specimens from 50 women	PCR–SSCP analysis and sequencing	A total of 46 (92%) of women had persitant infection. A total of 28 nucleotide substitutions were observed in the region (7172-7495), the most common nucleotide substitution was a transversion G7192T.
USA	Jiang *et al*., 2009	E6 and LCR	7764-90	Cervical isolate from 1 individual cloned ten times	PCR, Cloning and sequencing	Eight clones belonged to the Eur lineage and the remaining two clones were identical to the Afr2 lineage based on the LCR only, but showed similarities with the Eur lineage in the E6 region, therefore were termed recombinations. A total of nine nucleotide substitutions were observed in the region (7764-90), the most common nucleotide substitution was a transversion A90T.
Worldwide	Cornet *et al*., 2012	E6 and LCR	7175-83	953 cervical isolates	PCR and DNA sequencing	HPV 16 variants can be separated into nine sublineages as following; Eur, As, Afr1a, Afr1b, Afr2a, Afr2b, NA, AA1, and AA2, base on the E6/LCR. A total of 27 nucleotide substitutions were observed in the region (7175-83), the most common nucleotide substitutions were transitions G7489A, C7764T, C7786T and a transversions C7689A.
Costa Rica, USA	Chen *et al*., 2005	Whole genome	7175-83	12 cervico-vaginal isolates	PCR, Cloning and sequencing	Overall, a total of 313 (4.0%) of nucleotide positions were found to vary in the entire HPV 16 genome. A total of 49 nucleotide substitutions were observed in the region (7175-83), the most common nucleotide substitutions were a transversion G7193T & a transition G7521A.
Singapore	Ho *et al*., 1993a	LCR	7483-7840	Cervical and penile isolates from 8 couples	PCR, Cloning and sequencing	HPV 16 variant concordance was observed in Four out of eight couples. A total of 13 nucleotide substitutions were observed in the region (7483-7840), the most common mutation was at position 7519 (nucleotide substitution was not shown).
Japan	Kozuka *et al*., 2000	E6 and LCR	7193-24	51 cervical cancer biopsies	PCR and DNA sequencing	LCR variant analysis revealed that 14 isolates were from the Eur lineage, 35 isolates from the As lineage and two isolated were from the AA lineage. Mutations in the Y Y1 motifs in the LCR increase the expression of viral oncogens.
Mexico	Kammer *et al*., 2000	LCR	7060-83	7 cerviacal cancer isolates	PCR, Cloning and sequencing	The E6-proximal end of the LCR is responsible for the enhanced transcriptional activities of the AA and NA variants. A total of 53 mutations were observed in the region (7060-83), the most common were; G7060T, G7193T, C7681A, G7489A, G7521A, C7764T, and C7786T.

Country	Authors	Genome region	Region in the LCR covered	No of isolates Analyzed	Method	Results
Sweden and Finland	Kammer *et al*., 2002	E6 and LCR	7033-78	45 cervical isolates	PCR and DNA sequencing	The different oncogenic potential seen in HPV 16 Eur variants may be due to alterations of E6 proteins instead of altered activity of the P97 promoter. A total of 32 nucleotide substitutions were observed in the region (7033-78), the most common nucleotide substitution was a transversion G7193T and transitions G7521A.
Thailand	Pientong *et al*., 2013	E6 and LCR	7175-81	47 cervical isolates	PCR and DNA sequencing	LCR variant analysis revealed that 61% of isolates clustered with the As lineage, 8.5% clustered with AA lineage, 25.6% clustered with the Eur lineage, 3.7% clustered with the Afr2 lineage and the remaining 1.2% was termed J135C. A total of 38 nucleotide substitutions were observed in the region (7193-81), the most common nucleotide substitutions were transversions; A7175C, G7193T, A7287C , A7730C, G81T transitions T7177C, T7201C, G7521A, C24T & transition/tranversion G7842A/T
Paraguay	Mendoza *et al*., 2013	LCR	7485-7834	67 cervical isolates	PCR and DNA sequencing	Eur variants were detected among women with and without normal cytology while non-Eur variants were detected only in women with cervical lesions.A total of 15 nucleotide substitutions were observed in the region (7485-7834), the most common nucleotide substitution was a transition G7521A.

HPV 58 variants

Country	Author	Genome region	Region in the LCR covered	No of isolates Analyzed	Method	Results
Worldwide	Chan *et al*., 2011	E2, E5, E6, E7, L1 and LCR	7257-7429 and 7540-52	401 cervical isolates	PCR and DNA sequencing	HPV 58 variants can be grouped into four lineages revealing some degree of predilection in distribution. A total of 24 nucleotide substitutions were observed in the region (7257-7429 and 7540-52), the most common nucleotide substitution was a transversion C30G.
Italy	Cento *et al*., 2011	E6, E7, L1 and LCR	7181-7540	23 cervical isolates	PCR and DNA sequencing	There was no correlation between HPV 58 variant and abnormal cervical cytology. A total of 30 nucleotide substitutions were observed in the region (7118-7540), the most common nucleotide substitution was a transition C7266T.
Taiwan	Chang *et al*., 2011	E6, E7 and LCR	7421-7775	115 cervical isolates	PCR and DNA sequencing	The most prevalent lineage was the A3 lineage 56.52% followed by A2 (26.09%), A1 (15.65%) and C (1.74%). A total of 20 nucleotide substitutions were observed in the region (7421-7775), the most common nucleotide substitution was a transversions/transition A7714C/G
Brazil	Raiol *et al*., 2009	E6, L1 and LCR	7265-7788	8 cervical isolates	PCR and DNA sequencing	Phylogenetic analysis of HPV 58 isolates did not show ethnic clustering. A total of 12 nucleotide substitutions were observed in the region (7265-7788), the most common nucleotide substitution was a transition T7575C.
China	Wu *et al*., 2009	Whole genome	7714-86	37 cervical isolates	PCR and DNA sequencing	Twelve variants were identified and eight of the variants belonged to the prototype-like group. The remaining four variants belonged to the non prototype-like group. A total of nine nucleotide substitution were observed in the region (7714-86), the most common nucleotide substitution was transition/transversion A7793G/C
China	Liu *et al*., 2012	E6, E7, L1 and LCR	7257-7421 and 7540-52	174 cervical isolates	PCR and DNA sequencing	Based on the LCR, a total of 149 isolates belonged to the A lineage and 24 isolates could not be classified. Only one isolate belonged to the B lineage. A total of 21 nucleotide substitutions were observed in the region (7257-7421 & 7540-52), the most common nucleotide substitution were a transition C7266T and transition/transversions A7714G/C

HPV 53 variants

Country	Author	Genome Region	Region in the LCR covered	No of Specimen Analyzed	Method	Results
Slovenia	Kocjan *et al*., 2007	E6, E7 and LCR	7380-14	70 cervical isolates	PCR and DNA sequencing	HPV 53 phylogeny revealed two lineages which further formed star-like clusters. A total of 29 nucleotide substitutions were observed in this region (7380-14) consisting of 27 single nucleotide exchanges, one 1 bp (T) and one 4 bp (TGGG) insertions. The most common nucleotide change was the transition G7508A
Brazil	Wyant *et al*., 2011	E6, L1 and LCR	7422-7810	6 cervical isolates	PCR, Cloning and sequencing	HPV 53 phylogenetic analysis did not reveal any ethnogeographical clustering. A total of 21 nucleotide substitutions were observed in the region (7422-7810). The most common nucleotide change was the transition G7508A
Italy	Cento *et al*., 2012	E6, E7, L1 and LCR	7422-7810	30 exo- and endocervical isolates	PCR and DNA sequencing	No specific association was found between nucleotide substitutions and grade of cytological lesions.A total of 17 nucleotide substitutions were observed in the region (7422-7810), the most common change was the transition mutation C7428T transition

1.8 HPV life cycle

HPVs infect the basal cells of stratified epithelium of mucosal and cutaneous skin (McBride, 2008; Pittayakhajonwut and Angeletti, 2010). In a normal epithelium, only cells in the lower basal layers are dividing mitotically and after cell division, one daughter cell is pushed up to restore the overlying differentiated layers. These cells are eventually sloughed from the surface of the epithelium (McBride, 2008). However during HPV infection, the outcome of the infection depends on several factors which include the site of infection, HPV type, host's immune system and differentiation of the host cell (Burchell *et al*., 2006; Doorbar, 2005; Doorbar, 2006; Schlecht *et al*., 2003). The latter factor is due to the fact that HPV genome replication is determined by the host's mitotically active cells (Cheng *et al*., 1995; Culp *et al*., 2006; Kadaja *et al*., 2009; McBride, 2008).

There are two stages in which HPV replicates and amplifies its genome, and these are termed, non-productive and productive cycle (Pittayakhajonwut and Angeletti, 2010). In the non-productive stage, infection is initiated when the virus enters the epithelium through microabrasion and reaches the undifferentiated dividing basal cell layers (Figure 1.10) (Cheng *et al*., 1995; Leggatt and Frazer, 2007; McBride, 2008; Moody *et al*., 2010; Munoz *et al*., 2006; Pittayakhajonwut and Angeletti, 2010; Pyeon *et al*., 2009). The virus enters the basal epithelial cells and reaches the host cell's nucleus where it replicates and remains as a viral episome (McBride, 2008), and no progeny virus produced. Viral DNA replication is achieved with the assistance of the early genes, E1 and E2 which induce the host's cellular DNA synthesis in order for the virus to utilize cellular replication proteins under S phase control (de Sanjose *et al*., 2010; Frazer *et al*., 2011; Narisawa-Saito and Kiyono, 2007; Pittayakhajonwut and Angeletti, 2010; Schiffman *et al*., 2007). E6 and E7 proteins are required to maintain S-phase-like state so that cellular replication machinery is available (McBride, 2008). As the basal cells divide through mitosis they differentiate to produce identical copies of themselves containing viral genes as episomes (Cheng *et al*., 1995; Evander *et al*., 1997; Munoz *et al*., 2006).

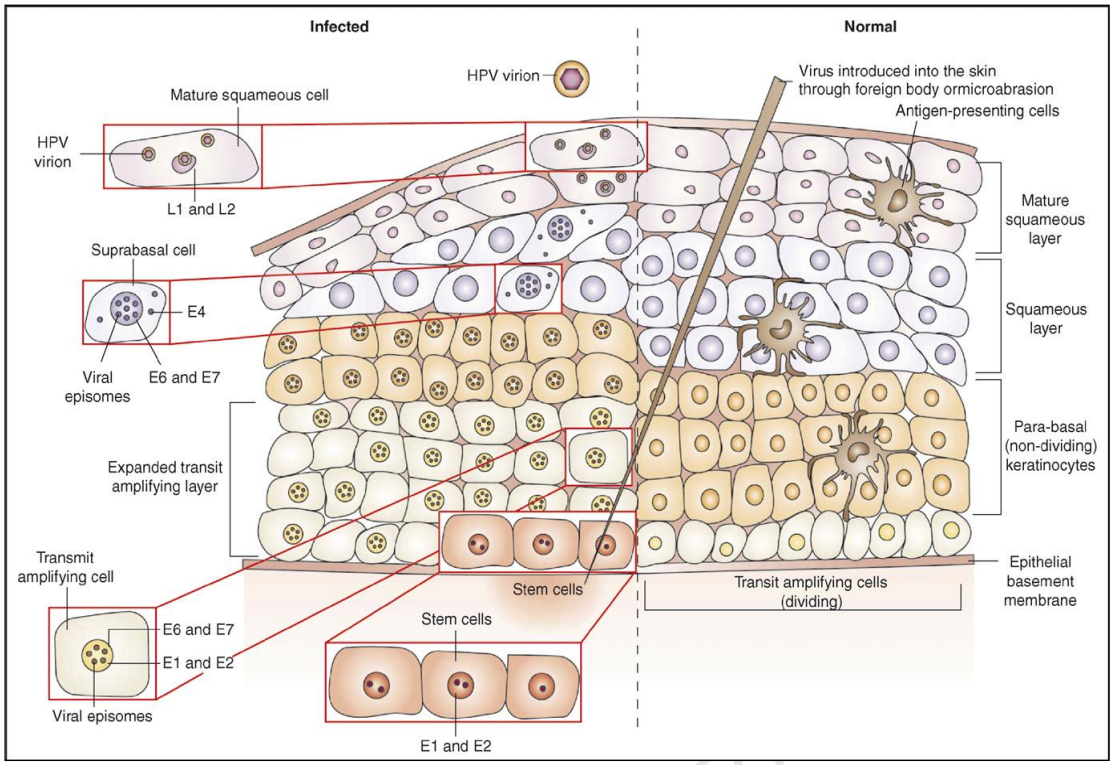


Figure 1.10 Biology of maturing squamous epithelium altered by an HPV infection (Leggatt and Frazer, 2007 reproduced from Munoz *et al*., 2006). The left side show expression sites of HPV non-structural proteins (E1, E2, E4, E6 and E7) and structural proteins (L1 and L2) as well as structural alterations to the epithelium caused by HPV infection. The right side of the Figure shows a normal maturing squamous epithelium.

The second stage of HPV amplification and replication is termed the productive cycle (Pittayakhajonwut and Angeletti, 2010). In the productive cycle the HPV DNA copy number is increased to hundreds or thousands per cell resulting in the production of progeny genomes which are ready to be encapsulated into virion particles (McBride, 2008). The process of viral genome amplification, late capsid protein synthesis and virion assembly occur in the upper, terminally differentiated cells of the epithelium (McBride, 2008; Pittayakhajonwut and Angeletti, 2010). Virions are assembled in the superficial differentiated layers and are found throughout the nuclei (McBride, 2008). The E4 gene assists in the expression of late genes in the squamous layer (Leggatt and Frazer, 2007; Munoz *et al*., 2006; Wilson *et al*., 2007). The late genes L1 and L2 make up the viral capsid which now encapsulates the HPV genome in the mature squamous layer (Florin *et al*., 2002; Leggatt and Frazer; Munoz *et al*., 2006). The process of viral encapsulation involves the expression of the minor coat protein L2 and L1 major coat protein (Doorbar *et al*., 2012). The cells containing viruses are now destined to be shed with dead cells from the epidermis (Florin *et al*., 2002; Kadaja *et al*.,

2009; Leggatt and Frazer, 2007; McBride, 2008; Moody *et al*., 2010; Narisawa-Saito and Kiyono, 2007).

1.8.1 Progression to cervical cancer

HPV infections result in the establishment of the viral genome as a stable episome in the basal layer cells characterized as low grade squamous intraepithelial lesions (LSIL) (Figure 1.11). In such lesions, the expression of gene products is carefully maintained in low levels with viral proteins produced at regulated levels at specific times as the infected cells move towards the epithelial surface (Doorbar, 2006; Kadaja *et al*., 2009; Moody *et al*., 2010).

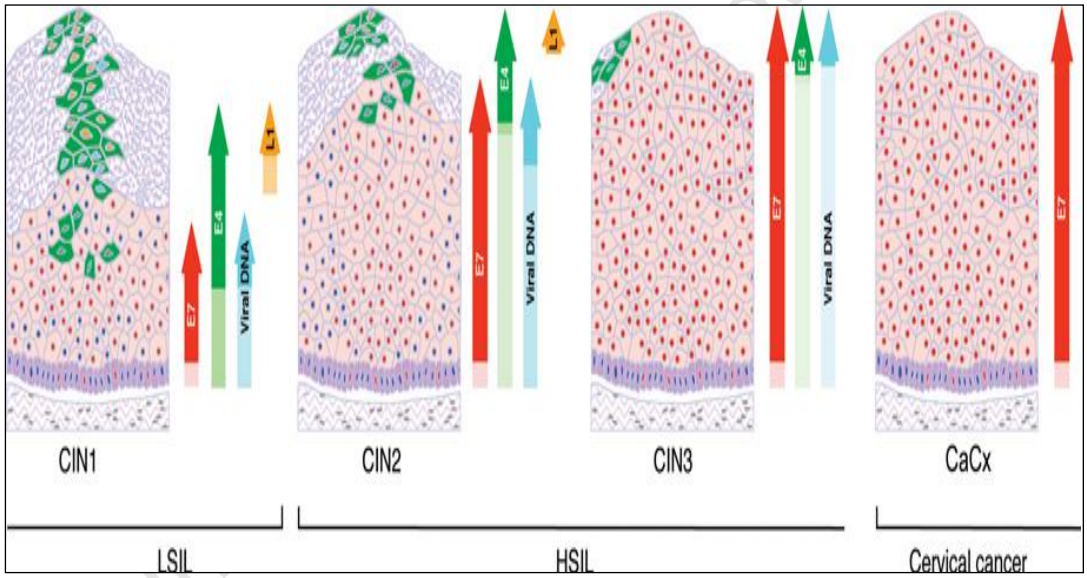


Figure 1.11 Changes in expression patterns that follow progression to cervical cancer (Doorbar, 2006). The pattern of viral gene expression changes during cancer progression. During LSIL (CIN1), the sequence of events is similar to that observed in productive lesions (diagrammatically on the far left). Whereas in HSIL, the start of late events is delayed, and although the order of events remains the same, the production of infectious virions becomes limited to smaller areas close to the epithelial surface. Integration of HPV sequences into the host cell genome can accur at the same time with these changes and may result to further deregulation in the expression of E7 and the loss of E1 and E2 replication protein. In cervical cancer (shown diagrammatically of the far right), the stages of the HPV life cycle are no longer sustained and viral episomes are normally lost.

Normally in low-grade lesions, HPV DNA exists as a double stranded circular episome (Huang *et al*., 2008). However in a majority of cervical cancer cells, a significant proportion of the HPV genome is integrated (Huang *et al*., 2008; von Knebel *et al*., 1991; von Knebel, 2002). The HPV genomic DNA

linearises and integrates itself with the host's genome in the cell nucleus. Linearization of HPV genome results in disruption of the E2 gene open reading frame (Huang *et al*., 2008; Palefsky, 2006). The disruption of the E2 gene results in the loss of ability to down regulate the expression of the transforming proteins E6/E7 (Palefsky, 2006; Schmidt *et al*., 2005). This leads to the inactivation of tumor suppressor gene products caused by E6 and E7 (Doorbar, 2006; zur Hausen, 2002; von Knebel *et al*., 1991). E6 blocks apoptosis by inhibiting p53 and E7 inhibits retinoblastoma protein (pRB), which normally stops cell cycle progression in keratinocytes. This results into uncontrolled cell division. Abnormal cells expressing large quantities of E6 and E7 multiply from the basal epithelial layers to the mature squamous layer (Munoz *et al*., 2006; zur Hausen, 2002; Schiffman *et al*., 2007). Elevated expression of the oncogenes protein E6/E7 contribute to cancer progression and the levels of E6 and E7 increase from LSIL to cervical cancer (Figure 1.11) (Doorbar, 2006).

1.9 Natural history of HPV infection

Cervical cancer usually develops after a long period of latency (Arnouk *et al*., 2009). The development occurs via a series of steps, namely, HPV infection, viral persistence, and progression of persistently infected cell to pre-cancer and eventually to cervical cancer (Figure 1.12) (Schiffman *et al*., 2007).

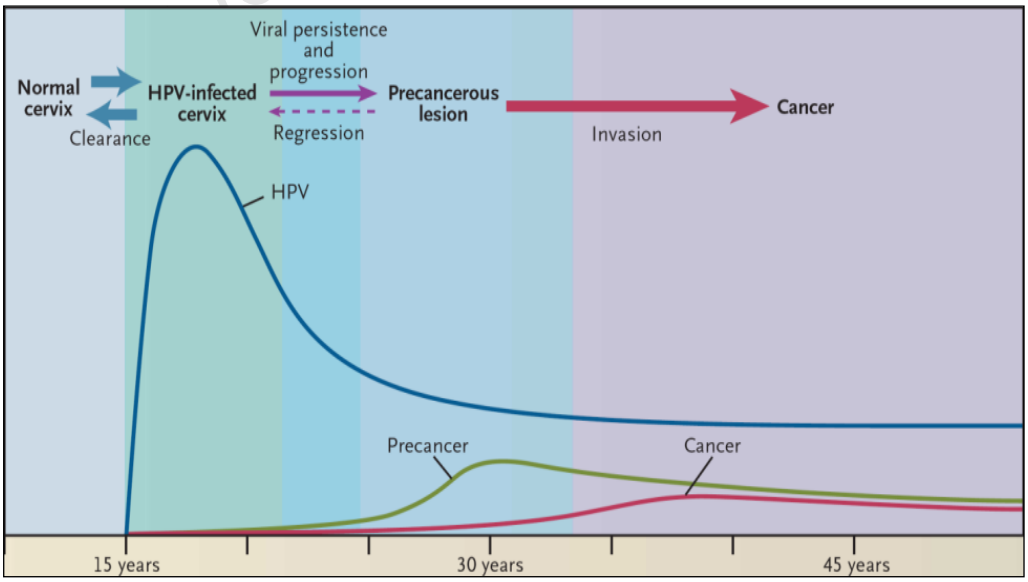


Figure 1.12 Stages leading to the development of cervical cancer after HPV infection (Schiffman and Castle, 2005).

HPV prevalence peak is observed in women less than 25 years of age (Cuschieri *et al*., 2004; de Sanjose *et al*., 2007). This is followed by a decline in DNA detection in the age group beyond 30 years (de Sanjose *et al*., 2007). The incidence of pre-cancers rise progressively for women over 25 years and cancer incidence are highest for women over 40 years (Bosch *et al*., 2008; Khan *et al.,* 2005; Schiffman and Kjaer, 2003; Snijders *et al*., 2006).

Anogenital HPVs are normally transmitted through sexual contact with an infected individual (Kjaer *et al*., 2001). The virus enters the cells of the epithelium through microabrasion in the skin and reaches the basal layers (Munoz *et al*., 2006). HPV infections of the cervical epithelium are associated with expression of non-structural viral proteins which induce minor alterations to epithelial cells causing squamous intraepithelial lesions (SIL) (Frazer *et al*., 2011). In the majority of individuals HPV infections are cleared and the clearance period of most HPV infections is approximately 6-12 months (Schmeink *et al*., 2013). However, in a minority of individuals whose immunity fails to clear the infection, the HPV infection will progress to cause precancerous lesions which is the next step towards advancing to cervical cancer (Doorbar, 2006). Around about 40% of untreated HSIL cases progress to invasive cancer (Castellsague *et al*., 2008; Peto *et al*, 2004). The period taken by HPV from initial infection to progress and cause HSIL takes about a decade (Ault, 2006; Plummer *et al*., 2007; Schiffman *et al*., 2007).

HPV 16 and 18 are the most persistent HR-HPV types, and together with other cofactors, such as; viral, environmental and host factors, can cause lesion that will eventually develop to cancer (section 1.3). The duration from HPV infection to cervical cancer normally takes between 10-20 years (Moscicki *et al*., 2006; Schiffman and Castle, 2005). Hence cervical cancer is very rare in women under 25 years, the incidence rises progressively for women over 25 years and is highest for women over 40 years (Khan *et al.,* 2005; Schiffman and Kjaer, 2003; Snijders *et al*., 2006).

In a study done by Matsumoto et al (2011), women with CIN1/2 were followed up at 3 to 4 month intervals, in order to evaluate the risk of disease progression associated with HPV genotypes. Findings showed that the cumulative probability CIN1/2 for progressing to CIN3 within 5 years for HPV 16, 18, 31, 33, 35, 52 and 58 was 20.5% while the probability for other HR types was 6% and 1.7% for LR types (Matsumoto *et al*., 2011). Sandri and co-workers (2009) evaluated the prevalence and association of HR-HPV genotypes with cervical lesions of different severity in 199 women. Results showed an increase in the prevalence of HPV 16 infection from 21.3% in CIN1 to 71.9% in CIN3 (Sandri *et al*., 2009). Among women that were surgically treated, genotype specific persistence was evaluated in order to predict the development of residual or recurrent disease during follow-up. Results revealed that HPV 16 was the most common HPV type associated with recurrent disease, with a frequency of 55.6%. The prevalence of HPV 16 increased with severity of cervical disease, from a prevalence of 23.1% in CIN1 to 45.5% in CIN2 and eventually to 75.6% in CIN3 (Venturoli *et al*., 2008).

HPV 18 is the second most frequent HPV type in ICC (Guan *et al*., 2012), and is responsible for a fifth of cancer cases (Schiffman *et al*., 2011). HPV 18 was found to have an absolute risk of future CIN2+ of 25.6% (Naucler *et al*., 2007). However it has been reported that studies using CIN3 as the principal outcome end point for cervical cancer underestimate the carcinogenic potential of HPV18 compared to other HR types (Bulk *et al*., 2007; Guan *et al*., 2012). This is because HPV 18 varies very little between normal cytology and CIN3 but has been found to increase significantly in ICC (Guan *et al*., 2012).

From these studies it is clear that the risk of persistence and progression varies among HR-HPV types to cause cervical cancer (Matsumoto *et al*., 2011). Hence it is useful to test for type-specific HPV in women with LSIL in order to identify women who are at increased risk of disease progression (Clifford *et al*., 2005a; Matsumoto *et al*., 2011; Sandri *et al*., 2009; Wheeler *et*

al., 2006). HPV typing is able to distinguish between HPV positive women with HR-HPV from those women with LR-HPV. Furthermore women who are positive for HPV 16 and / or HPV18 may be subjected to an intensive follow up since they have an elevated risk of CIN3+ as oppose to HR-HPV positive women who test negative for those two types (Meijer *et al*., 2006).

1.10 Age-specific prevalence of HPV and geographic variation of HPV prevalence in young women

Genital HPV infection is highly prevalent in young women that are sexually active (Ho *et al*., 1998; Iftner *et al*., 2010; Smith *et al* 2008). Overall HPV prevalence varies by age and country (Smith *et al*., 2008). However in most regions of the world HPV prevalence infection increases from 14 years through 25 years then decrease with increasing age thereafter (de Sanjose *et al*., 2007; Dunne *et al*., 2007; Kjaer *et al*., 2008; Smith *et al* 2008). There is a notable difference in HPV prevalence in middle aged women between the ages 35 to 50 years. In middle aged women, HPV prevalence was found highest (20%) in African women and Central/ South American women compared to women from Asia/Australia, Southern Europe/Middle East and Northern Europe with approximately 15% HPV prevalence (Smith *et al*., 2008).

In a worldwide meta-analysis study in which the age and genotype specific prevalence of cervical HPV was estimated in 157 879 women with normal cytology reported that approximately 10.4% women have HPV infection at any given point in time (de Sanjose *et al*., 2007). Overall distribution of HPV types in descending order was; HPV 16 (2.5%), HPV18 (0.9%), HPV 31 (0.7%), HPV 58 (0.6%) and HPV 52 (0.6%). HPV prevalence was highest (15.5%) in women from less developed countries compared to 10.0% in women from more developed countries (Figure 1.13) (de Sanjose *et al*., 2007).

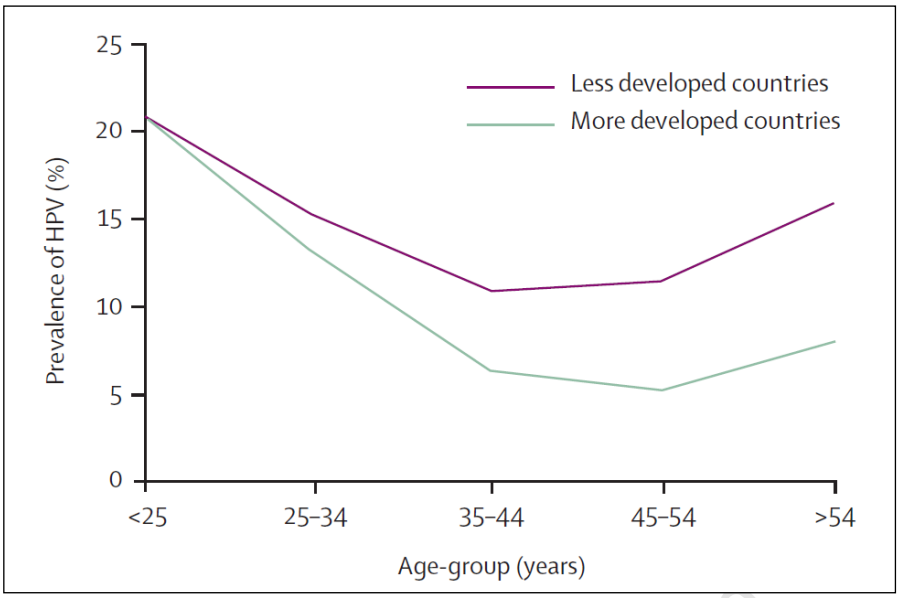


Figure 1.13 World estimates of age-specific HPV prevalence according to development status of specific country (de Sanjose *et al*., 2007).

A distinct geographic pattern which showed a second peak in women over the age of 44 years was observed in women from Africa, Northern America, Central and South America and Europe but not seen in Asian women (Figure 1.14) (de Sanjose *et al*., 2007).

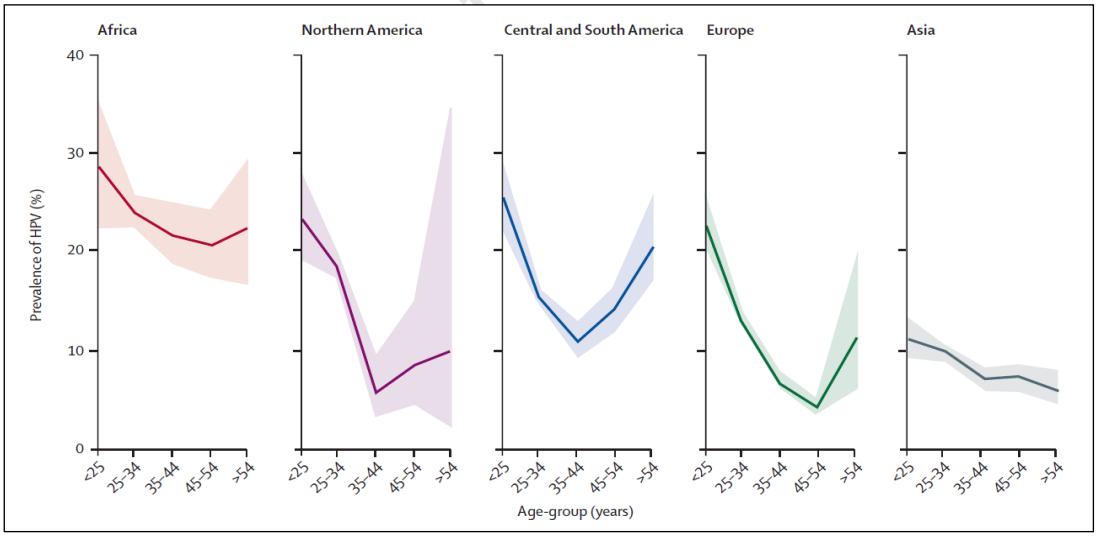


Figure 1.14 HPV prevalence stratified according age among women from various world regions with normal cytology (de Sanjose *et al*., 2007).

The highest prevalence of HPV was for African women with 22.1% when compared with women from other world regions (de Sanjose *et al*., 2007). A recent meta-analysis study of cervical HPV prevalence in 1 016 719 women with normal cytology revealed global HPV prevalence was 11.7%. Overall

HPV prevalence in descending order was as following; HPV 16 (3.2%), HPV 18 (1.4%), HPV 52 (0.9%), HPV 31 (0.8%) and HPV 58 (0.7%). HPV prevalence was slightly higher (11.8%) in women from less developed countries compared to 11.3% in women from more developed countries. HPV prevalence was highest in all regions in women younger than 25 years of age and declined thereafter at middle age. A second peak was observed at age 40 years or older and was clearly conspicuous in women older than 55 years in Africa. A second peak was not clearly pronounced in women from the Asian continent (Bruni *et al*., 2010).

In a South African study HPV prevalence was 20.4% (173/848) in women with normal cytology aged 21-59 years old. In descending order, the most frequently detected HR-HPV types were, HPV 16 (2.00%), HPV 52, (1.9%), HPV 45 (1.4%) and HPVs 18 and 68 (1.3%) (Allan *et al*., 2008). In another South African study of the prevalence of HR-HPV and distribution among HIV negative women aged 17-65 years with and without cervical intraepithelial neoplasia, 71.5% (902/1261) of women with normal cytology were HPV positive. HPVs 35, 16 and 58 were the three most frequently detected HR-HPV types in women with normal cytology (McDonald *et al*., 2012).

Since my M. Sc thesis investigated HPV prevalence in HIV negative women from Durban, South Africa aged 14-30 years, more detail is given reviewing HPV prevalence in the respective age bracket between 10 to 30 years of age. HPV prevalence trends among young women aged ≤30 years from various world regions are lowest among European women compared to North American and African women (Figure 1.15) (Castellsague *et al*., 2001; Healey *et al*., 2001; Iftner *et al*., 2010; Mbulawa *et al*., 2010; Silva *et al*., 2011). In a European study done by Silva and others (2011) in Portugal characterizing HPV infection status in 277 female adolescents aged 14 to 30 years old HPV was present in 46/277 (16.6%) women. The most common HPV types were HPV 31 (15.2%) followed by HPV 16 (13.0%) and finally HPV 53 and 61 both with a prevalence of 8.7% (Silva *et al*., 2011).

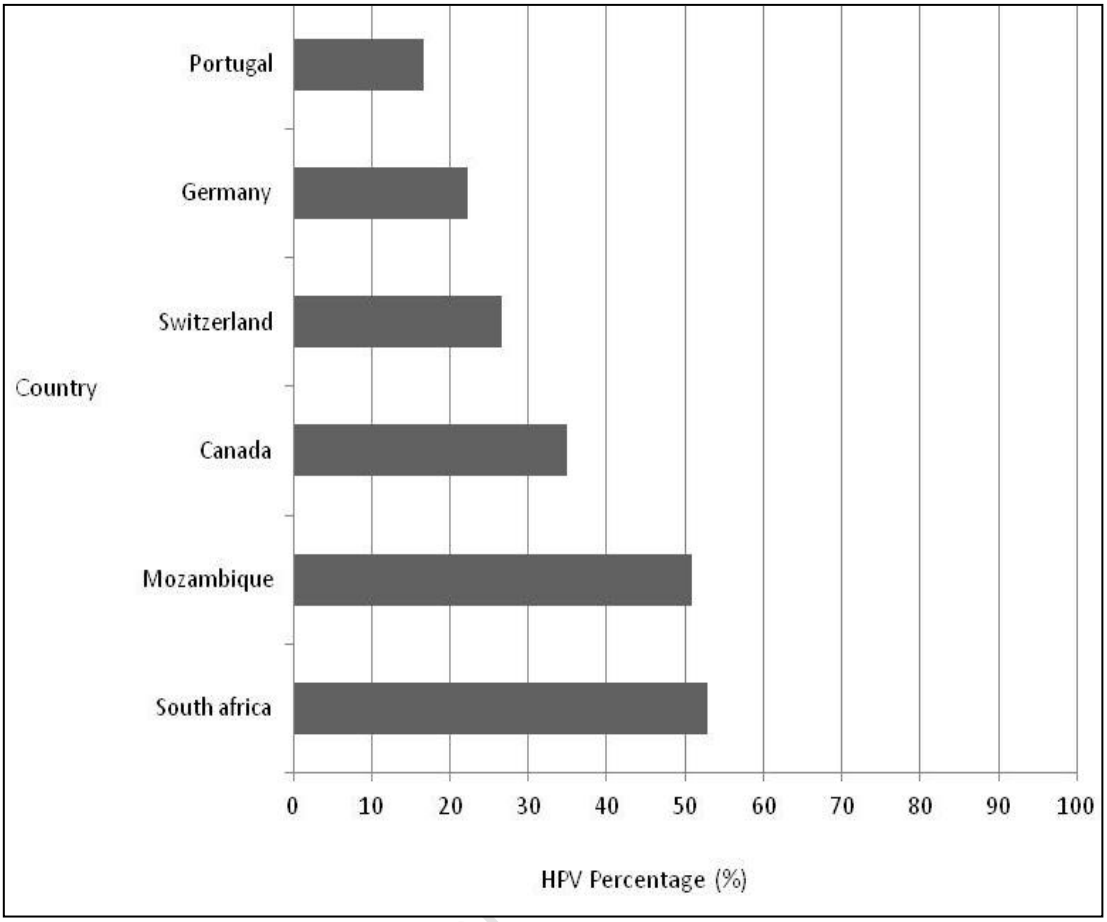


Figure 1.15 HPV prevalence variations in young women aged 10-30 years from various world regions (Castellsague *et al*., 2001; Dobec *et al*., 2009; Healey *et al*., 2001; Iftner *et al*., 2010; Mbulawa *et al*., 2010; Silva *et al*., 2011).

Among German females aged 10-30 years, HPV prevalence was 22.28% (377/1692). Females in the age group 20-22 years were found to have the highest prevalence of 28.3% (Iftner *et al*., 2010). In Switzerland HPV prevalence was 26.26% (47/179) among women aged ≥30 years. The prevalence of HPV was highest 28% (39/139) among women aged 21-30 years and prevalence of HR-HPV types was 15.1% (Dobec *et al*., 2009). The prevalence of HPV was determined among 2065 Dutch specimens obtained from women aged 18 to 29 years. Results showed that 393 (19%) women were HPV positive. The prevalence of HR-HPV types was 11.8% and 9.1% for LR types. The most prevalent HPV detected were HPV 16 (2.8%), HPV 51 (2.5%), and HPV 52 (2.5%) (Lenselink *et al*., 2008). In Italy, the HPV status of 1006 women aged 18-24 years was determined and findings showed that a total of 243 (24.1%) women were HPV positive, and out of the

243 HPV positive women, 175 (72%) women had HR-HPV types. The most common HPV type was HPV 16 with a prevalence of 4.5% followed by HPV 53 (2.7%), HPV 84 (2.6%), HPV 42 (2.5%), HPV 62 (2.4%) and lastly HPV 66 and HPV 89 both with a prevalence of 2.2% respectively (Ammatuna *et al*., 2008). Whereas among North American aboriginal women in Canada aged 13-30 years the prevalence of HPV was 34.86% (251/720). In this Canadian study the prevalence of HPV was 42.08% (101/240) in women aged 13-20 then declined to 31.25% (150/480) among women aged 21-30 years (Healey *et al*., 2001).

In South America, HPV was determined in 206 virgins from Costa Rica aged 18-26 years who were followed semi-annually for a median of 3.6 years after being sexually active. Results showed a cumulative HPV DNA detection of 110/206 (53.4%) women testing positive for at least one infection at least once. Out of 110 HPV positive women, the identity of type specific HPV genotypes was possible in 105 women. The number of women which tested positive once for HR types were 79 (38.4%) while 80 (38.8%) tested positive once for LR types. HPV 16 was the most common HR types with a prevalence of 19 (9.2%) followed by HPV 66 with a prevalence of 16 (7.8%) and HPV 52 with a prevalence of 15 (7.3%). Among LR-HPV types, HPV 53 was the most common with a prevalence of 19 (9.2%) (Rodriguez *et al*., 2007).

Among Asian women, the prevalence of HPV was determined in 1039 young Japanese women aged 20 to 25 years old at baseline. Their findings revealed that 355 (34.2%) women were HPV positive. HR-HPV types were detected in 304 (29.3%) and LR types were detect in 112 (10.8%). The most frequently detected HPV types was HPV 52 (8.1%) followed by HPV 16 (6.5%), HPV 51 (4.5%), HPV 18 (4.0%) and HPV 31 (3.8%) (Konno *et al*., 2011).

The prevalence of HPV has been reported to be highest among African women, with prevalence rates of 50.89% (57/112) and 52.78% (38/72) in

Mozambican and South African women aged ≤30 years respectively (Castellsague *et al*., 2001; Mbulawa *et al*., 2010), and may be influenced by HIV. High risk HPV prevalence was 55% in women aged 14-20 years and 48% in women aged 21-30 years from Mozambique (Castellsague *et al*., 2001). The overall prevalence of HR-HPV types in South African women aged ≤30 years was 62.30% (38/61). The most frequently detected HR-HPV types were HPV 53 (9.84%), HPV 35 (8.20%) and HPV 16, 18, 31 and 58 (6.56%) and the most frequently detected LR-HPV were HPVs 61, 62 and 70 (6.56%) (Mbulawa *et al*., 2010). In a recent African study done in a general population of Tanzanian women under the age of 25 (n=260) the prevalence of HR-HPV types was determined to be 24.6%. The most common HPV types were HPV 52 (6.5%), HPV 16 (5.4%), HPV 66 (4.2%), HPV 35 (3.8%) and HPV 56 (2.7%). Among 624 women in the age group of 25-29 years; the prevalence of HR-HPV types was 27.6%. The most common HPV types were HPV 52 (5.8%), HPV 51 (5.0%), HPV 16 (4.0%), HPV 35 (3.7%) and HPV 66 (3.5%) (Dartell *et al*., 2012).

In conclusion, results from the above studies indicate that genital HPV infections vary across different age groups in young women across geographical regions. This is due to differences in sexual behaviour, method used in selecting the study population and heterogeneity of HPV DNA detection techniques employed by various laboratories.

1.11 HPV sharing in couples

HPV is efficiently transmitted between sexual partners; consequently multiple transmissions within couples are also possible (Hernandez *et al*., 2008). HPV prevalence and concordance among couples depends on the couples' sexual behaviour, sensitivity of test used and the differences between acquisition rates among women and men (Parada *et al*., 2010). Furthermore, the natural history of HPV infections between men and women varies because of the differences between the epithelium of the penis to that of the cervical transformation zone (Parada *et al*., 2010).

HPV transmission occurs more frequently in sexually active couples in a monogamous relationship that use non-barrier forms of contraception compared to partners in transient relationships (Hernandez *et al*., 2008; Widdice *et al*., 2010). This may be due to the fact that partners in transient relationship have less time to transmit HPV infections (Widdice *et al*., 2010). The transmission rate of HPV is higher from female to male than male to female (Widdice *et al*., 2013).

A recent study examined HPV concordance and transmission rates in 25 heterosexual couples with a relationship of at least 3 months' duration in which women were aged 13- 21 and the respective male partners were ≥18 years from California. HPV DNA from genitals was sampled 5 times over a 6 weeks period. Results of the study showed genital HPV concordance between the couples ranged from 64 to 95% for at least 1 HPV type. Transmission rates from the male genital to female genital ranged from 14.5-100 per 100 person-months and for female genital to male genital ranged from 26.8- 187.5 per 100 person-months (Widdice *et al*., 2013). HPV transmission rates were found lower in a Hawaiian study in which HPV transmission was examined in 25 heterosexual monogamous couples followed up at 2 months interval for an average period of 7.5 months. The mean age for men was 26 years and the mean age for women was 28 years. Finding showed that male to female transmission occurred in 7 couples and female to male transmission occurred in 12 couples. Infections transmitted from the male to the female all originated in the penis and the overall transmission rate was 4.9 per 100 person-month of exposure. By contrast infections transmission from the cervix and/ or urine to penis consisted of female to male events and the overall transmission rate was 17.4 per 100 person-months of exposure (Hernandez *et al*., 2008). HPV transmission rates were lowest among Canadian couples in which women were aged 18-24 years and their respective male partners were ≥ 18 years old. Transmission rates were report to be 3.5 per 100 person-months for male to female and 4.0 per 100 person-months for female to male transmission. The duration of the relationship was ≤ 6 months and HPV transmission was observed in 40.78%

(73/179), and which 83.6% (61/73) involved transmission of single HPV types. HPV transmission of 2 types was observed in 13.7 % (10/73) couples and 2.7% (2/73) involved the transmission of 3 types (Burchell *et al*., 2011).

Among African couples, HPV concordance was found in 15.7% couples for any HPV types and 8.4% for high-risk HPV types among 166 Rwandan couples. The median age of the men was 31 years and the median age for women was 28 years (Veldhuijzen *et al*., 2012). HPV concordance was determined in 486 heterosexually active black South African couples with a mean age of 38 years for men and 35 years for women. A total of 10% of HIV negative couples had concordant HPV types while 48% of HIV positive couples shared HPV types. In HIV discordant couples in which the male partner was HIV positive, 11% were HPV concordant while in HIV discordant couples where the female partner was HIV positive, 30% shared HPV types (Mbulawa *et al*., 2010).

HPV prevalence and concordance among couples from the above studies vary and are influenced by many factors such as the couple's sexual behaviour, HIV status and sensitivity of test employed. Investigating HPV transmission is mostly determined by genotyping for type specific concordance among partners in a couple. This approach is not precise and tends to overstate HPV transmission rate between couples. Therefore a better approach would be to investigate the transmission of HPV by identifying HPV variants through sequencing, which will enable the investigator to determine HPV transmission rates with a higher level of precision (Ho *et al*., 1993a; Lee *et al*., 2009). In the only study in which HPV transmission among respective partners in couples was determined by evaluating variant concordance, findings revealed that eight couples were HPV 16 positive; however HPV 16 variants were only shared in four out of eight couples. Partners in the other four couples had discordant HPV variants (Ho *et al*., 1993a).

1.12 HPV vaccines

Currently, there are two prophylactic HPV vaccines available; one was developed by Merck termed Gardasil and the other developed by GlaxoSmithKline termed Cervarix (Verheijen, 2011). These vaccines consist of virus like particles (VLP) derived from L1 major capsid proteins (Figure 1.16) (Chen *et al*., 2000). The L1 major structural protein is used to produce vaccines (De la Rosa *et al*., 2009), since this protein has the ability to self-assemble into VLPs in the absence of other papillomavirus proteins. The VLPs consists of 72 pentamers of L1 protein and are morphologically similar to native virions, but contain no DNA and (Chen *et al*., 2000). Although VLP are not infectious, they have the ability to induce neutralizing antibodies (Kirnbauer *et al*., 1992).

Cervarix contains VLPs assembled from the L1 proteins of HPV 16 and HPV 18 (Deschuyteneer *et al*., 2010). Gardasil also contains L1 VLPs from HPV 16 and HPV 18 but in addition contains L1 VLPs derived from HPV 6 and HPV 11 (Donavan *et al*., 2011; Shank-Retzlaff *et al*., 2006). Together HPV 6 and HPV 11 are responsible for 90% of genital warts (Aubin *et al*., 2008).

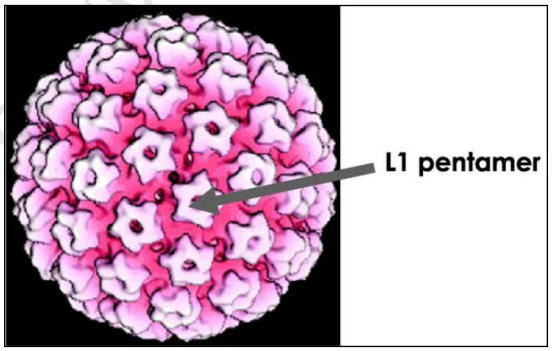


Figure 1.16 HPV L1 virus like particle (VLP) structure (Stanley *et al*., 2006)

Both vaccines are administered intramuscularly on three separate occasions, with Gardasil and Cervarix having 100% and 98% efficacy against cervical diseases caused by HPV 16 and 18, respectively (Table 1.6) (Botha, 2009). Cross protection against any lesions caused by non-vaccine oncogenic HPV types is 47.7% and 23.4% for Cervarix and Gardasil respectively (Brown *et al*., 2009; Jit *et al*., 2011). It is recommended that HPV vaccine should be

given to young children in the public health domain before sexual debut to decrease the burden associated with HPV (Moscicki, 2008), with Gardasil being the better option since it is the only registered vaccine for boys (CDC, 2011; Georgousakis *et al*., 2012). The efficacy of Gardasil against HPV 16, 18, 6 and 11 related external genital lesions in males aged 16-26 years was reported to be 79.0% among men who had sex with men and 92.4% among heterosexual men (Giuliano *et al*., 2011).

Table 1.6 Commercially available vaccines (modified from Botha, 2009)

Vaccine	**Gardasil**	**Cervarix**
HPV types	6, 11, 16, 18	16, 18
Dosing schedule	0, 2, 6 months	0, 1, 6 months
Adjuvant	AAHS (amorphous aluminium hydroxyphosphate sulphate)	ASO_4 (aluminium hydroxide + MPL)
VLP preparation system	Yeast	Insect cell
Efficacy against disease caused by HPV types in the vaccine	100%	98.1%
Efficacy against non-vaccine oncogenic HPV type	23.4%	47.7%

In a recent study done by Draper et al (2011) in which the frequency and titre of neutralizing antibodies was determined against pseudoviruses representing genetically related HPV types from the *Alpha* (α) genera, species 7 consisted of five HPV types; 18, 39, 45, 59 and 68, and species 9 consisted of six HPV types; 16, 31, 33, 35, 52 and 58. Determination of frequency and titre of neutralizing antibodies was done using sera from recently vaccinated young girls aged 13-14 years using the bivalent Cervarix vaccine. Results showed that, neutralizing antibodies against non-vaccine HPV types 31 and 33 from α9 were detected, while only neutralizing antibodies against non-vaccine HPV type 45 from α7 was detected (Draper *et al*., 2011).

Wheeler et al (2009) evaluated the impact of a quadrivalent HPV vaccine on infection related to 10 non-vaccine HPV types; 31, 33, 35, 39, 45, 51, 52, 56, 58 and 59. This was done to determine whether administration of the vaccine reduced the incidence of infection of ≥ 6 month's duration. The rate of HPVs; 31, 33, 45, 52 and 58 infections were reduced by 17.7% after vaccination.

The rate of HPV related CIN 1-3 was reduced by 18.8% and the rate of HPV types: 31, 58, 59 related CIN1 were reduced by 26%, 28.1% and 37% respectively (Wheeler *et al*., 2009). Therefore administration of quadrivalent HPV vaccine to partially infected individuals reduces the incidence of infection (Wheeler *et al*., 2009).

The studies by Draper *et al*., 2011 and Wheeler *et al*., 2009 both demonstrated the vaccines' ability to cross-protect against non-vaccine types, and may also be provided to women at risk for co-infection with both vaccine and non-vaccine HPV types (Draper *et al*., 2011; Wheeler et al., 2009). In a head to head encounter, Einstein and co-workers (2009) compared the immune response elicited by both vaccines using the same methodology for assessment of immune response and reactogenicity. They found that HPV 16 and 18 neutralising antibody levels induced by Cervarix at month 7 were higher than those induced by Gardasil. Neutralizing antibodies levels against HPV 16 were 2.3-4.8 fold higher in the Cervarix group compared to the Gardasil group after the completion of the vaccination course. The level of HPV 18 neutralizing antibody induced by the two vaccines was found to be 6.8-9.1- folds higher in Cervarix when compared to Gardasil. Individuals vaccinated with Cervarix produced a higher positivity rate of cervicovaginal secretion neutralizing antibody against HPV16 and HPV 18 (Einstein *et al*., 2009).

Epidemiological data on HPV prevalence and type distribution at baseline is crucial to evaluate the impact of HPV vaccines. Therefore, the data presented in this study serves as a starting point to aid in the understanding of the burden of HPV infection. Consequently, this information will be valuable for policy making regarding to HPV screening that would maximize the potential benefits of an HPV vaccine. Hence, we aim to establish cervical HPV type-specific prevalence, incidence and clearance rates in young rural and urban HIV negative woman who participated in an HIV prevalence trial preparedness study in KwaZulu Natal.

Chapter 2: A Study of HPV Types In Young South African Women

2.1 Introduction

Viral prevalence is the result of acquisition of new infections (incidence) and duration (persistence) (Castle *et al*., 2005). HPV prevalence varies by country, regions within the country and population subgroups due to different sexual behaviour and screening methods used (Bruni *et al*., 2010; Clifford *et al*., 2005b; Dartell *et al*., 2012; De Vuyst *et al*., 2010; Dols *et al*., 2012; Menendez *et al*., 2010; Michala *et al*., 2012; Smith *et al*., 2008; Ting *et al*., 2010; Thomas *et al*., 2004; Vinodhini *et al*., 2012).

Data on HPV acquisition, clearance and persistence is needed in different populations to effectively design cervical cancer prevention strategies (Datta *et al*., 2012; Rositch *et al*., 2012). HPV acquisition is age dependant and is most common among young women (Castle *et al*., 2005; Liaw *et al*., 2001; Syrjannen *et al*., 2004; Thomas *et al*., 2000). Women younger than 35 years accumulate incident HPV infections more frequently than older women. The rate of acquisition is reported to be 3.7% per month in women younger than 20 years (Syrjannen *et al*., 2004).

Women 15-29 of age are more likely to clear HPV infections when compared to women older than 30 years of age (Safaeian *et al*., 2008). It is known that HPV infections clear within 6 months (Goodman *et al*., 2008), however it has been found that in majority of women, 80% of HPV infections are cleared within 18 months (Oh *et al*., 2008). LR-HPV types clear more rapidly compared to HR-HPV types (Datta *et al*., 2012; Goodman *et al*., 2008). The rate of clearance among LR-HPV types is highest for HPV 11, HPV 61, HPV 72 and HPV 84 (Datta *et al*., 2012). Among the HR-HPV types, HPV 16 is less likely to clear compared to other HR-HPV types (Bulkmans, *et al*., 2007).

Clearance rates are lower for HR-HPV types and consequently persistence rates are higher for HR types compared to LR types (Datta *et al*., 2012; Rositch *et al*., 2012). It has been reported that HPV 16, 31, 33 and HPV 52 are the most persistent HPV types (Rositch *et al*., 2012).

A persistent HR-HPV type is necessary for the development of cervical cancer (Franco *et al*., 1999; Chen *et al*., 2011b). Although HPV persistence rate varies across different studies and is largely influenced by region of study and HPV type, analyzing the dynamics of HPV clearance against HPV persistence are useful in understanding the natural history of HPV. This will in turn aid to the prediction of whether a SIL will regress spontaneously or will persist and progress to cause HSIL and eventually cervical cancer (Datta *et al*., 2012; Fukuchi *et al*., 2009; Rositch *et al*., 2012). However an obstacle in determining such dynamics is that there are limited studies available for specific age groups (Rositch *et al*., 2012).

2.2 The aims of the study

(i) To investigate the prevalence of genital HPV infection in HIV-negative women aged 14-30 years in KwaZulu Natal

(ii) To investigate HPV acquisition, clearance and persistence

(iii) To examined the prevalence of HPV and type distribution according to cervical cytology

2.3 Materials and methods

2.3.1 Study population

Cervico-vaginal lavages were obtained from 434 HIV negative young women aged 14-30 years who were enrolled in the CAPRISA sero-incidence clinical-trial preparedness study during 2005-2007. These participants were assessed for STI and received HIV testing in addition to HIV counseling on a three monthly basis for up to 13 months. There were 223 women at baseline, 106 women at second visit (defined as 2-4 months after enrollment) and 69 women at third visit (defined as 5-7 months after enrollment). There was an additional of 27 women at fourth visit (defined as 8-10 months after enrollment) and 9 women at fifth visit (defined as 11-13 months after enrollment).

During enrollment these women also received cytological assessment of the cervix by conventional Papanicolaou (Pap) smear and results were provided to the respective participants during the study. Specimens were collected with voluntary informed consent and maintained in storage at -85°C. The University of KwaZulu Natal Biomedical Research Ethics committee reviewed and approved the protocol entitled “HIV and STI incidence rates in Vulindlela: A prevention preparedness study. Q Abdool Karim, CAPRISA” (Ref: E197/03). This study was also approved by the Research Ethics Committee of the University of Cape Town (Ref: 459/2011).

2.3.2 Preparation of cervico-vaginal lavages (CVL)

A plastic bulb pipette was cut below the bulb and the tip was fastend to a 10ml syringe then 5ml of phosphate buffered saline (PBS) was drawn. A warm water lubricated speculum was inserted into the vagina and the pipette was introduced through the speculum into the vagina. The pipette was directed towards the cervical os and squeezed to bath the cervix. The fluid was allowed to pool into the posterior fornix and aspirated into the same pipette. The procedure was repeated 3 times using the same fluid which was

then dispensed into a sterile 30ml urine container. The CVL samples were stored on ice and transported on ice to the laboratory in less than 6 hours. The CVL specifmen were then aseptically transferred into a 15ml screw-capped centrifuge tube. The specimens were centrifuged at 800g for 10 minutes to fractionate the cellular component from the supernatant. The supernatant was aliquoted into cryovials and stored at -85°C.

2.3.3 DNA extraction

DNA from cervico-vaginal lavage (CVL) cell pellet was extracted for HPV typing using the Magna Pure Compact Nucleic Acid Isolation Kit I (Roche Diagnostics, Mannheim, Germany). This kit was used in conjunction with the MagNA Pure Compact Instrument (Roche Diagnostics, Mannheim, Germany) employing an automated extraction procedure performed according to the manufacturer's instruction. The procedure utilizes; reagent cartridges, tip trays and elution tubes, which were all placed into their respective positions in the MagNA Pure Compact Instrument (Roche Diagnostics, Mannheim, Germany). Total volume of 400µl cervical samples in 2ml tubes were loaded onto the sample tube rack. Sealed reagents cartridges containing the following reagents; Proteinase K, lysis buffer, magnetic glass particles (MGPs), isopropanol, wash buffer I and II, and elution buffer were used. Lysis buffer containing guanidine salts was used to lyse the cells, Proteinase K and the salts were used to denature contaminating proteins, including nucleases. The DNA was immobilized onto the MGPs and then precipitated by the addition of isopropanol. The denatured protein and all other contaminants were then removed by washing with buffers containing absolute ethanol. DNA was released from the beads at 80°C and was eluted into a sterile 2ml tube in 200µl of elution buffer consisting of nuclease-free, sterilized water. The extracted DNA samples were then stored at -20°C until further use. DNA extraction was carried out in an area dedicated for nucleic acid extraction only, with the use of nuclease-free sterile tips and tubes. These precautionary measures were conducted in order to eliminate contamination of DNA.

2.3.4 PCR amplification and HPV genotyping

The Roche Linear Array HPV Genotyping Test (Roche Diagnostics, Mannheim, Germany) was used for the detection of HPV types in specimens. This test was used because; it is very sensitive and specific. The test is suitable for high-throughput testing and can be used in large epidemiological studies. In addition the test has the ability to differentiate specific HPV types within the high and low risk groups, a feature which is not possible with the hybrid capture 2 test (Iftner and Villa, 2003). The hybrid capture 2 test is easy to perform in clinical settings and is suitable for automation. In addition it does not require special facilities to avoid cross-contamination since it does not rely on target amplification as compared the Roche Linear Array HPV Genotyping Assay to achieve high sensitivity. A limitation of the hybrid capture 2 test is that it only detects high risk HPV types (Iftner and Villa, 2003).

Preparation of specimens for PCR were carried out in a PCR-clean area, which is a room dedicated for making up PCR master-mixes. The precautionary measures were implemented to avoid contamination with amplified products. The amplification and detection of HPV types were done in a post-amplification area. DNA from clinical specimens was amplified using Roche Linear Array HPV genotyping and detection assay according to the manufacturer's instructions (Roche Diagnostics, Mannheim, Germany). This assay utilizes the PGMY09/11 primers used to amplify the conserved L1 region of 37 HPV genotypes in a single PCR. The reverse primers are biotin labelled at the 5' end to allow the capture of alkaline denatured amplicons onto the streptavidin coated strips during post-amplification detection.

HPV master mix contains the PGMY09/11 primers to amplify HPV DNA from 37 types resulting in a 450 bp product and GH20/PC04 primers which were concurrently used to amplify 268 bp of the β-globin gene. The human β-globin gene was co-amplified in the same PCR reaction to assess sample adequacy, DNA extraction and amplification efficiency. Furthermore, the

master mix also contains the enzymes AmpliTaq® Gold DNA polymerase and uracil-N-glycosylase (AmpErase); dNTPs including dUTPs. A Master Mix reagent in a total volume of 580µl was then mixed with 125µl of Magnesium chloride solution to give a working HPV master mix. The Mg^{2+} enhances the activity of AmpliTaq® Gold DNA polymerase. This enzyme extended the primers bound to target DNA for the amplification of HPV and human DNA. The AmpErase enzyme was used to eliminate carry-over contamination in PCR by catalyzing destruction of amplicons containing dUTP. This was done to ensure selective amplification of the required target DNA only. Fifty microlitres (µl) of DNA sample was added to 50µl aliquots of working HPV master mix and amplified in the amplification and detection area.

Negative and positive controls from the kit were included in each amplification procedure. The negative control (50µl) has no HPV DNA while the positive control (50µl) has HPV 16 DNA. The post-amplification detection of 37 HPV types (HPV6, 11, 16, 18, 26, 31, 33, 35, 52, 58, 39, 40, 42, 45, 51, 53-56, 59, 61, 62, 64, 66-73, 81-84, IS39 and CP6108) was done using the Linear Array Detection Kit (Roche Diagnostics, Mannheim, Germany). IS39 is a subtype of HPV 82 and CP6108 is also known as HPV 89 (Clifford *et al*., 2005b).

The following PCR conditions were performed on the GeneAmp PCR System 9700 (Applied Biosytems) 96-well gold-plated machine at a ramp rate of 50%: 50°C for 2 min to activate AmpErase, 95°C for 9 min, followed by 40 cycles consisting of 95°C for 30 sec, 55°C for 1 min, 72°C for 1 min and a final hold at 72°C for 5 min. The reaction was kept at 72°C until the addition of an alkaline denaturation solution from Linear Array Detection Kit (Roche Diagnostics, Mannheim, Germany) to denature the amplicons and AmpErase.

The denatured amplicon mixture (75µl) were then transferred to appropriate wells in the typing tray containing warm hybridization buffer and a single

Linear Array HPV Genotyping strip which were included in the kit. The strip was coated with HPV and β-globin oligonucleotide probes immobilized onto membrane. Each amplicon hybridized to the strip at 53°C and was bound to the complementary probe with a matching sequence. There was one cross-reactive oligonucleotide probe that hybridized with HPV 33, 35, 52 and 58. The Roche Linear Array HPV Genotyping test does not directly detect HPV 52; instead it combines a set of probes detecting HPV 33, 35, 52 and 58 as a group. The specimens that test negative for HPV 33, 35 and 58 individually but test positive for the group will be deemed to be HPV 52 positive. Specimens that test positive for the group and for HPV 33 or 35 or 58 with uncertain HPV 52 status will be deemed HPV 52 negative, however HPV 52 infection cannot be ruled out. Following hybridization and stringent washes with buffers containing SDS and sodium salts, streptavidin-horseradish peroxidase conjugate (SHPC) was added to the strip to bind to the biotin-labeled amplicons hybridized to the probes on strip. A substrate mixture of hydrogen peroxide (H_20_2) and 3,3',5,5'-tetramethylbenzine (TMB) was added to each strip. The TMB substrate oxidised to a blue color by catalysis action of streptavidin-horseradish peroxidase in the presence of H_20_2. The blue colour then precipitated at the probe positions and allow for the reading of strips.

2.3.5 Statistical analysis

X^2 test (Epi Info version 5 Statcalc) was used in comparing prevalence between groups. P values ≤0.05 were considered significant.

2.4 Results

2.4.1 Description of the Roche Linear Array HPV Genotyping Test

Each run included both the negative and positive control to ensure good quality of results. To ensure that there is no cross contamination between specimens in each run an empty well between specimens was left. Figure 2.1A demonstrates an example from 10 specimens showing positive

hybridization signals for HPV types as well as positive and negative control. The positive and negative controls were correctly determined therefore the run was successful. Samples; 172, 175, 183 and 186 were negative for HPV while samples 178 and 185 had only one type and specimens 174, 179, 180, and 184 have multiple HPV types. In this run we did not observe similar pattern of HPV types detected between specimens, therefore these findings indicate no evidence of cross contamination between specimens in a run. Figure 2.1B illustrates how the Linear Array HPV genotyping test reference guide was used to determine HPV types. In this sample both β-globin high and β-globin low test were positive as well as for HPV 18, 31, 35, 45, 53, 59 and 83.

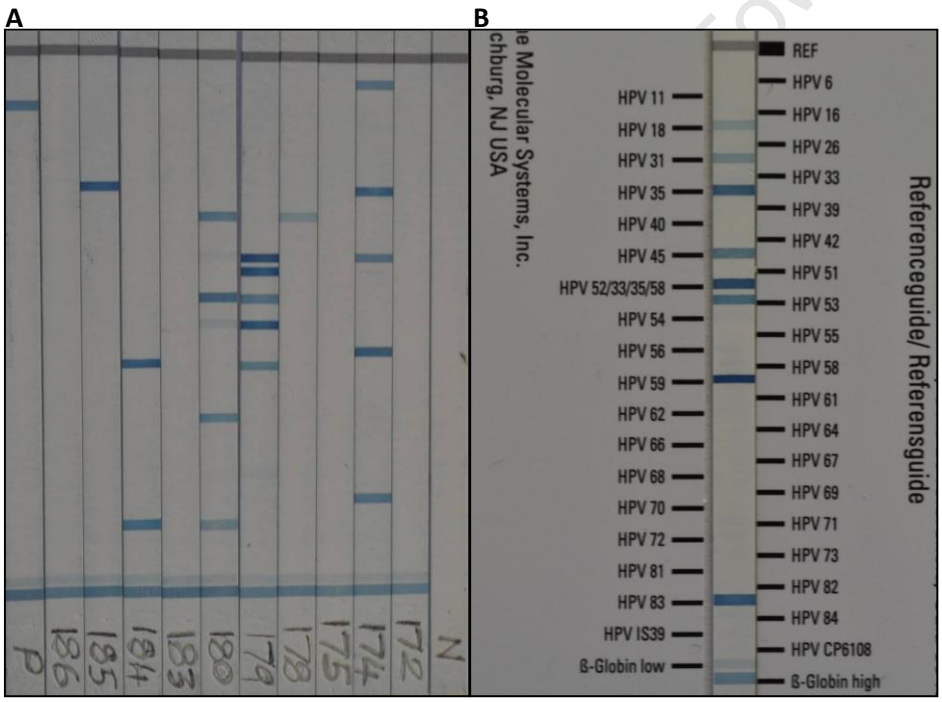


Figure 2.1A Roche linear array HPV genotyping strips demonstrating multiple HPV types without cross contamination. **(A)** Demonstrates results from 10 patients showing positive hybridization signals for HPV types as well as positive controls. Each run included both the negative (N) and positive (P) control to ensure that each run was successful. Samples 172, 175, 183 and 186 were all HPV negative while samples 178 and 185 had only one type and specimens 174, 179, 180, and 184 were all HPV positive with have multiple HPV types. (**B)** Roche linear array HPV genotyping strip demonstrating positive results with multiple HPV types (i.e.) 18, 31, 35, 45, 53, 59 and 83 as well as both β-globin high and β-globin low.

2.4.2 HPV prevalence at baseline

Overall baseline HPV prevalence in this cohort was 67.26% (150/223). The prevalence of single infections was 36% (54/150) while 64% (96/150) of specimens harbored multiple infections. Out of all infections observed, the prevalence of HR-HPV was 57.18% (219/383) and prevalence of LR-HPV was 42.82% (164/383). The most frequently detected HPV types were HPVs 58, 62 (14.67%); HPV 61 (14.0%); HPV 51 (12.67%); HPVs 16, 52 (11.33%); HPV 53 (10.67%); HPV 68 (10.00%); HPVs 6, 39, 45 (9.33%). The least prevalent HPV types were; HPVs 11, 67, 69 (1.33%); HPVs 40, 82 (2.00%); HPVs 71, 89 (2.67%); HPV 26 (3.33%); HPVs 31, 56, 70 (4. 00%); HPV 42, IS39 (4.67%) and HPV 33 (5.33%) (Figure 2.2).

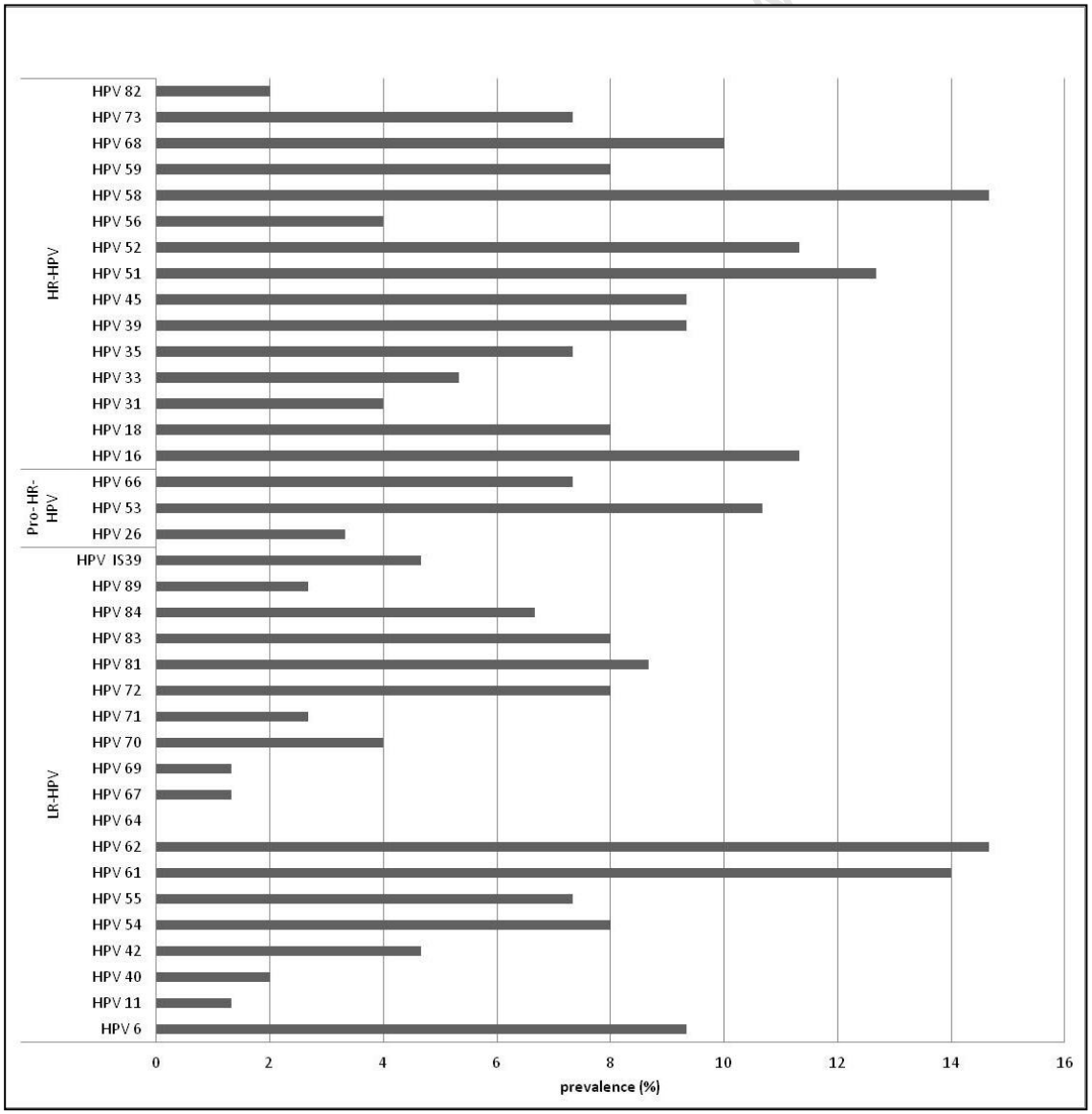


Figure 2.2 Baseline prevalence of genital HPV types in HIV- negative young women

From the 223 positively genotyped specimens, 46.64% (104/223) of specimens were from women aged 14-20 years and 53.36% (119/223) were from women aged 21-30 years. HPV prevalence according to age group was 69.23% (72/104) in women aged 14-20 years and 65.55% (78/119) in women aged 21-30 years.

2.4.3 HPV prevalence at second visit

At the second visit defined as 2-4 months after enrollment, the prevalence of HPV was 65.09% (69/106). The prevalence of single HPV infections was 27.54% (19/69) and the prevalence of multiple HPV infections was 72.46% (50/69). The prevalence of HR-HPV type was 62.87% (105/167) and LR-HPV prevalence was 37.13% (62/167). The most frequently detected HPV types were HPV 58 (21.74%); HPV 59 (15.94%); HPVs 16, 35 (14.49%); HPV 62 (11.59%); HPVs 39, 51, 52, 54, 55 (10.14%); HPVs 18, 61, 70, 73 (8.70%) and HPVs 45, 53, 72, 83 (7.25%). The least frequently detected HPV types in ascending order were as following HPVs 11, 71 (1.45%); HPVs 6, 33, 42, 89, IS39 (2.90%); HPVs 26, 56 (4.35%) and HPVs 31, 66, 81, 84 (5.80%) (Figure 2.3).

2.4.4 HPV prevalence at third visit

The HPV prevalence at third visit defined as 5-7 months after enrollmnt was 60.87% (42/69). The prevalence of single HPV infections was 26.19% (11/42) and the prevalence of multiple HPV was 73.81% (31/42). HR-HPV prevalence was 55.12% (70/127) and prevalence of LR-HPV types was 44.88% (57/127). The most frequently detected HPV types were HPVs 58, 61 (19.05); HPV 51 (16.67%); HPVs 18, 35, 52, 54, 59, 62 (14.29); HPVs 6, 16, 26, 45 (11.90%); HPVs 53, 56, 70, 89 (9.52%); HPVs 69, 71, 83, 84 and IS39 (7.14%). The least frequently detected HPV types in ascending order were HPV 55 (2.38%) and HPVs 11, 31, 39, 42, 66, 72, 73, 81 (4.76%) (Figure 2.4).

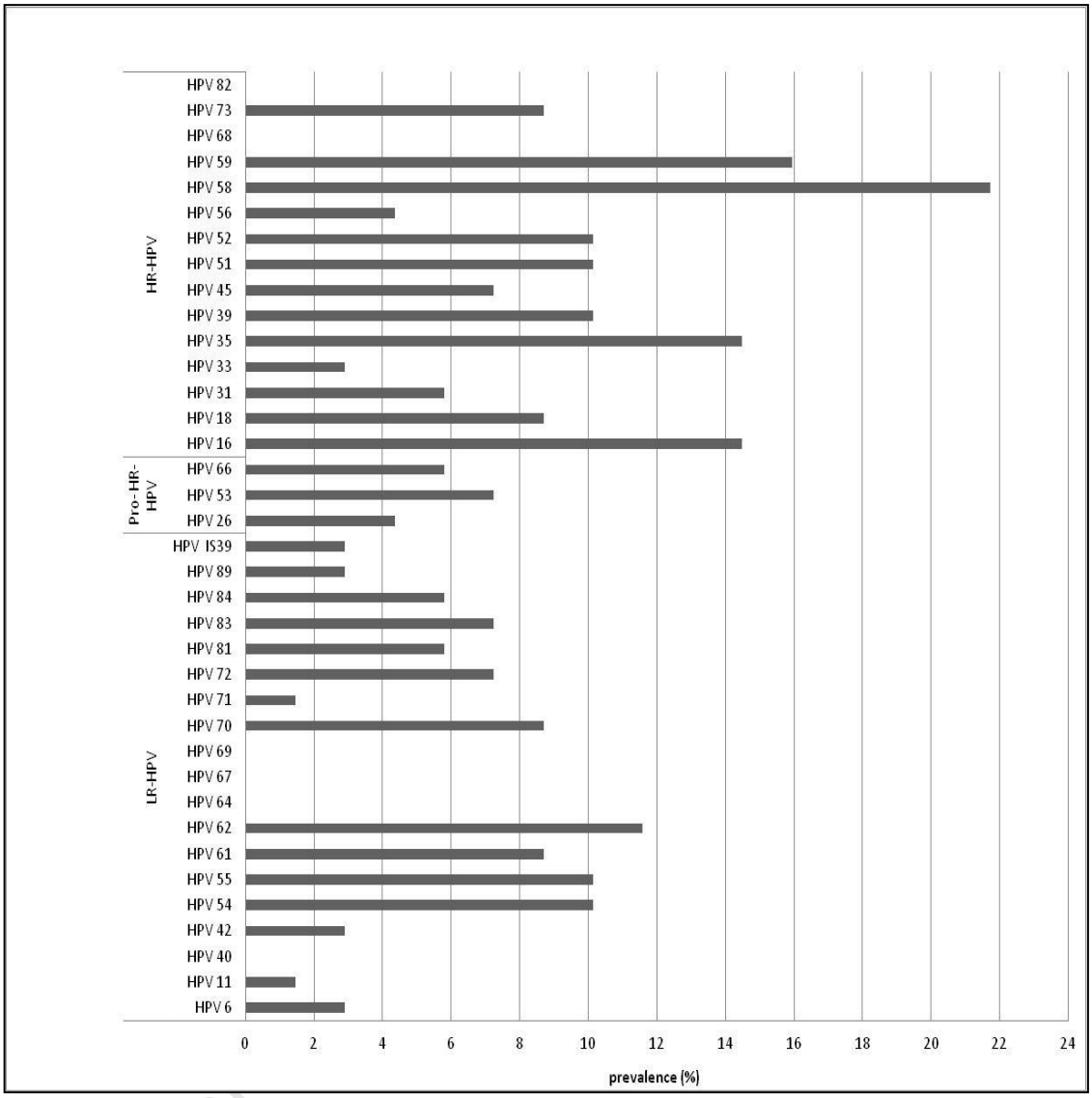


Figure 2.3 HPV prevalence in HIV- negative young women at 2-4 months after first visit

2.4.5 HPV prevalence at fourth visit

There were very few (27 women) followed up to fourth visit defined as 8-10 months after enrollment. HPV prevalence was 59.26% (16/27) and the prevalence of multiple HPV types was 68.75% (11/16) while that of single HPV types was 31.25% (5/16). HR-HPV prevalence was 62.79% (27/43) and the prevalence of LR-HPV types was 37.21% (16/43). The most commonly detected HPV types were HPV 18 (18.60%), HPV 62 (9.30%), and HPVs 51 and 58 with 6.98% respectively.

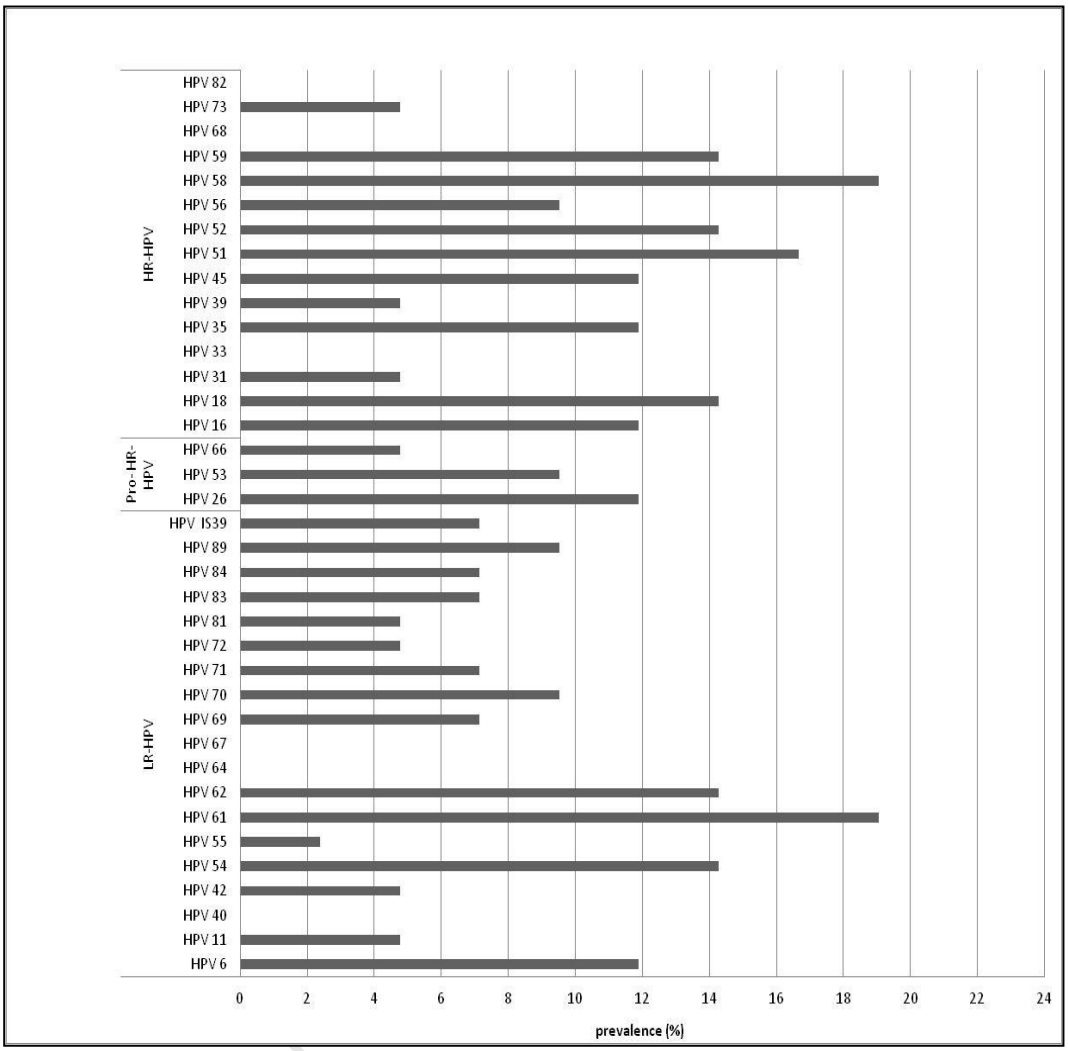


Figure 2.4 HPV prevalence in HIV- negative young women at 5-7 months after first visit

2.4.6 HPV prevalence at fifth visit

HPV prevalence was 66.67% (6/9) at fifth visit defined as 11-13 months after enrollment. The prevalence of both multiple and single HPV types was 50% (3/6). HR-HPV prevalence was 46.15% (6/13) and that of LR-HPV was 53.85% (7/13). The most commonly detected HPV types were HPVs 53 and 70 (15.38%).

2.4.7 Prevalence of HPV stratified according to species

When HPV types were grouped according to species, at baseline the overall prevalence of HPV in descending order is as following; α3 HPV 62.67%

(94/150), α9 HPV 55.33% (83/150), α7 HPV 48.67% (73/150), α5 HPV 24% (36/150), α6 HPV 22% (33/150), α10 18% (27/150), α13 HPV 8.00% (12/150), α11 7.33% (11/150), α1 HPV 4.67% (7/150), α14 HPV 2.67% (4/150) and α8 HPV 2.00% (3/150). Single infections in α3, α7 and α9 HPV species were observed in 11.70% (11/94), 10.96% (8/73) and 18.07% (15/83) respectively. The trends observed during follow up were increasingly significant for α7 HPV species ($R^{2=}$ 0.966) and α13 HPV species ($R^{2=}$ 0.967) (Table 2.1).

Table 2.1 Prevalence of HPV species at different visits during follow-up

	Baseline visit	**Second visit**	**Third visit**	**R^2 Value**
α1 HPV species	4.67	2.9	4.76	0.002
α3 HPV species	62.67	49.28	66.67	0.048
α5 HPV species	24.00	17.39	42.86	0.509
α6 HPV species	22	17.39	23.81	0.075
α7 HPV species	**48.67**	**50.72**	**54.76**	**0.966**
α8 HPV species	2.00	0.00	0.00	0.750
α9 HPV species	55.33	69.57	64.29	0.387
α10 HPV species	18.00	14.49	19.05	0.048
α11 HPV species	7.33	8.70	4.76	0.413
α13 HPV species	**8.00**	**10.14**	**14.29**	**0.967**
α14 HPV species	2.67	1.45	7.14	0.557

2.4.8 HPV incidence, persistence and clearance

HPV acquisition was defined as the detection of a new HPV type that was not initially detected at previous visit. Acquisition of new HPV infection was 46.23% (49/106) at second visit and of which 67.09% (53/79) were HR-HPV types and 32.91% (26/79) were LR-HPV types. The most frequently acquired HPV type was HPV 59 with a frequency of 10.13% (8/79) and the most frequently acquired HPV species was α9 with a frequency of 30.38% (24/79) (Table 2.2).

Table 2.2 HPV incidence, persistence and clearance at different visits during follow-up

	Second visit				Third visit			
	Overall Percentage	**Frequency of HPV group**	**Most frequent HPV type**	**Most frequent HPV species**	**Overall Percentage**	**Frequency of HPV group**	**Most frequent HPV type**	**Most frequent HPV species**
New HPV infections	46.23% (49/106)	HR-HPV 67.09% (53/79) LR-HPV 32.91% (26/79)	HPV 59 10.13% (8/79)	α9 30.38% (24/79)	43.48% (30/69)	HR-HPV 40.79% (32/63) LR-HPV 49.21% (31/63)	HPVs; 18, 51, 52, 53, 54, 56, 59, 69, 89 4.76% (3/63)	α3 19.05% (12/63)
HPV Clearance	41.51% (44/106)	HR-HPV 50.63% (40/79) LR-HPV 49.37% (39/79)	HPV 6 7.59% (6/79)	α3 22.78% (18/79)	42.03% (29/69)	HR-HPV 64% (32/50) LR-HPV 36% (18/50)	HPV 39 12.00% (6/50)	α7 36.00% (18/50)
Persistent HPV infections	48.11% (51/106)	HR-HPV 62.37% (58/93) LR-HPV 37.63% (35/93)	HPV 58 8.60% (8/93)	α9 24.73% (23/93)	46.38% (32/69)	HR-HPV 59.68% (37/62) LR-HPV 40.32% (25/62)	HPV 58 11.29% (7/62)	α9 27.42% (17/62)

When the prevalence of HPV acquisition was stratified according to HPV status, we noted that women who were HPV negative at baseline 27.03% (10/37) were less likely to acquire new HPV infections at second visit compared to women that were HPV positive at baseline 56.52% (39/69) (P = 0.004).

A total of 30.94% (69/223) women were followed up to third visit consisting of 62.32% (43/69) women which were HPV positive at baseline and the remaining 37.68% (26/69) were women that were HPV negative at baseline. Acquisition of new HPV infections was 43.48% (30/69) at third visit. The prevalence of HR-HPV was 40.79% (32/63) while the prevalence of LR-HPV types increased from 32.91% (26/79) at second visit to 49.21% (31/63) at third visit (P=0.049) . The most frequently acquired HPV types were HPVs; 18, 51, 52, 53, 54, 56, 59, 69 and 89 with a prevalence of 4.76% (3/63). The most frequently acquired HPV species was the α3 19.05% (12/63).

We also noted that women who were HPV negative at baseline 23.08% (6/26) compared to women that were HPV positive at baseline 58.13% (25/43) were less likely to acquire new HPV infections at third visit (P=0.005). It is possible that of the 25 women that acquired new HPV infections, 32.00% (8/25) had reactivation of HPV infection or acquired new HPV infections or their HPV type infection was below detectable levels at second visit. This is because the same HPV type(s) were detected at baseline visit and third visit but not at second visit. These HPV types were HPVs 6, 16, 18, 51, 53, 54, 59, 66, 72, 81, 83 and 84.

HPV clearance was defined as the absence of an HPV type that was present during previous visit. A total of 41.51% (44/106) women cleared HPV infections at second visit and 42.03% (29/69) women cleared HPV infections during third visit. At second visit, the most frequently cleared type was the LR-HPV 6 with 7.59% (6/79) and HPV types within the α3 species were the most frequently cleared with 22.78% (18/79) (Table 2.2). The most frequently cleared HR-HPV type was HPV 53 with 6.33% (5/79).

HPV persistence was defined as the detection of an HPV type that was present at previous visit. HPV persistence was 48.11% (51/106) at second visit and 46.38% (32/69) at third visit. Both at second and third visit HR types were the most persistent HPVs; 62.37% (58/93) at second visit and 59.68% (32.38%) at third visit (Table 2.2). The persistence frequency of LR-HPV was 37.63% (35.63) at second visit and 40.32% (25.62) at third visit. At both visits HPV 58 was the most persistent HR-HPV type with a frequency of 8.60% (8/93) at second visit and 11.29% (7/62) at third visit. The most persistent HPVs were from α9 species on both visits with a detection frequency of 24.73% (23/93) at second visit and 27.42% (17/62) at third visit.

2.4.9 HPV prevalence and type distribution according to cervical cytology

Pap smear results were available for a total of 171 women. A total of 146/171 (85.38%) of women had normal cytology while 12.28% (21/171) women had

abnormal cytology. Abnormal cytology consisted of 14 ASCUS and 7 LSIL. Overall, only 2.34% (4/171) specimens had invalid results due to specimen inadequacy. HPV prevalence in women with normal cytology was 65.75% (96/146) while in women with abnormal cytology the prevalence of HPV was 80.95% (17/21). The prevalence of single HPV infection in normal cytology was 40.63% (39/96) and that of multiple HPV infections was 59.38% (57/96). In abnormal cytology consisting of 17 women with HPV infection and four women who are HPV negative inspite of having abnormal cytology results, the prevalence of single HPV infections was 17.65% (3/17) and that of multiple infections was 82.36% (14/17). There were few (n=17) study participants with abnormal cytology, therefore no further analysis could be reported on HPV prevalence according to cervical cytology.

2.5 Discussion

The study presented in this report is a longitudinal analysis of age-specific prevalence of HPV DNA in a population of young women from KwaZulu Natal. The baseline HPV prevalence in this study was 67.26% which is quite high considering the fact that this cohort consists of HIV negative women. When compared to women ≤30 years from another region in South Africa, the prevalence of HPV was higher in our study compared to 41.8% and 52.78% observed among HIV negative women from Cape Town (Allan *et al*., 2008; Marais *et al*., 2008; Mbulawa *et al*., 2010). HPV prevalence is known to vary by region and population subgroups due to sexual behaviour (Bruni *et al*., 2010; Clifford *et al*., 2005b; Dartell *et al*., 2012; De Vuyst *et al*., 2010; Dols *et al*., 2012; Menendez *et al*., 2010; Michala *et al*., 2012; Smith *et al*., 2008; Ting *et al*., 2010; Thomas *et al*., 2004; Vinodhini *et al*., 2012). Therefore, the observed difference between our study and the studies conducted in Cape Town could be due to different methods used in selecting the study population and diversity of HPV DNA detection techniques used by the same laboratory.

HPV prevalent trends reveal that the high HPV prevalence of 67.26% in this Durban cohort consisting of HIV negative women aged 14-30 years is almost

statistically similar to that of HIV positive women (77.27%; 85/110) from the study of Mbulawa et al (2010) (P=0.059). Although the P-values were not significant, they were approaching significance. An obstacle as to why the p-values were not significant could be possibly due to the small number of HIV positive young women aged ≤30 years from the cohort of Mbulawa et al (2010). Among HIV positive women, the most frequently detected HPVs types were HPV 62 (23.53%); HPVs 45, 61 (18.82%); HPV 53 (17.65%); HPV 52 (16.47 %); HPV 70 (15.29%); HPV 54 (14.12%) and HPVs 16, 66, 84 with 12.94% (Mbulawa *et al*., 2010). Whereas in the current study the most commonly detected HPV types were HPV 58, 62 (14.67%); HPV 61 (14.0%); HPV 51 (12.67%); HPV 16 ,52 (11.33%); HPV 53 (10.67%) and HPV 68 (10.00%) (Figure 2.3). Among HIV positive women from Cape Town and HIV negative women from the current study, the most frequently detected HPV type was HPV 62 followed by HPV 61. It has been reported that HPV 62 is overrepresented in HIV positive individuals (Riva *et al*., 2007). When HPV types were stratified according to species, the most frequently detected HPV species in our study was α3 (62.67%) (Table 2.2) whereas among HIV positive women from Cape Town the α3 (70.00%) was also the most commonly detected HPV species (Mbulawa *et al*., 2010). The frequency of single HPV infections within species in our study was α3 (11.70%) and α7 (10.96). Among HIV positive women from Cape Town the frequency of single HPV infection within species was α3 (11.69%) and α7 (12.28%) (Mbulawa *et al*., 2010). Therefore it may be possible that the high prevalence of HIV infections in KwaZulu Natal which is the highest in South Africa (Moodley, 2006) might be influencing the distribution of HPV types in the general population of KwaZulu Natal. However speculation about the distribution of HPV being influenced by HIV still need further investigation for clarification in order to stand valid.

When the prevalence of HPV acquisition was stratified according to HPV status, we noted that women who were HPV negative at baseline were less likely to acquire new HPV infections at second visit and third visit compared to women that were HPV positive at baseline. The observed phenomenon is

in agreement with results reported elsewhere, whereby it was also noted that HPV negative women at baseline have a lower probability of acquiring new HPV infections during follow up compared to women that are HPV positive at baseline (Liaw *et al*., 2001; Rousseau *et al*., 2001). This may be due to the fact that an HPV positive status at enrolment could be an indication of a higher risk sexual behaviour (Rousseau *et al*., 2001). At third visit, the most frequently acquired HPV types were from the α3 species, similar findings have been reported elsewhere, that the risk of an incident HPV infection was greatest among α3 papillomavirus (Goodman *et al*., 2008).

The most frequently cleared type was the LR-HPV 6 and other studies have also reported that LR-HPV types clears more rapidly compared to HR-HPV types (Datta *et al*., 2012; Goodman *et al*., 2008). Castle and others found that the number of cleared LR and HR-HPV types were higher than that of acquired types (Castle *et al*., 2005). This observation is in agreement with our results for only LR types at second visit and HR types at third visit. At second visit HPV types within the α3 species were the most frequently cleared and similar finding have been noted by other studies (Goodman *et al*., 2008). However at third visit the most frequently cleared HPV type was HPV 39 which is a HR type and collectively the most cleared HPV types were from the α7 species.

Both at second and third visit HR types were the most persistent HPVs and this observation has been reported elsewhere (Datta *et al*., 2012; Rositch *et al*., 2012). It has been suggested that in order to estimate HPV persistence with the highest degree of accuracy, time interval should be narrowed down between visits to a month. This is due to the fact that more frequent testing intervals would detect transient infections that may be missed when testing with longer intervals (Rositch *et al*., 2012). The current study estimated HPV persistence between three month intervals and hence there may not be many transient HPV infections which may have been missed.

When the prevalence of HPV was stratified according to age groups, we found that women aged 14-20 years had a higher HPV prevalence (69.23%) compared to women aged 21-30 years (65.55%), however the difference was not statistically significant. The observed trend is consistent with other studies stating that the highest prevalence of HPV is reported in young women then decreases thereafter with increasing age (Bruni *et al*., 2010; Castellsague *et al*., 2001; de Sanjose *et al*., 2007; Healey *et al*., 2001; Liaw *et al*., 2001; Menendez *et al*., 2010; McDonald *et al*., 2012; Smith *et al*., 2008; Swangvaree *et al*., 2010; Thomas *et al*., 2000).

When women were stratified according to cytology, a majority (85.38%) of women had normal cytology while a minority of women (12.28%) had abnormal cytology. This observation was influenced by the few number of cytological results from this study (n=171). It has been reported that in women younger than 30 years HPV infections are most likely to be associated with ASCUS or LSIL (Gage *et al*., 2012). As expected, HPV prevalence was higher in women with abnormal cytology 80.95% compared to women with normal cytology 65.75%. Unsurprisingly, the prevalence of HR-HPV type was higher in women with abnormal cytology than in women with normal cytology (Allan *et al*., 2008).

It is known that the impact of HPV vaccines depends upon the distribution of HR-HPV in a given population (Adler *et al*., 2008), since HR-HPVs are a necessary causes of cervical cancer (Moscicki *et al*., 2006). In our current study the prevalence of HPV types targeted by the quadrivalent vaccine is as following, HPV 16 (11.33%); HPV 18 (8.00%); HPV 6 (9.33%) and HPV 11 (1.33%). The vaccine has some degree of cross protection against HPV 31 and HPV45 which are not present in the vaccine (Wheeler *et al*., 2009). The prevalence of HPV 31 is 4.00% and that of HPV 45 is 9.33%. Although the distribution of HPV types in women with normal cytology varies across the different regions of the world, it has been noted that HPV 16, 18, 31 and 58 were consistently found among the 10 most frequent types in all regions (Bruni *et al*., 2010). Furthermore results show that in sub-Saharan Africa, the

five most frequently detected HR-HPV types among women with normal cytology are HPV 16, 52, 18, 58, and HPV 53 (Louie *et al*., 2009b; HPV Information Centre, 2009). However in our study, the most prevalent HR-HPV types in normal cytology in descending order were; HPV 58 (14.67%); HPV 51 (12.67%); HPV 52, 16 (11.33%); HPV 53 (10. 67%); HPV 68 (10.00%) and HPV 39, 45 (9.33%). Interestingly HPV 58 was also found to be the most prevalent HPV type among Zimbabwean women aged 18-49 years (Fukuchi *et al*., 2009). Several studies have proposed that this type should also be included in the HPV vaccine (Castellsague *et al*., 2001; Chan, 1999; Chan, 2012; Zhang *et al*., 2010) since neoplastic lesions caused by this type are likely to progress to cervical cancer (Castellsague *et al*., 2001).

Epidemiological data on HPV prevalence and type distribution at baseline is criticall to evaluate the impact of HPV vaccines (Bruni *et al*., 2010). Therefore, the data presented in this study serves as a starting point to help in the growing knowledge of understanding the burden of HPV infection. Hence, this data will be useful concerning policy making regarding to HPV screening that would maximize the potential benefits of an HPV vaccine (Marais *et al*., 2008; Smith *et al*., 2008). In general, the Pap smear is used to screen asymptomatic women and consequently this has lead to drastic reduction in the occurrence of cervical cancer in populations where there is wide coverage. In South Africa, women are recommended to have three smears every 10 years from the age of 30 years onwards. Internationally, women are recommended to have their first Pap smear taken within the first three years of sexual debut or at age 21 years annually until at the age of 30 years. After the age of 30 years, women are then recommended to have their Pap smear taken every three years until the age of 65-70 years (Botha *et al*., 2010). However screening women between ages 20-24 years has little or no effect on the incidence of cervical cancer in women younger than 30 years (Sasieni *et al*., 2009). In the current study there were only 12.57% (21/167) women with abnormal cytology and those women were aged 14-25 years. At the age of 30 a majority of these women will have cleared most of these infections leading to the regression of some lesion. Therefore screening

women over the age of 30 years will be effective due to the fact that it will only be those women with persistent HR-HPV types in which lesions could progress and eventually lead to cervical cancer if not treated.

2.6 Conclusion

In conclusion, HPV prevalent trends reveal that the high HPV prevalence in this Durban cohort consisting of HIV negative women aged 14-30 years is similar to that of HIV positive women from Cape Town. Therefore, it could be possible that the high prevalence of HIV infections in KwaZulu Natal might be influencing the distribution of HPV types in the general population of KwaZulu Natal.

We noted that women who were HPV negative at baseline were less likely to acquire new HPV infections at second and third visit compared to women that were HPV positive at baseline. We also noted that the LR-HPV 6 was the most frequently cleared type and that both at second and third visit HR types were the most persistent HPVs.

When the prevalence of HPV was stratified according to age groups, we found that women aged 14-20 years had a higher HPV prevalence compared to women aged 21-30 years; however the difference was not statistically significant.

The prevalence of HPV types targeted by the quadrivalent vaccine is as following, HPV 16 (11.33%); HPV 18 (8.00%); HPV 6 (9.33%) and HPV 11 (1.33%) and that of HPVs 31 and 45 are 40.00 and 9.33%. Results from this study will inform policy makers about which HPV types are circulating in the country and will provide valuable data for assessing the impact of HPV vaccination programs in this country. However it should be noted that decisions on HPV types in vaccine should be based on HPV in cervical cancer data not normal cytology data.

Chapter 3: HPV variants in South African Couples

3.1 Introduction

An HPV genotype is defined as being more than 10% different from all others in their L1 gene sequence (Bernard *et al.,* 2010; Chow *et al*., 2010; de Villiers *et al*., 2004). Genotypes occur in the form of numerous genomic variants that differ by 2% in most genes and 5% in less conserved regions (Bernard *et al*., 2010; Ho *et al*., 1993a). The most variable region of the HPV genome is the LCR (Calleja-Macias *et al*., 2005; de Villiers *et al*., 2004; Schmidt *et al*., 2001); this region is therefore the most appropriate for defining HPV variants.

HPV 16 LCR variants are grouped into 5 distinct phylogenetic branches namely; Asian (As), Asian American (AA), African 1 (Afr1), African 2 (Afr2) and European (Eur) variant (Ho *et al*., 1993a; Mendoza *et al*., 2013; Tanzi *et al*., 2009; Yamada *et al*., 1997). HPV 58 variants are divided into four lineages (A-D) based on nucleotide variation in the LCR. Lineage A is the most common branch and is divided into sublineages A1, A2 and A3. The B lineage is divided into B1 and B2, while lineage C is not divided. Lineage D is divided in D1 and D2 (Chen *et al*., 2011a; Godinez *et al*., 2013). HPV 53 variants can be divided into two clusters which are termed prototype like (P-L) and non prototype like (Non P-L) (Kocjan *et al*., 2007; Prado *et al*., 2005).

Most studies that investigate HPV sharing and transmission between heterosexual partners, consider type-specific concordance which only evaluates sharing of one or more HPV genotypes between partners but not the variant shared (Abalos *et al*., 2012; Benevolo *et al*., 2008; Bleeker *et al*., 2005; Burchell *et al*., 2011; Burchell *et al*., 2010; Castellsague *et al*., 1997; Giovannelli *et al*., 2007; Hernandez *et al*., 2008; Mbulawa *et al*., 2010; Mbulawa *et al*., 2009; Parada *et al*., 2010). Although this method of investigation is accurate, it is not precise and has the tendency to overstate HPV transmission between couples. Therefore the identification of HPV variants by sequencing enables one to investigate the transmission of HPV with a higher level of precision (Ho *et al*., 1993a; Lee *et al*., 2009). Only one study has analysed HPV variant transmission between partners in couples,

and it was found that from a total of eight HPV 16 positive couples, variant concordance was found in only four couples (Ho *et al*., 1993a).

In South Africa, there is only one study which analyzed HPV 16 variant distribution. In 93 Black South African women the Eur variant of HPV 16 was found to be the most common (47%) while the African variants (Afr1 and Afr2) were 41% of all isolates (Tu *et al*., 2006). No studies have analysed the prevalence of HPV 53 and 58 variants in the South African population.

3.2 The aims of the study

In this chapter we examined LCR sequence variation in HPV 16, 53 and 58 positive heterosexual couples to determine which variants were present. The current study used samples from an HPV couples study cohort described by Mbulawa et al (2010) in which the role of factors that influence HPV prevalence was examined in 486 heterosexually active black South African couples. The ultimate goal of this study is to determine HPV variant-specific concordance of HPV types 16, 58 and 53 between partners in a couple.

3.3 Materials and methods

3.3.1 Description of cohort

This study made use of stored specimens and data from an HPV couples cohort study (Mbulawa *et al.,* 2010). The cohort participants were 486 heterosexually active black couples with a mean age of 35 years ranging from 18 to 66 years. These participants were recruited from Empilisweni centre, Cape Town, South Africa. Cervical and penile specimens were collected every 6 months for 24 months. DNA was extracted from both penile and cervical specimens using the MagNA Pure Compact Nucleic Acid Isolation kit (Roche) and an automated MagNA Pure Compact machine (Roche).

Type specific HPV concordance among couples was defined as the presence of the same HPV genotype in both penile and cervical cells of partners in a

relationship. A total of 10% of the HIV negative couples had concordant HPV types while 48% of HIV positive couples shared HPV types. In HIV discordant couples in which the male partner was HIV positive, 11% were HPV concordant while in HIV discordant couples where the female partner was HIV positive, 30% shared HPV types (Mbulawa *et al*., 2010).

A total number of seventeen HPV 16 (Table 3.1), eight HPV 58 (Table 3.2) and sixteen HPV 53 (Table 3.3) concordant couples were identified from the study of Mbulawa et al (2010). Stored DNA specimens isolated from cervical and penile specimens from these couples were used in this study to analyse the variants of HPV 15, 53 and 58 based on the sequence of the LCR.

Table 3.1 HPV 16 concordant couples identified using the Roche Linear Array Genotyping test (Mbulawa *et al*., 2010).

Sample identity	Gander	Age	HIV status	HPV types	cytology
181-1	Female	28	-	**16**, 58	LSIL
181-2	Male	32	-	**16**	-
282-1	Female	19	-	**16**, 18	normal
282-2	Male	26	-	**16**	-
368-1	Female	46	+	**16**, 52	normal
368-2	Male	63	+	6, **16**, 35, 45, 53, 68, 70, 72, 83, 89	-
440-1	Female	18	+	6, **16**, 39, 45, 58, 59, 61, 62, 66, 73, 81, IS39	LSIL
440-2	Male	22	-	6, **16**, 45, 59, 61, 62, 81, 66	-
460-1	Female	31	+	**16**, 58, 61, 66, 68, 69	normal
460-2	Male	40	-	6, **16**, 58, 61, 66, 68, 69	-
501-1	Female	41	+	6, **16**, 35, 39, 51, 53, 62, 66, 68, 72, 73, 81	ASCUS
501-2	Male	37	+	**16**, 35, 53, 68	-
511-1	Female	30	+	11, **16**, 42, 58	LSIL
511-2	Male	33	+	11, **16**, 39, 42, 55, 58, 61, 68	-
529-1	Female	35	+	**16**, 35, 51, 53, 55, 58, 62, 68, 81 ,84	HSIL
529-2	Male	45	-	**16**, 35, 51, 53, 55, 58, 62, 68, 81 ,82	-
571-1	Female	28	+	**16**,18, 52, 53, 55, 56, 61, 62, 72, 84, IS39	LSIL
571-2	Male	25	-	**16**,18, 45, 52, 53, 55, 56, 61, 62, 66, 72, 84 IS39	-
605-1	Female	21	-	**16**	normal
605-2	Male	25	-	**16**	
621-1	Female	33	+	**16**	LSIL
621-2	Male	36	-	**16**, 53, 61	-
626-1	Female	23	+	**16**, 52, 70, 84	normal
626-2	Male	25	+	**16**, 45, 53, 66, 68, 70, 83	-
727-1	Female	42	+	**16**, 53, 73, 81	LSIL
727-2	Male	53	-	**16**, 73, 81	-

Sample identity	Gander	Age	HIV status	HPV types	cytology
744-1	Female	22	-	**16**, 31, 70	normal
744-2	Male	26	+	11, **16**, 59, 89	-
763-1	Female	36	+	**16**, 40, 83	LSIL
763-2	Male	29	-	**16**, 54, 61, 83, 84	-
811-1	Female	30	-	**16**, 31, 35, 39, 45, 54, 55, 56, 62, 68, 70, 72, 84	LSIL
811-2	Male	23	+	**16**, 18, 45, 51, 53, 54, 68, 69, 70, 84	-
821-1	Female	41	+	**16**	HSIL
821-2	Male	39	+	**16**, 62, 84	-

Table 3.2 HPV 58 concordant couples identified using the Roche Linear Array Genotyping test (Mbulawa *et al*., 2010).

Sample identity	Gander	Age	HIV status	HPV types	cytology
153-1	Female	38	+	**58**	LSIL
153-2	Male	49	-	**58**, 62	-
313-1	Female	39	+	26, **58**	normal
313-2	Male	40	+	26, **58**	-
460-1	Female	31	+	16, **58**, 61, 66, 68, 69	normal
460-2	Male	40	-	6, 16, **58**, 61, 66, 68, 69	-
511-1	Female	30	+	11, 16, 42, **58**	LSIL
511-2	Male	33	+	11, 16, 39, 42, 55, **58**, 61, 68	-
529-1	Female	35	+	16, 35, 51, 53, 55, **58**, 62, 68, 81 ,84	HSIL
529-2	Male	45	-	16, 35, 51, 53, 55, **58**, 62, 68, 81 ,82	-
623-1	Female	25	+	26, **58**	LSIL
623-2	Male	24	-	26, **58**	-
683-1	Female	36	+	51, 53, **58**, 61, 83, 89	normal
683-2	Male	29	+	51, 53, **58**, 59, 61, 62, 83	-
762-1	Female	30	+	33, 35, 42, **58**, 61, 62, 68, 70, 71, 72, 81, 83	normal
762-2	Male	39	+	33, 35, 42, 45, **58**, 62, 68, 70, 71, 81, 83	-

Table 3.3 HPV 53 concordant couples identified using the Roche Linear Array Genotyping test (Mbulawa *et al*., 2010).

Sample identity	Gander	Age	HIV status	HPV types	cytology
308-1	Female	45	+	**53**, 58	LSIL
308-2	Male	49	+	**53**	-
339-1	Female	19	-	**53**, 54	normal
339-2	Male	21	-	**53**, 56, 59	-
364-1	Female	25	+	31, 42, 45, 52, **53**, 61, 62	ASCUS
364-2	Male	29	-	31, 45, 52, **53**, 54, 61, 62	-
408-1	Female	28	-	**53**, 55, 61, 62, 71, 83, 84, 89	normal
408-2	Male	34	-	**53**, 55, 59, 61, 62, 66, 70, 71, 83, 84, 89	-
461-1	Female	25	+	39, 45, **53**, 58, 83	normal
461-2	Male	41	-	**53**	-
501-1	Female	41	+	6, 16, 35, 39, 51, **53**, 62, 66, 68, 72, 73, 81	ASCUS
501-2	Male	37	+	16, 35, **53**, 68	-

Sample identity	Gander	Age	HIV status	HPV types	cytology
529-1	Female	35	+	16, 35, 51, **53**, 55, 58, 62, 68, 81 ,84	HSIL
529-2	Male	45	-	16, 35, 51, **53**, 55, 58, 62, 68, 81 ,82	-
547-1	Female	31	+	**53**, 61, 66,	LSIL
547-2	Male	32	-	42, **53**, 56, 89	-
571-1	Female	28	+	16,18, 52, **53**, 55, 56, 61, 62, 72, 84, IS39	LSIL
571-1	Male	25	-	16,18, 45, 52, **53**, 55, 56, 61, 62, 66, 72, 84, IS39	-
669-1	Female	47	+	16, **53**, 83	ASCUS
669-2	Male	30	-	39, 40, **53**, 62, 66	-
683-1	Female	36	+	51, **53**, 58, 61, 83, 89	normal
683-2	Male	29	+	51, **53**, 58, 59, 61, 62, 83	-
693-1	Female	37	+	33, 51, **53**, 61, 62, 84	LSIL
693-2	Male	45	-	31, 33, 51, **53**, 61, 62, 84	-
696-1	Female	27	+	6, 35, **53**, 55, 62, 84	LSIL
696-2	Male	36	-	35, 40, **53**	-
702-1	Female	27	+	**53**	normal
702-2	Male	32	+	35, 52, **53**, 58, 66, 67	-
799-1	Female	25	-	**53**, 70	normal
799-2	Male	35	+	**53**, 58, 62, 72, 84	-
805-1	Female	36	+	52, **53**, 62, 68, 83	LSIL
805-2	Male	36	-	31, 45, 52, **53**, 62, 66, 70, 72	-

3.3.2 PCR amplification of HPV 16, 58 and 53 LCR

A total number of seventeen HPV 16, eight HPV 58 and sixteen HPV 53 positive couples were identified in the HPV couples cohort using the Roche Linear Array HPV Genotyping data from the study of Mbulawa et al (2010). To analyse the sequence variants of HPV genotypes 16, 53 and 58 in these specimens the LCR was amplified by PCR and sequenced. Primers specific to HPV 16, 58 and 53 LCRs were used to target their respective LCRs. The primers are shown in Tables; 3.4-3.6.

Table 3.4 Primers used in the amplification of HPV 16 (Kurvinen *et al*., 2000)

Primer name	Sequence and length of primers	Position	Length (bp)
HPV16-LCR-F	5'-CCTCATCTACCTCTACAACTGCTAAACGC-3' (29)	7108-7136	1028
HPV16-LCR-R	5'-CGTCGCAGTAACTGTTGCTTGCAGTACACAC-3' (31)	222-192	

Table 3.5 Primers used in the amplification of HPV 58 (Chan *et al*., 2011)

Primer name	Sequence and length of primers	Position	Length (bp)
HPV58-LCR-out-F	5'-GATCAGTTTCCTTTGGGACG-3' (20)	7017-7036	948
HPV58-LCR-out-R	5'-ACCTCAGATCGCTGCAAAGT-3' (20)	215-234	
HPV58-LCR-in-F	5'-CCTTAAAGCAAAGCCCAGAC-3' (20)	7058-7077	
HPV58-LCR-in-R	5'- ATGCACAGATGTCTCCAACG-3' (20)	162-181	

Table 3.6 Primers used in the amplification of HPV 53 (Prado *et al*., 2005)

Primer name	Sequence and length of primers	Position	Length (bp)
HPV53-LCR-F	5'-TTTGCATGTTGTTAATAAATA T-3' (22)	7270-7291	602
HPV53-LCR-R	5'-AGTAGGATTGTTACTTTC-3' (18)	1-18	

PCR reactions for the amplification of the different LCRs of each HPV type were set up in a total volume of 25µl. The volumes and concentrations of PCR components were as follows: 500 nM of each primer mixed with 12.5µl of ImmoMix™ (Bioline, Massachusetts, USA) which is a mixture that consists of 2mM dNTP mix, IMMOLASE™ DNA polymerase enzyme, 32 mM $(NH_4)_2SO_4$, 125 mM Tris-HCl (pH 8.3), Stabilizer, 0.02% Tween 20 and 1.5mM magnesium salt to enhance enzyme activity. A total of 1µl DNA sample was added to each PCR. The ImmoMix™ cocktail contains a dNTP mixture consisting of the nucleotides (dCTP, dTTP, dATP and dGTP) for the extension of primers. The high-specific IMMOLASE™ DNA polymerase enzyme was used to efficiently amplify the LCR. For HPV 16 LCR amplification, the protocol was carried out according to Kurvinen et al (2000) with minor modifications: initial denaturation at 95°C for 7 minutes, 35 cycles consisting of 94°C for 1 minute, 56.5°C for 1 minute and 72°C for 2 minutes to allow for primer extension. Final extension was at 72°C for 10 minutes for the final elongation of primers.

The amplification protocol for the LCR of HPV 58 was followed according to Chan et al (2011) with minor modifications. Thermal conditions for outer primers were as following; initial denaturation at 95°C for 7 minutes, 40 cycles consisting of 94°C for 30 seconds, 58°C for 30 seconds and 72°C for 1 minute to allow for primer extension. Final extension was at 72°C for 7 minutes for the final elongation of primers. Thermal conditions for inner primers were as following; initial denaturation at 95°C for 7 minutes, 40 cycles consisting of 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 1 minute to allow for primer extension. Final extension was at 72°C for 7 minutes for the final elongation of primers. The amplification protocol for the HPV 53 LCR was followed according to Prado et al (2005) with modifications, and consisted of

an initial denaturation at 95°C for 7 minutes, followed by 40 cycles of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 minute to allow for primer extension. Final extension was at 72°C for 7 minutes for complete synthesis of all the PCR products. Agarose gel electrophoresis of HPV 16, 58 and 53 DNA was performed on 1% agarose gels at 80 volts for DNA fragments larger than one kilobase (kb). DNA separates on agarose gel according to size of DNA fragment. Electrophoresis of PCR amplicons less than one kb was carried out on 2% agarose gels at 100 volts. The separation of DNA fragments larger than 5kb was performed using 0.8% agarose gel. All agarose gels contained ethidium bromide at a concentration of 0.5µg/ml for DNA visualization on a UV transilluminator at 260nm (Syngene). 1X Tris-Borate EDTA (TBE) was the buffer medium used for gel electrophoresis. DNA fragment size was estimated by using the one kb O'Gene Ruler (range 250bp-10000bp) and 100bp O'Gene Ruler (range 100bp-1000bp) (Fermentas, USA) for large and small DNA fragments.

Samples in which the LCR could not be successfully amplified using this method were non-specifically amplified by rolling circle amplification (RCA) as described in section 3.3.3.

3.3.3 Rolling circle amplification of DNA for PCR

Rolling circle amplification (RCA) was performed using the Illustra™ TempliPhi 100 Kit (Amersham Biosciences, GE Healthcare, UK). This is the most commonly used kit (Rector *et al*., 2004) and is most suitable to use when there is a tiny amount of starting DNA (Nelson *et al*., 2002). The advantages of using RCA to amplify HPV genome are that this technique is simple, rapid and efficient method for the non-specific enrichment of circular DNA (Johne *et al*., 2009; Marincevic-Zuniga *et al*., 2012). In addition to that, this technique does not make use of any special instrument to cycle the temperature except the PCR machine which is widely used for DNA diagnostic work (Demidov, 2002).

The reactions were carried out according to manufacturer's instructions. DNA specimens for amplification (1µl) were mixed with 5µl of sample buffer in a 0.2ml thin wall PCR tube. The sample buffer contained exonuclease-protected random hexamers. The mixture was heated at 95°C for 3 minutes to denature the double stranded (ds) DNA and cooled to 4°C. A premix of the TempliPhi enzyme mix (0.2µl), TempliPhi reaction buffer (5µl) and additional 0.45µl 10mM dNTP were prepared in a separate tube. The enzyme mix contained the Φ29 DNA polymerase, extra dNTPs and extra hexamers in 50% glycerol. The TempliPhi reaction buffer contained dNTPs and salts for enzyme activity. Five microlitres of the premix were added to the cooled sample and mixed by vortexing. The reaction was incubated at 30°C for 18 hours on an AB9700 machine (Applied Biosystems).

The amplification reaction was initiated when random hexamers hybridized to the circular DNA (Figure 3.1). The (Φ29) DNA polymerase extends each of the primers until it reaches the next primer downstream. The (Φ29) DNA polymerase then displaces the next downstream primers, rendering the dsDNA to single strand after displacement. The displaced single strand then acts as an available template to which more hexamers can anneal for further extension. The extension process results in the formation of dsDNA again. This process is repeated several times and results in the formation of repeated whole HPV genome sequences which are termed concatemers (Demidov, 2002; Johne *et al*., 2009; Nelson *et al*., 2002; Rector *et al*., 2004). This cycle can continue until the supply of free nucleotides are exhausted (Demidov, 2002; Nelson *et al*., 2002).

The RCA was carried out for 18 hours, then the mixture was incubated at 65°C for 10 minutes to inactivate the bacteriophage Φ29 DNA polymerase. This step was then followed by diluting the mixture by adding 30µl of HPLC water to each tube. Two microlitres of sample were electrophoresed on 0.8% agarose gel at 85 volts. From the RCA amplicons, the respective LCRs of the different HPV types were amplified using type-specific primers for the

respective HPV types according to the former described protocols. The rest of sample was stored at -20°C for further use in subsequent reactions.

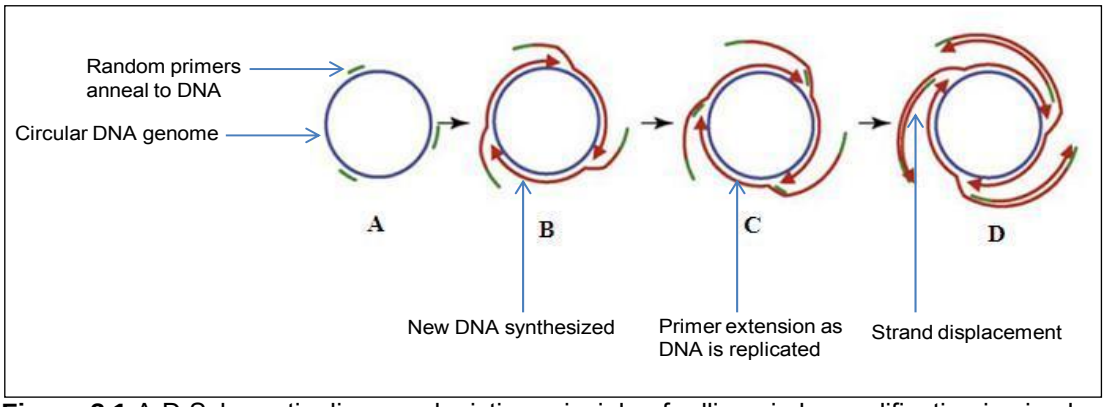


Figure 3.1 A-D Schematic diagram depicting principle of rolling-circle amplification in circular genomes. Blue lines indicate target DNA sequence, green line depicts oligonucleotide primers and red lines represent new DNA synthesized by the polymerase. Random oligonucleotide primer anneal to circular DNA at multiple sites (A). DNA polymerase begins to synthesize complementary strand at the bound primer (B). The complementary DNA strand is extended until it reaches the next bound primer (C), where strand displacement occurs and DNA synthesis continues for additional cycles. Primers that are still present in the mixture continue to bind to the displaced strand. Consequently giving rise to additional initiation points. This process results to long concatemiric products of double-stranded DNA molecules (D) (Johne *et al*., 2009).

3.3.4 DNA Sequencing and phylogenetic analysis

The amplified LCR PCR products were checked for the presence of a unique HPV LCR band on an agarose gel. The PCR products were sent to Stellenbosch Central Analytical Facility (Stellenbosch University) where they were purified and sequenced directly using a Big Dye Terminator v3.1 Cycle sequencing kit in an ABI DNA sequencer 3130 (Applied Biosystems). The forward and reverse primers used for the PCR amplification were used for sequencing both strands of the amplicons.

Sequences were aligned using Muscle (Edgar, 2004) and ambiguous alignments refined with GBlocks (Castresana, 2000). Phylogeny was estimated by maximum likelihood with PhyML 3.0 (Guindon *et al*., 2010) assuming a General Time Reversible (GTR) nucleotide substitution model and gamma distributed substitution rate among sites as the best fit for the LCR sequences (Chan *et al*., 2011; Sun *et al*., 2012). Confidence was estimated by 1000 bootstrap replicates.

To confirm that sequence variations identified in the LCR were not as a result of Taq polymerase errors during amplification the PCR amplification and sequencing were independently repeated.

3.4 Results

3.4.1 PCR amplification of HPV 16, 58 and 53 LCR

The LCR was successfully amplified in twelve out of the seventeen HPV 16 positive couples. A representative agarose gel of the PCR products for eight male specimens is shown in Figure 3.2. The LCR was amplified in seven specimens; 181-2, 368-2, 440-2, 460-2, 501-2, 621-2 and 727-2 and not amplified in one specimen 605-2. The amplicon size was 1028bp and the same amplicon size was obtained from specimens 181-1, 368-1, 440-1, 460-1, 501-1, 511-1, 511-2, 529-1, 529-2, 571-1, 571-2, 621-1, 626-1, 626-2, 727-1, 811-1 and 811-2 (results not shown).

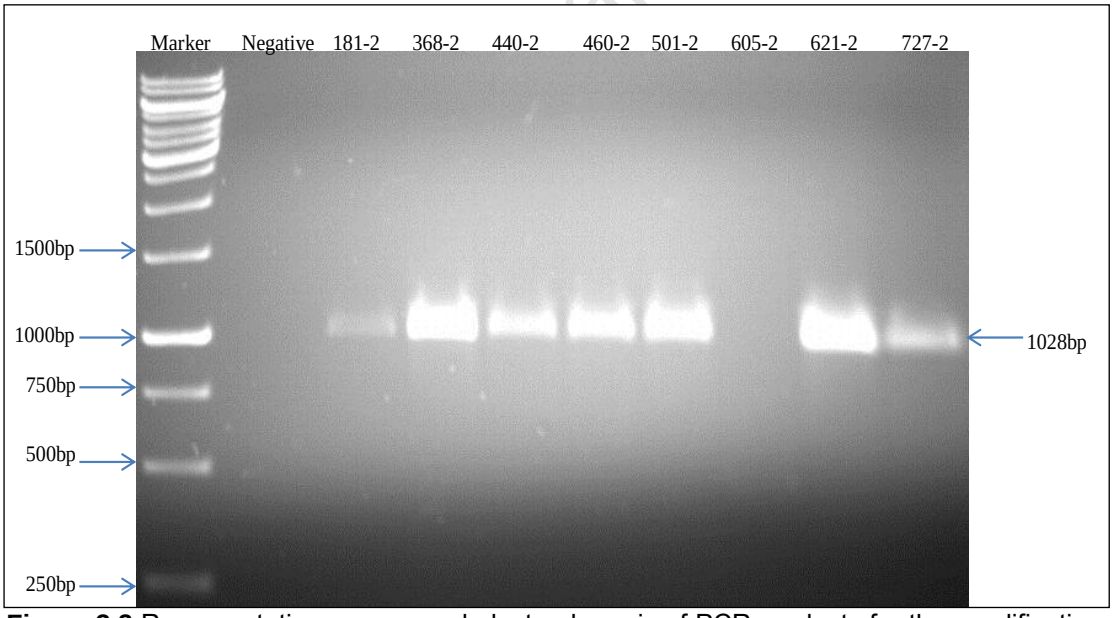


Figure 3.2 Representative agarose gel electrophoresis of PCR products for the amplification of the HPV 16 LCR from clinical specimens. The expect size of the amplification product (1028bp) is indicated on the right hand side. Specimen numbers are indicated at the top of the gel. A negative control PCR which had water instead of template was included (Negative). The sizes of the DNA fragments in the DNA size marker (Marker) are indicated on the left.

The HPV 58 LCR was successfully amplified in all eight HPV 58 positive couples. A representative agarose gel is shown in Figure 3.3 with the PCR products from three female specimens 460-1; 511-1 and 529-1. The amplicon size was 948bp and the amplicons of the same size were obtained from the other thirteen specimens 153-1, 153-2, 313-1, 313-2, 460-2, 511-2, 529-2, 623-1, 623-2, 683-1, 683-2, 762-1 and 762-2 (results not shown).

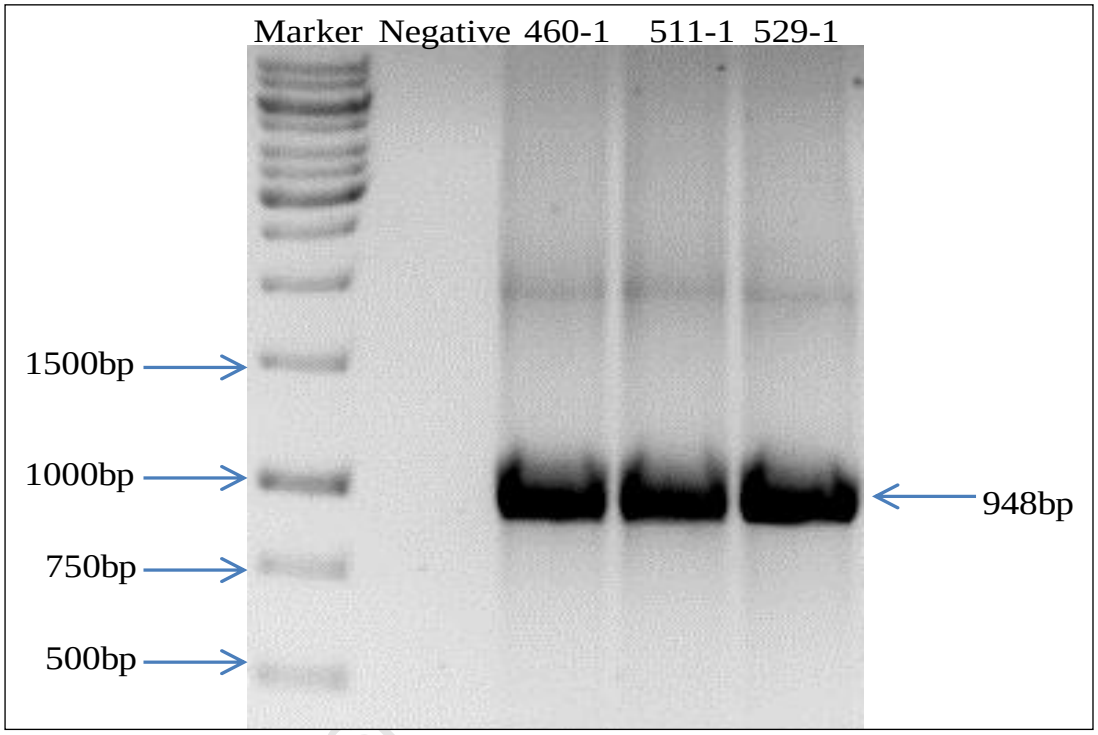


Figure 3.3 Representative agarose gel electrophoresis of PCR products for the amplification of the HPV 58 LCR from clinical specimens. The expect size of the amplification product (948bp) is indicated on the right hand side. Specimen numbers are indicated at the top of the gel. A negative control PCR which had water instead of template was included (Negative). The sizes of the DNA fragments in the DNA size marker (Marker) are indicated on the left.

The HPV 53 LCR was successfully amplified in twelve out of sixteen HPV 53 positive couples. A representative agarose gel is shown in Figure 3.4, with the PCR products from six female specimens; five specimens were LCR positive and one specimen was LCR negative. The amplicon size was 602bp and the same amplicon size was obtained from the remaining nineteen HPV 53 positive specimens 308-2, 339-1, 339-2, 364-1, 364-2, 408-1, 408-2, 461-1, 461-2, 501-2, 529-1, 529-2, 571-2, 669-2, 683-1, 683-2, 696-2, 702-1 and 702-2 (results not shown).

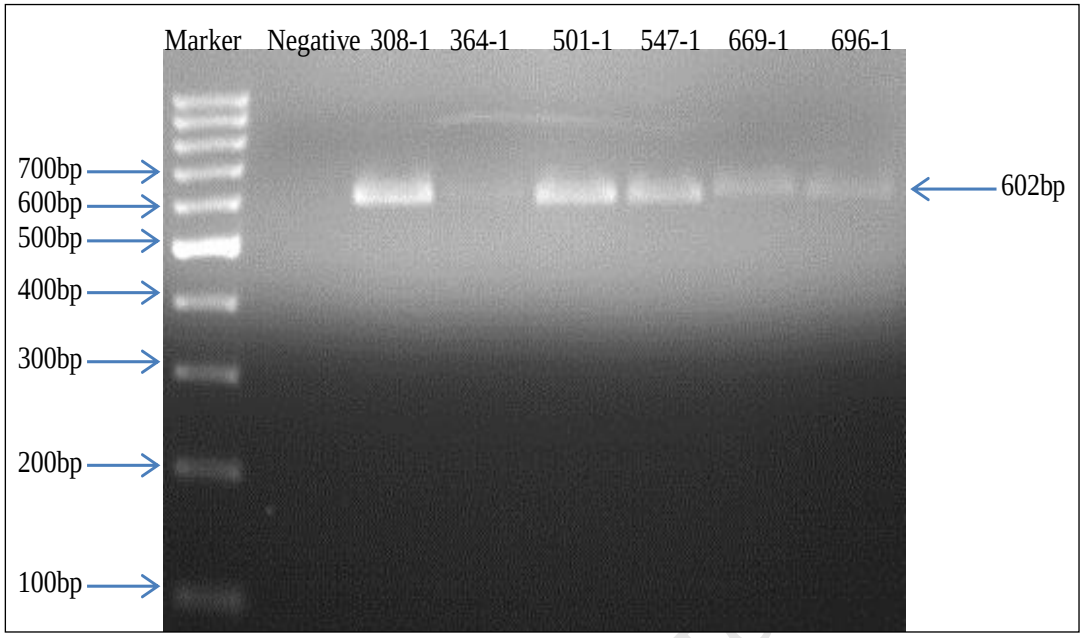


Figure 3.4 Representative gel picture of amplified HPV 53 LCR from clinical specimens after electropheresis. The expect size of the amplification product (602bp) is indicated on the right hand side. Specimen numbers are indicated at the top of the gel. A negative control PCR which had water instead of template was included (Negative). The sizes of the DNA fragments in the DNA size marker (Marker) are indicated on the left.

HPV 16, 58 and 53 positive samples in which the LCR was not successfully amplified were subjected to non-specific amplification using RCA. Amplification products were analysed by agarose gel electrophoresis. As RCA forms concatemers of the amplified DNA the product is visible as a high molecular weight band that migrates to a position above the 8000 bp DNA size marker as indicated in Figure 3.5. All the specimens in which the LCR could not be efficiently amplified were successfully amplified by RCA. Figure 3.5 shows results from five male's (from 511-2 to 626) specimens that were HPV 16 positive.

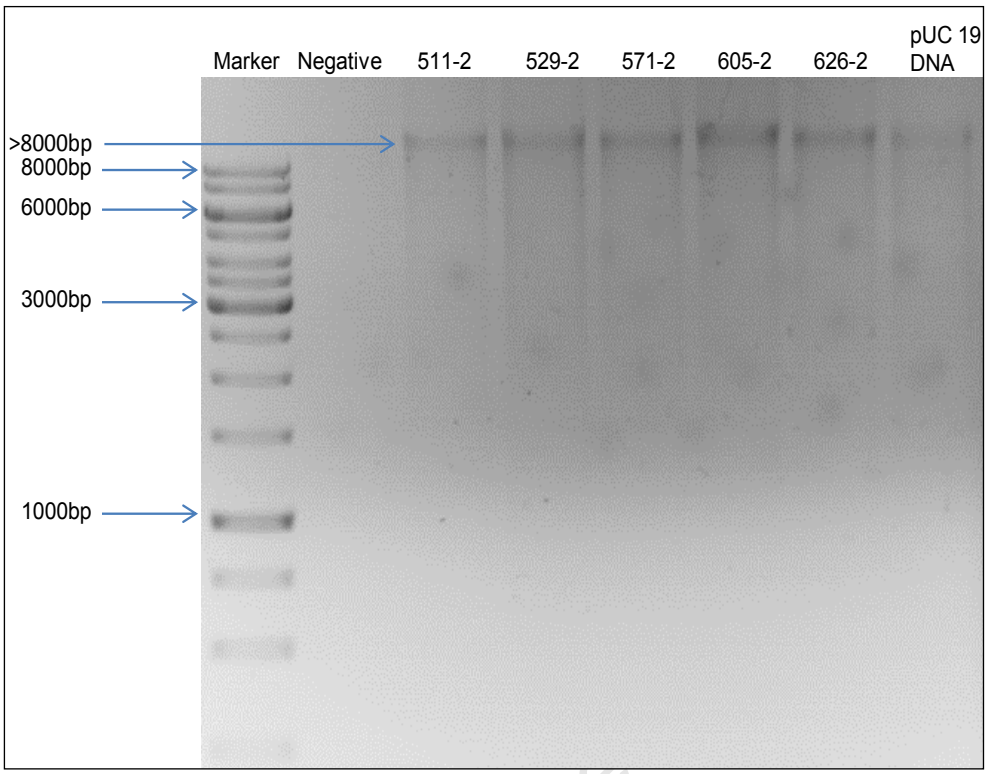


Figure 3.5 Representative agarose gel electrophoretic analysis of products of rolling circle amplification of DNA from clinical specimens**.** Specimen numbers are indicated at the top of the gel. Water was used as negative control. Positive control was the plasmid pUC 19 plasmid.

Following RCA, enriched specimens were used as templates for the amplification of the LCR for HPVs 16, 58 and 53 using the type-specific primers previously described. Non-specific amplification of the template or target sequence by RCA increased the template concentration allowing for efficient amplification of the LCR. This is illustrated in Figure 3.6, where a specimen was amplified using PCR only, and only a feint DNA band of the expected size (1028 bp) was visible (A). However when the same specimen was subjected to RCA prior to the LCR PCR (B) a bright band was observed following agarose gel electrophoretic analysis.

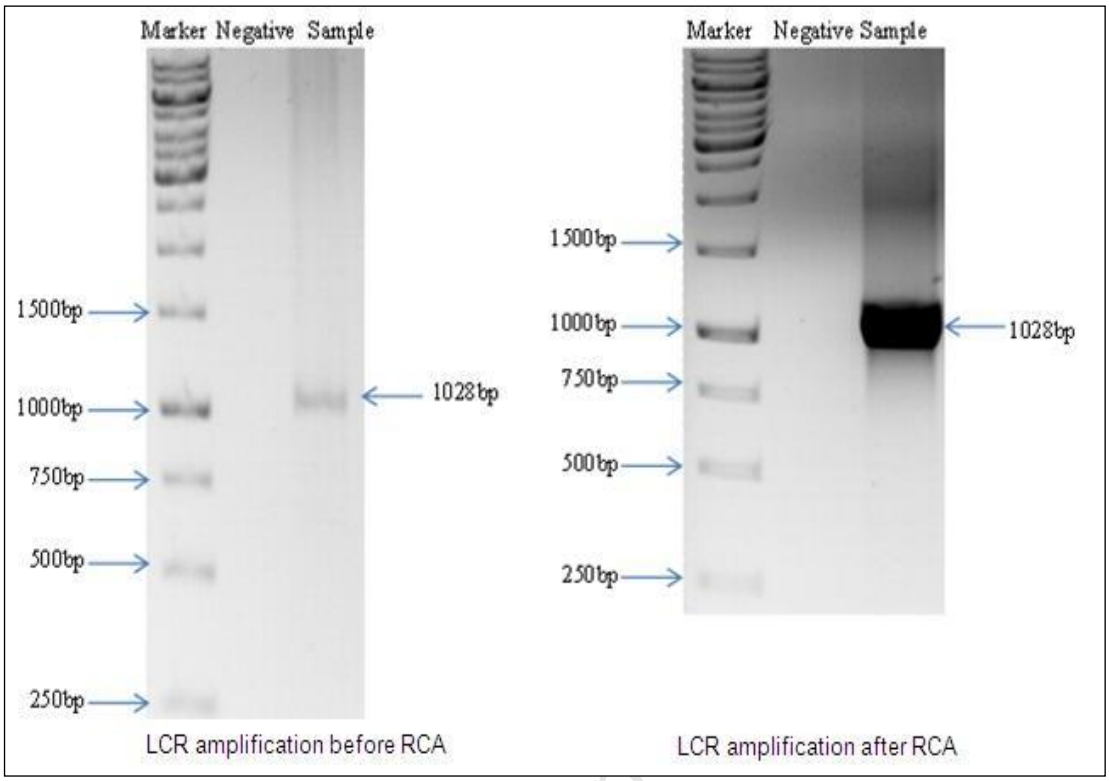


Figure 3.6 Representative gel picture of amplified HPV 16 LCR after electropheresis. The agarose gel picture on the left shows PCR amplified LCR from clinical specimen prior to rolling circle amplification (A). While the agarose gel picture on the right (B) shows PCR amplified LCR produced from RCA amplified specimen, both samples were from the same specimen.

3.4.2 HPV 16 LCR sequence analysis

All the amplicons were sequenced directly in order to identify nucleotide differences within the LCR in comparison to the prototype HPV 16 European variant. This study made use of the HPV 16 corrected European reference sequence (Eur) retrieved from Papillomavirus Episteme or PAVE published in GenBank with the accession number (NC_001526) (van Doorslaer *et al*., 2013). The positions of nucleotide substitutions in the different variants identified are shown in Table 3.7. The LCR was successfully amplified and sequenced in a total of 12/17 (70.6%) couples and revealed a total of 19 different point mutations, including 10 transitions and 9 transversions.

Common mutations were defined as mutations that occurred in ≥ ten isolates. A total of fourteen common mutations were observed. These mutations were at the following positions; G7385C, C7432A, A7485C, G7489A, G7521A,

C7669T, C7689A, C7764T, C7786T, G7826A, G7834T, A7837C, A7839G, and C31T (Table 3.7).

Table 3.7 Nucleotide mutations identified within the LCR of HPV 16

Base position in HPV 16R		7308	7314	7372	7385	7432	7485	7489	7521	7669	7689	7729	7764	7786	7807	7826	7834	7837	7839	31	Total
Base in HPV16R sequence		C	A	A	G	C	A	G	G	C	C	A	C	C	T	G	G	A	A	C	19
Variant	Sample Identity																				
AA	181-1	T					C	A	A	T	A	C	T	T			T				10
	181-2	T					C	A	A	T	A	C	T	T			T				10
Eur (I)	368-1								A												1
	368-2								A												1
Afr2a (I)	440-1				C	A	C	A	A	T	A		T	T		A	T	C	G	T	14
	440-2				C	A	C	A	A	T	A		T	T		A	T	C	G	T	14
Afr2a (I)	460-1				C	A	C	A	A	T	A		T	T		A	T	C	G	T	14
	460-2				C	A	C	A	A	T	A		T	T		A	T	C	G	T	14
Afr2a (I)	501-1				C	A	C	A	A	T	A		T	T		A	T	C	G	T	14
	501-2				C	A	C	A	A	T	A		T	T		A	T	C	G	T	14
Afr2a (II)	511-1			C		A	C	A	A	T	A		T	T	C	A	T	C	G	T	15
	511-2			C		A	C	A	A	T	A		T	T	C	A	T	C	G	T	15
Eur	529-1																				0
	529-2																				0
Eur (II)	571-1		C																		1
	571-2		C																		1
Afr2a (I)	621-1				C	A	C	A	A	T	A		T	T		A	T	C	G	T	14
	621-2				C	A	C	A	A	T	A		T	T		A	T	C	G	T	14
Eur	626-1																				0
	626-2																				0
Afr2a (I)	727-1				C	A	C	A	A	T	A		T	T		A	T	C	G	T	14
Eur (I)	727-2								A												1
Eur (II)	811-1		C																		1
Afr2a (I)	811-2				C	A	C	A	A	T	A		T	T		A	T	C	G	T	14

Single nucleotide base differences within HPV 16 LCR variants in comparison to the European prototype sequence HPV 16R, GenBank accession number (NC_001526), retrieved from Papillomavirus Episteme or PAVE (van Doorslaer *et al*., 2013). Sample identity denotes the specimen number and the letters (I-II) refer to variations within the sublineage analysed in the various specimens. Total indicates the overall number of mutations in each specimen analysed. The green highlighted base pair positions denote the position where the most frequent mutation occurred among all the isolates. Red bases indicate novel mutations.

The number of mutations in each isolate ranged from 0-15. The most common mutation was a transition point mutation G7521A (highlighted in Table 3.7), which was detected in 70.8% (17/24) of the specimens. We observed two nucleotides (nt) changes that have not been reported before.

These two nt changes were at positions A7372C and T7807C both detected in Afr2a sublineages.

There were some mutations observed in transcription factor binding regions of the LCR. A transversion mutation G7385C was observed in the glucocorticoid response elements (GRE) region in 41.67% (10/24) isolates from the Afr2a (I) variants. A transversions mutation G7489A found in the GRE transcriptional binding site in 58.33% (14/24). The mutations; A7485C, G7521A, C7786T and G7826A were found within the YY1 binding site. Transversion and transition mutations; A7729C and C7764T were detected at the nuclear factor 1 (NF1) transcription factor binding site. A transversions mutation C7689A was found in the TEF1 binding site, a transversions mutation G7834T was found in the octamer-binding factor 1 (Oct1) and the transition C31T was found in the Sp1 transcription factor (Lei *et al*., 2011; Mendoza *et al*., 2013; Tanzi *et al.,* 2009; Pande *et al.,* 2008; Pientong *et al*., 2013).

3.4.3 HPV 58 LCR sequence analysis

All the amplicons were sequenced in order to identify mutations within the LCR in comparison to the HPV 58 prototype. This study made use of the HPV 58 prototype sequence published in GenBank with the accession number NC_001443/D90400 as the reference sequence (Cento *et al*., 2011; Chan *et al*., 2011; Chang *et al*., 2011; Liu *et al*., 2012; Raiol *et al*., 2009; Wu *et al*., 2009). The positions of mutated bases in the different variants identified are shown in Table 3.8. The LCR was successfully amplified and sequenced in all eight couples. LCR sequence analysis revealed 30 different point mutations in 28 nucleotide positions, with 14 transitions and 16 transversions.

Table 3.8 Nucleotide mutations identified within the LCR of HPV 58

Base position in HPV 58R		7181	7182	7186	7188	7189	7199	7207	7265	7266	7284	7345	7369	7421	7431	7435	7483	7522	7686	7714	7730	7745	7779	7786	30	52	53	54	62	Total
Base in HPV 58R sequence		G	T	A	G	T	C	T	C	C	C	T	T	G	T	A	T	T	G	A	C	G	A	T	C	C	T	G	C	28
Variant	**Sample Identity**																													
A2 (I) variant	153-1									T										G										2
A2 (I) variant	153-2									T										G										2
B2 variant	313-1			C			T	A			G					G		C	A		A	C	G		G	G		A	A	14
B2 variant	313-2			C			T	A			G					G		C	A		A	C	G		G	G		A	A	14
C variant	460-1	T	G	G	T	C		A			G	C	G		G		G				A		G		G	T	C	A	A	18
C variant	460-2	T	G	G	T	C		A			G	C	G		G		G				A		G		G	T	C	A	A	18
A2 (I) variant	511-1									T										G										2
A2 (I) variant	511-2									T										G										2
A2 (II) variant	529-1								G	T				A						G										4
A2 (II) variant	529-2								G	T				A						G										4
A2 (I) variant	623-1									T										G										2
A2 (I) variant	623-2									T										G										2
C variant	683-1	T	G	G	T	C		A			G	C	G		G		G				A		G		G	T	C	A	A	18
A2 (III) variant	683-2								G	T										G				G						4
C variant	762-1	T	G	G	T	C		A			G	C	G		G		G				A		G		G	T	C	A	A	18
C variant	762-2	T	G	G	T	C		A			G	C	G		G		G				A		G		G	T	C	A	A	18

Single nucleotide base differences within HPV 58 LCR variants in comparison to the prototype sequence HPV 58R, GenBank accession number NC_001443/D90400. Sample identity denotes the specimen number. Identical variants within the A2 lineage were grouped into subgroups I to III. Total indicates the overall number of mutations in each specimen analysed. The green highlighted base pair positions denote the position where the most frequent mutation occurred among all the isolate. Red bases indicate novel mutations.

Common mutations were defined as mutations that occurred in ≥ seven isolates. In most of the isolates from eight couples, we observed a total number of 9 common mutations. These mutations were at the following positions; T7207A, C7266T, C7284G, A7714G, C7730A, A7779G, C30G, G54A and C62A (Table 3.8). A total of four different point mutations were found in two different nucleotide positions. These mutations were transitions at A7186G and C52T, and within the same positions tranversions were found A7186C and C52G. The number of mutations in each isolate ranged from 2-18. The most common mutations were transitions point mutations C7266T and A7714G (highlighted in Table 3.8), which were detected in 56.3% (9/16) of the specimens.

There were two nt changes that have not been reported before. These nt changes occurred in the following positions; C7199T and T7786G. Among the two nt changes; C7199T was detected in B2 variant while T7786G was detected in only one out of nine A2 variant which was termed A2 (III) variant. Only one mutation was observed in transcription factor binding regions of the LCR. A transition mutation A7714G was observed at the NF1 transcription factor binding site (Raiol *et al*., 2009).

3.4.4 HPV 53 LCR sequence analysis

All the amplicons were sequenced in order to identify mutations within the LCR in comparison to the prototype HPV 53. This study made use of the HPV 53 prototype sequence published in GenBank with the accession number X74482 as the reference sequence (Cento *et al*., 2012; Kocjan *et al*., 2007; Wyant *et al*., 2011). The positions of all mutated bases in the different variants identified are shown in Table 3.9. The LCR was successfully amplified and sequenced in a total of (12/16; 75%) couples. LCR sequence analysis revealed 29 different point mutations in 34 nucleotide positions, with 19 transitions, 9 transversions, four insertions and five deletions.

Common mutations were defined as mutations that occurred in ≥ 12 isolates. In most of the isolates from the twelve couples, we observed 19 mutations common. These mutations were at the following positions; T7422A, C7428T, C7432G, G7508A, T7528C, TGGGA insertion between 7530 and 5731, G7531A, T7538C, T7618G, T7619C, T7634C, A7655G, C7676G, C7685T, C7756T, G7758A, C7784T, A7797C and C7810T (Table 3.9). The most common mutations was a transitions point mutation at position G7508A (highlighted in Table 3.9), which was detected in all of the specimens. There were a total of nine nt changes and five deletions that have not been reported before. These nt changes occurred in the following positions; A7398G, A7566G, G7608A, T7609G, G7656T, T7709G, A7751G, T7760C, A7766G and deletions at the following positions; A7533, T7534, A7535, T7536, C7537. Among the nine nt changes (5/9 55.56%); A7398G, G7608A, T7609G, G7656T, A7751G and all five deletions; A7533, T7534, A7535, T7536, C7537 were detected in (non P-L) (IV) variants. Mutation at position A7566G was detected in prototype-like (P-L) (II) variants. Another nt change at position T7709G was detected in the non P-L (II) variants. The mutation T7760C was detected in both the non P-L (I) and non P-L (V) variants. Another mutation A7766G was also detected in the non P-L (V) variants.

There were some mutations observed in transcription factor binding regions of the LCR. Mutations T7422A and G7508A were found in the YY1 binding site, mutations T7528C and C7810T were found in the E2 binding site. The following three nucleotide substitutions; A7655G, C7756T and G7758A were found in NF-1binding site while T7538C and T7648G were found in the Oct-1 and AP-1 transcriptional binding sites respectively (Wyant *et al*., 2011).

Table 3.9 Nucleotide mutations identified within the LCR of HPV 53

Base position in HPV 53R		7398	7422	7428	7432	7508	7528					7531	7533	7534	7535	7536	7537	7538	7566	7608	7609	7618	7619	7634	7648	7655	7656	7676	7685	7709	7751	7756	7758	7760	7766	7784	7797	7810	Total
Base in HPV53R sequence		A	T	C	C	G	T	-	-	-	-	G	A	T	A	T	C	T	A	G	T	T	T	T	T	A	G	C	C	T	A	C	G	T	A	C	A	C	29
Variant	**Sample Identity**																																						
P-L (II)	**308-1**					A													G																				2
	308-2					A													G																				2
P-L (I)	**339-1**					A																																	1
	339-2					A																																	1
Non P-L (I)	**364-1**		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T			T	A	C		T	C	T	23
	364-2		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T			T	A	C		T	C	T	23
Non P-L (II)	**408-1**		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T	G		T	A			T	C	T	23
P-L (I)	**408-2**					A																																	1
Non P-L (III)	**461-1**		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T			T	A			T	C	T	22
P-L (I)	**461-2**					A																																	1
P-L (III)	**501-1**					A																		C					T									3
	501-2					A																		C					T									3
Non P-L (III)	**529-1**		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T			T	A			T	C	T	22
	529-2		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T			T	A			T	C	T	22
Non P-L (III)	**571-1**		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T			T	A			T	C	T	22
	571-2		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T			T	A			T	C	T	22
Non P-L (IV)	**669-1**	G				A	C	T	G	G	G	A	-	-	-	-	-			A	G	G	C	C	G	G	T		T		G	T	A			T		T	27
	669-2	G				A	C	T	G	G	G	A	-	-	-	-	-			A	G	G	C	C	G	G	T		T		G	T	A			T		T	27
Non P-L (I)	**683-1**		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T			T	A	C		T	C	T	23
	683-2		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T			T	A	C		T	C	T	23
P-L (II)	**696-1**					A													G																				2
	696-2					A													G																				2
Non P-L (V)	**702-1**		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T			T	A	C	G	T	C	T	24
	702-2		A	T	G	A	C	T	G	G	G	A						C				G	C	C		G		G	T			T	A	C	G	T	C	T	24

Single nucleotide base differences within HPV 53 LCR variants in comparison to the prototype sequence HPV 53R, GenBank accession number X74482. Sample identity denotes the specimen number and the letters (I-V) refer to changes within the variants analysed in the various specimens. Total indicates the overall number of mutations in each specimen analysed. The green highlighted base pair positions denote the position where the most frequent mutation occurred among all the isolate.

3.4.5 Phylogenetic classification of HPV 16, 58 and 53 variants

HPV 16, 58 and 53 phylogenetic trees were generated in order to classify the LCR sequences into their respective lineages. Reference sequences for the respective HPV variants were obtained from GenBank. The LCR sequence alignments are included in the appendix (Figures 1-3).

Reference sequences for HPV 16 LCR representing different phylogenetic branches/lineages were as following; AF402678 (AA), AF534061 (As), AF472508 (Afr1), AF472509 (Afr2) and prototype (Eur) reference sequence retrieved from Papillomavirus Episteme or PAVE (van Doorslaer *et al*., 2013). From the phylogenetic tree, 4 distinct lineages were observed and these corresponded to AA, Afr1, Afr2 and Eur-As variants (Figure 3.7), similar observations were noted elsewhere (Tanzi *et al*., 2009).

The HPV 16 variants identified in this study were separated into distinct lineages based on the presence of diagnostic single nucleotide polymorphisms (SNPs) and phylogenetic analysis. Several diagnostic SNPs within the HPV 16 LCR and E6 genes have been identified that can be used to specifically distinguish some of the variant lineages (Cornet *et al*., 2012). The diagnostic SNPs covered by the LCR PCR products in this study are given in Table 3.10 (Cornet *et al*., 2012). The major European-Asian branch of HPV 16 variants can be distinguished from the African and Asian-American branches by three LCR nucleotide positions (G7489, C7764 and C7786). These SNPs were present in the LCR sequences from samples 368-1, 368-2, 529-1, 529-2, 571-1, 571-2, 626-1, 626-2, 727-2 and 811-1. The classification of the HPV 16 variants from these samples as belonging to the Eur-Asian lineage is well supported by the maximum-likelihood tree in Figure 3.7 While the Eur lineage cannot be distinguished by any SNPs, the As lineage has six diagnostic SNPs in the LCR (T7177C, T7201C, C7270T, A7287C, G7842A/T and C24T). The HPV 16 LCR sequences for 368-1, 368-2, 529-1, 529-2, 571-1, 571-2, 626-1, 626-2, 727-2 and 811-1 (covering nts 7308-31) do not have G7842A/T and C24T typical of As, and were therefore classified in this study as Eur.

Table 3.10 Diagnostic and common single nucleotide polymorphisms in the long control region of HPV 16 (modified from Cornet *et al*., 2012).

Lineage	Sublineage	LCR nucleotide position																									
		7287	**7339**	**7348**	**7394**	**7395**	**7435**	**7485**	**7489**	**7507**	**7669**	**7689**	**7729**	**7730**	**7743**	**7764**	**7786**	**7826**	**7834**	**7837**	**7839**	**7842**	**7876**	**7886**	**24**	**25**	**31**
HPV16 ref		A	A	A	C	C	G	A	G	A	C	C	A	A	T	C	C	G	G	A	A	G	C	C	C	T	C
EAS	Eur	-	—	—	—	—	—	—	—	—/C	—	—	—	—/C	—	—	—	—	—	—	—	—	—	—	—/G	—	—
	As	C	—	—	—	—	—	—	—	—	—	—	—	C	—	—	—	—	—	—	—	**A/T**	—	—	**T**	—	—
AFr1	Afr1a	-	—	—	—	—	—	—	A	—	—	A	—	—	—	T	T	—	T	—	—	—	—/A	—	—	—	T
	Afr1b	-	—	**C**	—	—	—	—	A	—	—	A	—	—	—	T	T	—	T	—	—	—	A	—	—	**C**	T
AFr2	Afr2a	-	—	—	—	—	**A**	C	A	—	T	—/A	—	—	—	T	T	A	—/T	C	**G**	—	—	—	—	—	T
	Afr2b	-	—	—	—	—	—	C	A	—	T	A	—	—	—	T	T	A	T	C	—	—	—	—	—	—	T
NA/AA	NA	-	—	—	—	—	—	C	A	—	T	A	—/C	—	—	T	T	—	T	—	—	—	—	—	—	—	—
	AA1	-	T	—	T	T	—	C	A	—	T	A	C	—	**G**	T	T	—	—	—	—	—	—	G	—	—	—
	AA2	-	T	—	T	—/T	—	C	A	**G**	T	A	C	—	—	T	T	—	—	—	—	—	—	G	—	—	—

[a]Includes only nucleotide positions that can distinguish at least two of the nine identified sublineages from each other. Dashes indicate no nucleotide exchange from the HPV16 reference sequence. Nucleotides separated by a slash are different nucleotides that can be in the given position for the given sublineage. Diagnostic SNPs are in boldface.[b]EAS, European-Asian; EUR, European; As, Asian; AFr, African; NA, North American; AA, Asian-American.

The major African branch can be specifically distinguished from the European-Asian branch by the presence of a C31T SNP in the LCR. This SNP was found in the LCR sequences for specimens 440-1, 440-2, 460-1, 460-2, 501-1, 501-2, 511-1, 511-2, 621-1, 621-2, 727-1 and 811-2. The classification of the HPV 16 isolates from these samples clustered into the African lineage and is well supported by the maximum-likelihood tree in Figure 3.7. The two Afr sublineages (Afr1 and Afr2) can be separated based on four SNPs in the LCR, where the Afr2 lineage uniquely has A7485C, C7669T, G7826A and A7837C (Cornet *et al*., 2012). These SNPs were found in the LCR sequences for 440-1, 440-2, 460-1, 460-2, 501-1, 501-2, 511-1, 511-2, 621-1, 621-2, 727-1 and 811-2, and they were therefore classified as Afr2. The presence of the diagnostic SNP A7839G in these LCR sequences

further allowed classification into the Afr2a sublineage. Although the HPV 16 isolates from 511-1 and 511-2 could be classified as Afr2a based on diagnostic SNPs the LCR sequence had two additional nucleotide changes not present in the Afr2a reference sequence (A7372C and T7807C) and they therefore clustered separately in the phylogenetic tree (Figure 3.7). The LCR sequences from couple 181 clustered with the Asian-American reference sequence in the phylogenetic tree (Figure 3.7) and while no SNPs in the LCR can specifically distinguish AA variants (Cornet *et al*., 2012) they have the six SNPs common to this lineage (A7485C, G7489A, C7669T, C7689A, C7764T and C7786T).

HPV 16 phylogenetic classification of LCR variants revealed that the majority half (12/24) of HPV 16 variants belonged to the Afr2a sublineage and 41.67% (10/24) Eur lineage while only 8.33% (2/24) belonged to the AA lineage. None of the LCR sequences identified in this study belonged to the As and Afr1 branch.

HPV 58 LCR reference sequences representing different phylogenetic lineages were as following; HQ339045 (A1), HQ338950 (A2), QH339150 (A3), HQ339176 (B1), HQ339309 (B2), HQ339310 (C), HQ338954 (D1) and HQ339312 (D2). HPV 58 phylogenetic tree revealed 4 distinct branches, namely branch (A) divided into A1, A2, A3; branch (B) subdivided into B1 and B2, branch (C) and finally branch (D) subdivided into D1 and D2 (Figure 3.8). The HPV 58 variants identified in this study were separated into distinct lineages based on the presence of diagnostic SNPs and phylogenetic analysis. Chan and co-workers (2011) identified several diagnostic SNPs or signature sequences in the LCR that can be used to separate the A1, A2, A3, B1, B2, C, D1 and D2 lineages of HPV 58 (Table 3.11). A maximum likelihood phylogenetic tree was generated using the LCR sequences from the HPV 58 positive couples and reference sequences for each lineage (Figure 3.8).

Table 3.11 Diagnostic and common single nucleotide polymorphisms in the long control region of HPV 58 (modified from Chan *et al*., 2011).

Lineage	LCR- 173 bp												LCR- 337 bp											
	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7		
	2	2	2	2	3	3	3	3	3	3	4	4	5	6	6	7	7	7	7	7	7	7		
	5	6	7	8	0	1	3	4	6	9	2	2	4	1	8	1	3	4	5	7	7	9	3	5
	7	6	7	4	4	3	2	5	9	5	1	9	0	9	6	4	0	5	5	8	9	2	0	2
Prototype	T	C	C	C	A	A	G	T	T	G	G	T	A	G	G	C	C	G	A	T	A	C	C	C
A1	-	-	-	-	-	-	-	-	-	-	-	-	R^7	-	-	-	-	-	-	-	-	-	-	-
A2	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-
A3	-	-	IN	-	G	-	-	-	-	-	-	-	-	-	-	C	-	-	R^8	-	-	-	-	T
B1	G	-	-	S^5	-	C	-	-	-	-	-	-	G	A	-	-	-	-	-	-	-	-	G	T
B2	-	-	-	G	-		-	-	-	-	-	-	-	-	-	-	A	C	-	-	G	-	G	K^{10}
C	-	-	-	G	-	-	-	C	K^6	-	-	-	-	-	-	-	A	-	-	-	R^9	-	G	T
D1	-	-	-	G	-	-	A	-	-	-	-	G	-	-	-	-	A	-	-	C	G	T	G	T
D2	-	-	-	G	-	-	A	-	-	A	A	G	-	-	A	-	A	-	-	C	G	-	G	T

A dash indicates the same nucleotide as in the prototype sequence. IN, indicates an insertion of 12 bp. for S^5, 67% isolates have C and 33% have G; for K6, 86% isolates have G and 14% have T; for R7, 78.6% isolates have A and 21.4% have G; for R8, 96% isolates have G and 4% have A; for R9, 86% isolates have G and 14% have A; and for K10, 67% isolates have T and 33% have G.

The majority of the LCR sequences (9/16) belonged to the A2 lineage. The LCR from couples 153, 511, 529, 623 and sample 683-2, clustered with the A2 reference sequence (Figure 3.8) and had the signature sequence for the A2 lineage (C7266T, (A/C)7714G). This lineage is known to dominate in Africa (Chan *et al*., 2011). Additional SNPs were present in the LCR from couple 529 (C7265G and G7421A) and 683-2 (C7265G and T7786G). A total of five of the LCR sequences clustered with the lineage C reference sequence in the phylogenetic tree (couples 460 and 762 as well as sample 683-1) and had the SNPs diagnostic for this lineage (C7284G, T7345C, T7369G, C7730A, A7779G, C30G and C52T). The LCR from the remaining couple (313) clustered with the B2 lineage (Figure 3.8) and had the diagnostic SNPs for this lineage (C7284G, C7730A, G7745C, A7779G, C30G, and C52G).

Two reference HPV 53 sequences were used for the construction of HPV 53 phylogenetic tree, representing the prototype and non-prototype lineages.

The reference sequence X74482 was used to represent the prototype branch and AY949073 sequence was used to represent the non-prototype branch (Kocjan *et al*., 2007). The HPV 53 LCR phylogenetic tree revealed three separate branches based on the LCR (Figure 3.9). Half 50% (12/24) of HPV 53 analysed sequences belonged to the non P-L branch whereas 41.67% (10/24) sequences belonged to the P-L branch. Two sequences from couple 669 belonged to a separate group in with significant bootstrap values.

The two clusters of variants (non P-L and P-L) can be distinguished by a 4bp (TGGG) insertion at position 7530 in the LCR of non-protypelike variants highlighted in green (Figure 3.10). This insertion was present in the majority 58.33% (14/24) of the LCR couples; 364, 529, 571, 683, 702 and samples 408-1 and 461-1. The LCR from couple 669 had a novel deletion of 5 bases at positions 7533-7537 highlighted in green that has not been reported to date (Figure 3.10), and does not cluster with the non P-L or P-L variants (Figure 3.10).

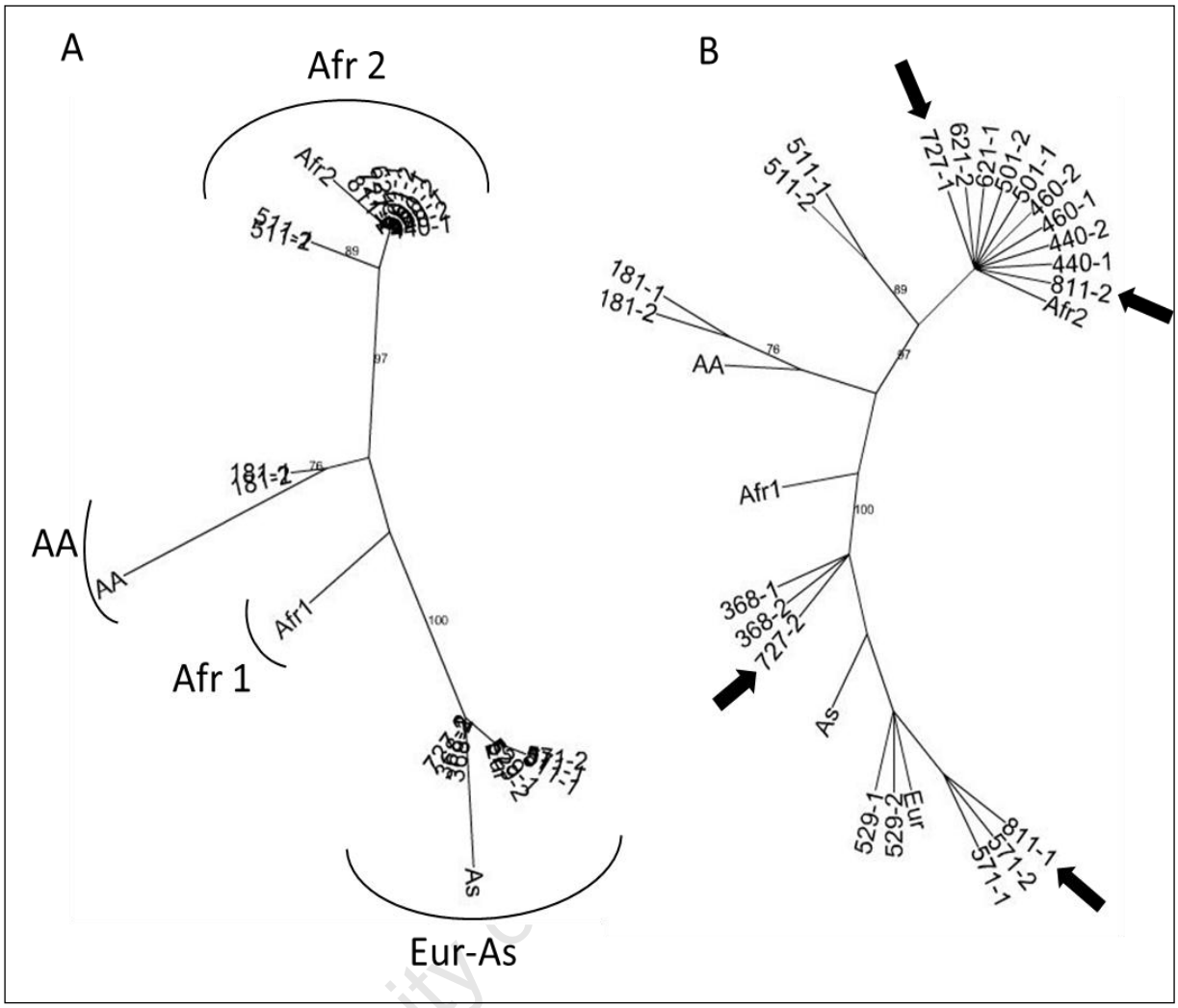


Figure 3.7 Maximum likelihood phylogenetic tree of LCR sequences (645 nts) from HPV 16 positive couples. Bootstrap support for the unrooted tree is shown as a percentage of 1000 replicates, with values greater than 70% shown. (A) The LCR reference sequences for the different lineages were retrieved from GenBank with the following accession numbers (Afr1 (AF472508) Afr2 (AF472509), As (AF534061), AA (AF402678). The corrected European reference sequence (Eur) was retrieved from Papillomavirus Episteme or PAVE (van Doorslaer *et al*., 2013). (B) The same tree with branches with less than 40% support collapsed and branch length ignored. Arrows indicate the position of isolates from couples that segregated into different major branches.

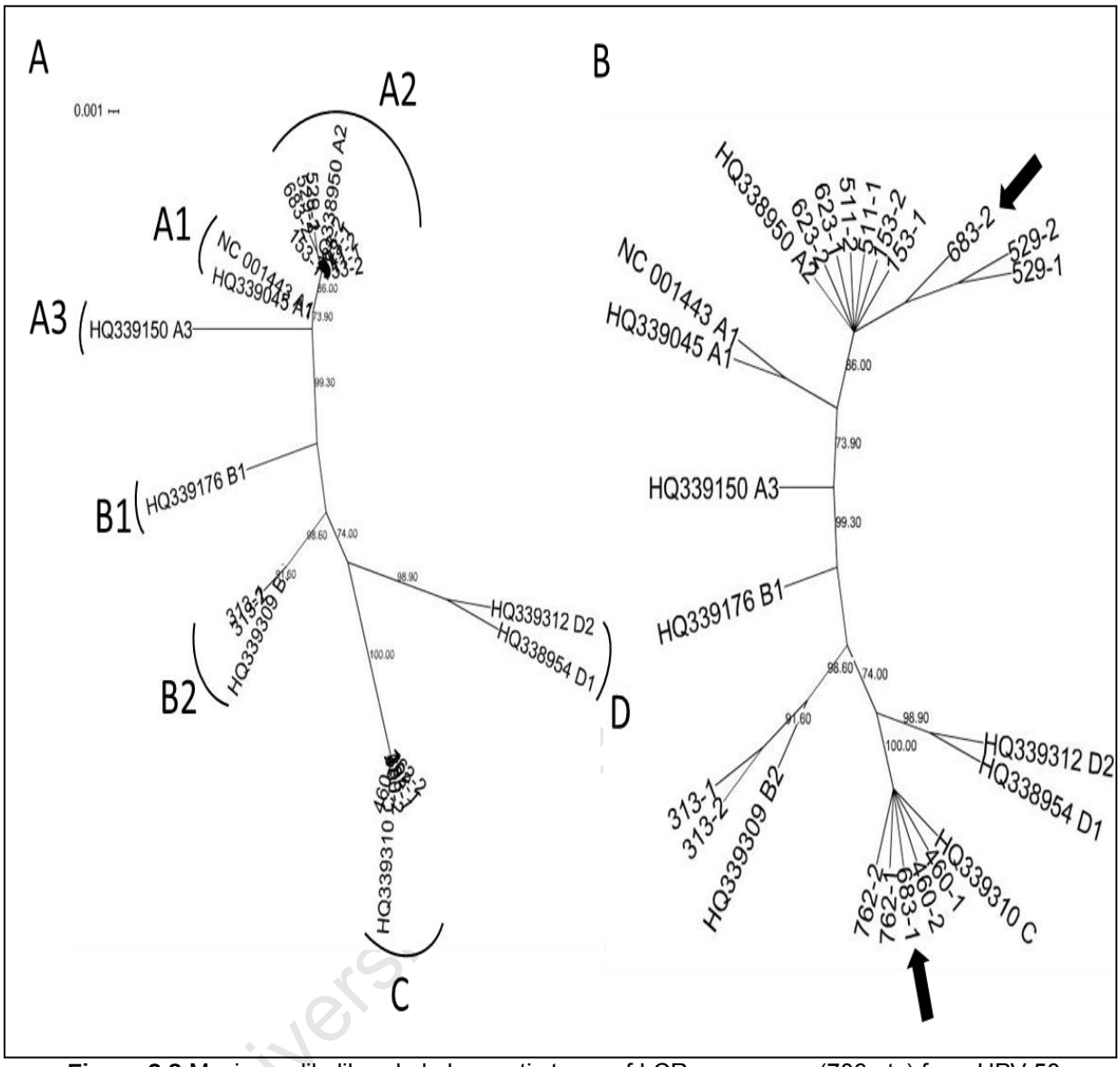


Figure 3.8 Maximum likelihood phylogenetic trees of LCR sequences (706 nts) from HPV 58 positive couples. Bootstrap support for the unrooted tree is shown as a percentage of 1000 replicates, with values greater than 70% shown. (A) The LCR reference sequences for the different lineages were retrieved from GenBank with the following accession numbers; HQ339045 (A1), HQ338950 (A2), QH339150 (A3), HQ339176 (B1), HQ339309 (B2), HQ339310 (C), HQ338954 (D1) and HQ339312 (D2). (B) The same tree with branches with less than 40% support collapsed and branch length ignored. Arrows indicate the position of isolates from couples that segregated into different major branches.

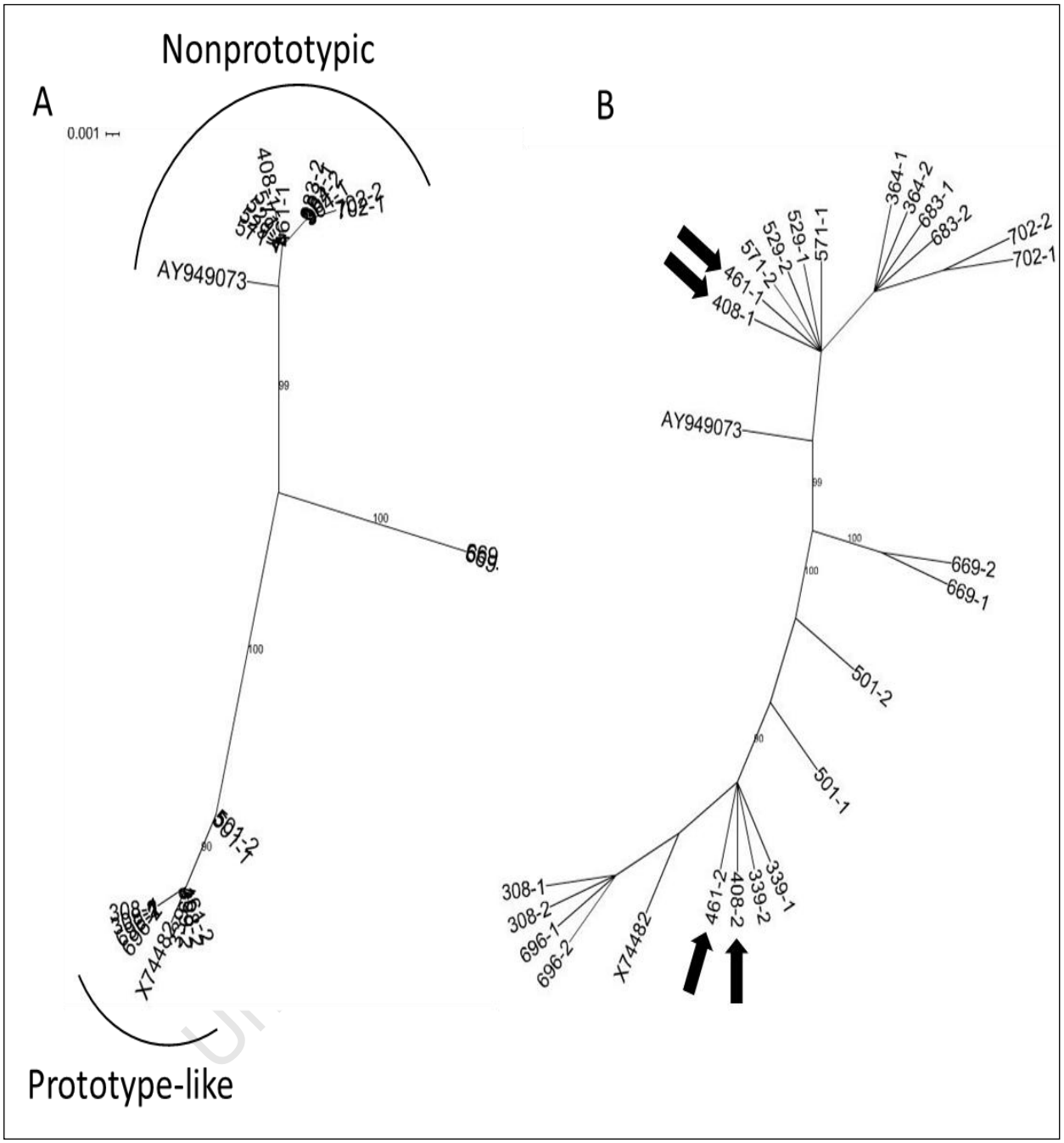


Figure 3.9 Maximum likelihood phylogenetic tree of LCR sequences (417 nts) from HPV 53 positive couples. Bootstrap support for the unrooted tree is shown as a percentage of 1000 replicates, with values greater than 70% shown. (A) The LCR reference sequences for the different lineages were retrieved from GenBank with the following accession numbers; X74482 (P-L) (Prototype-Like) and AY949073 (Non (P-L) Prototype-like). (B) The same tree with branches with less than 40% support collapsed and branch length ignored. Arrows indicate the position of isolates from couples that segregated into different major branches.

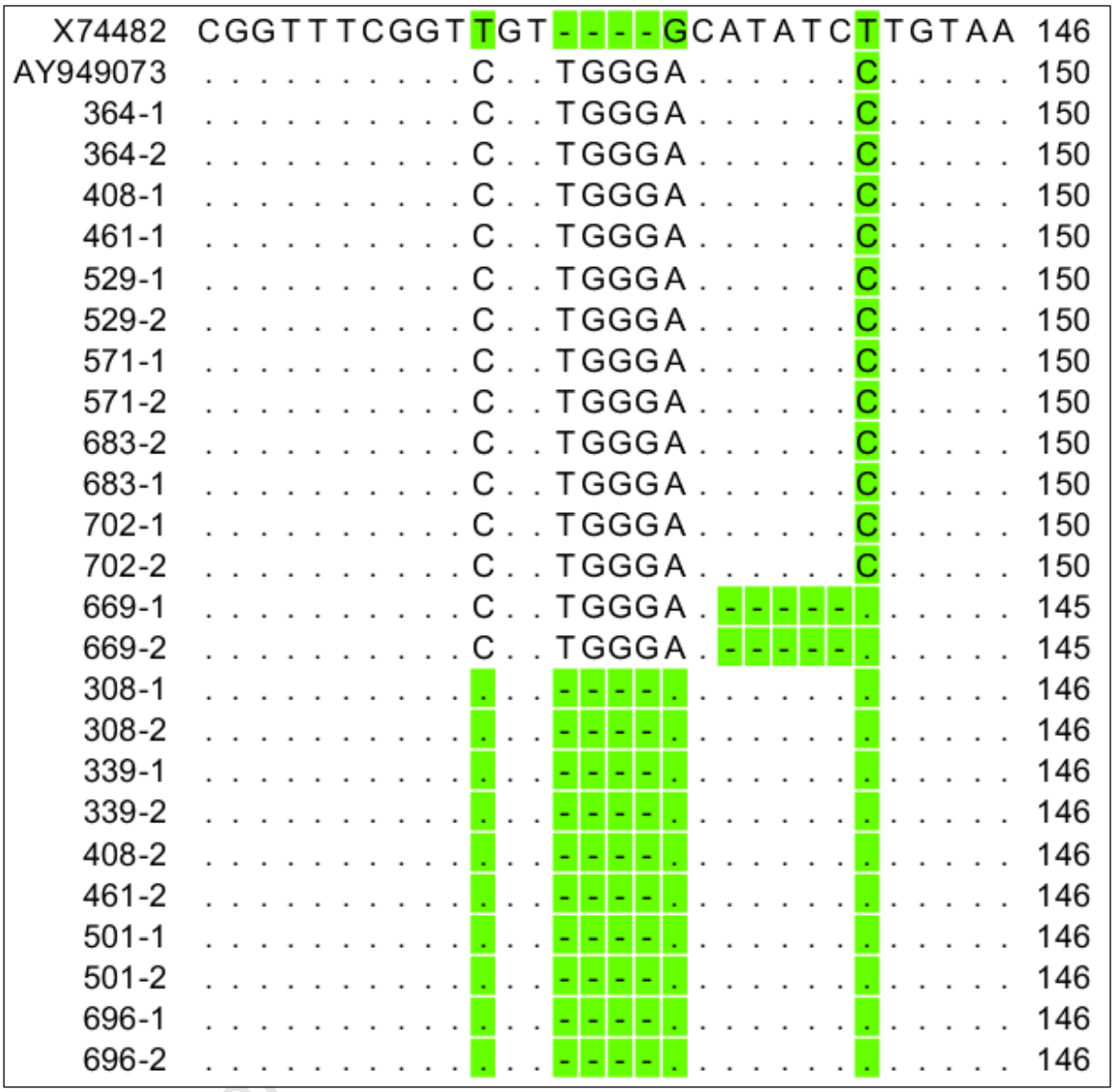


Figure 3.10 HPV 53 diagnostic single nucleotide polymorphisms (SNPs) distinguishing the P-L variants from non P-L variants. Non P-L variants have a T7528C nucleotide substitution, 4bp (TGGG) insertion at position 7530 and a T7538C in the LCR. Couple 669 had a unique five bases deletion at positions 7533-7537.

3.4.6 HPV 16 variants in couples

Of the 12 couples sharing HPV 16, ten were variant-specific concordant, sharing the same variant. The analysis of LCR variants showed that the majority (5/12; 41.7%) of couples shared the Afr2a sublineage. From the five couples sharing the Afr2a sublineage, four couples (440, 460, 501 and 621) were concordant with the Afr2a (I) sublineage. While partners in couple 511 were concordant with a different Afr2a sublineage termed Afr2a (II). The

difference between the two Afr2a (I) and Afr2a (II) sublineage is that, Afr2a (I) has the mutation G7385C while Afr2a (II) has A7372C and T7807C. Couple 181 were found to share the AA variant.

Four out of twelve (4/12; 33.3%) couples were concordant with the Eur lineage. Both partners in couples 529 and 626 were concordant with the Eur prototype variant. Couple 368 were concordant with the Eur (I) variant and the partners in couple 571 shared the Eur (II) variant. The difference between Eur (I) and Eur (II) variant is that, Eur (I) has a mutation at the position G7521A, while Eur (II) has a mutation at position A7314C.

Two couples (2/12; 16.7%) were found to be discordant. The male partner in couple 727 had the Eur (I) variant while his respective female partner has the Afr2a (I) variant. In couple 811 the male partner had the Afr2a (I) variant while his female partner had the Eur (II) variant (Figure 3.7).

3.4.7 HPV 58 variants in couples

Of the eight couples sharing HPV 58, seven were variant-specific concordant, sharing the same variant. The phylogenetic classification of HPV 58 LCR variants showed that half (4/8; 50%) of the couples shared the A2 lineage. Partners in three couples; (153, 511 and 623) were concordant with the A2 (I) variant. Partners in couple 529 both had the A2 (II) variant. The difference between the two A2 variants is that, the A2 (II) has two extra mutations at positions C7265G and G7421A.

Partners in couples 460 and 762 (2/8; 25%) were concordant with the C variant and partners in couple 313 (1/8; 12.5%) were concordant with the B variant. The only HPV 58 variant discordant couple was couple 683; here the female partner had the C variant, while her male partner had the A2 (III) variant. The A2 (III) variant differs from the A2 (II) variant by having a T7786G mutation whereas the A2 (II) variant has a G7421A mutation.

3.4.8 HPV 53 variants in couples

Of the 12 couples sharing HPV 53, ten were variant-specific concordant, sharing the same variant. The phylogenetic classification of HPV 53 LCR variants showed that (6/12; 50%) of couples shared the non P-L lineages. From the six couples sharing the non prototype-like lineage; partners from couples 364 and 683 shared the non P-L (I) variant. Partners in two other couples; 529 and 571 were concordant with the non P-L (III) variant. In couple 669, partners were concordant with the non P-L (IV) variant and finally in couple 702 shared the non P-L (v) variant. The non P-L (I) variant differs by the mutation T7760C to the non P-L (II) variant which has T7709G mutation. The non prototype-like (I) variants differs by one extra/additional mutation in the position T7760C which is absent in the non P-L (III) variant. The non P-L (IV) variant is clearly distinguishable from all other non P-L variant in numerous positions. Among these positions are the mutations A7398G, five deletions at positions (A7533, T7534, A7535, T7536, C7537), G7608A, T7609G, T7648G, G7656T and A7751G. The non P-L (V) variant differs by one extra nucleotide at position A7766G from the non P-L (I) variant.

Four couples (4/12; 33%) were concordant with the P-L lineages, and in these four couples; partners in couple 339 were concordant with the P-L (I) variant. Partners in couples; 308 and 696 were concordant with the P-L (II) variant. Partners in couples 501 shared the P-L (III) variant. The P-L (II) variant differs by an extra/additional mutation at position A7566G from the P-L (I) variant. The P-L (III) variant differs to the P-L (I) variant by two extra/additional mutations at positions T7634C and C7685T.

Two couples were HPV-58 variant specific discortdant. Partners in couples 408 and 461 were discordant, with the female partner in couple 408 having the non P-L (II) variant while her respective male partner has the P-L (I) variant. Lastly in couple 461 the female partner had the non P- L (III) variant and her respective male partner had the P-L (I) variant.

3.4.9 HPV variants stratified according to cervical cytology

The HPV 16, 58 and 53 variants found in the cervical samples are summarised in Table 3.12 together with the cervical cytology data. A total of 25 females were investigated for HPV 16, 58 and 53 variants. A total of nineteen women had single HPV type variant detected, five women had two HPV type variants detected; and the respective women had normal, ASCUS or LSIL. Only one woman (highlighted in green) had variants of all three HPV types investigated and was diagnosed with HSIL (Table 3.12).

Table 3.12 HPV types 16, 58 and 53 variant distribution according to cervical cytology

Female partner	HPV types	HPV variant	Cervical cytology
313-1	58	B2	Normal
339-1	53	P-L	Normal
368-1	16	Eur	Normal
408-1	53	Non P-L	Normal
460-1	16 58	Afr2/C	Normal
461-1	53	Non P-L	Normal
626-1	16	Eur	Normal
683-1	53 58	C/non P-L	Normal
702-1	53	Non P-L	Normal
762-1	58	C	Normal
364-1	53	Non P-L	ASCUS
501-1	16 53	Afr2/P-L	ASCUS
669-1	53	Non P-L	ASCUS
153-1	58	A2	LSIL
181-1	16	AA	LSIL
308-1	53	P-L	LSIL
440-1	16	Afr2	LSIL
511-1	16 58	Afr2/A2	LSIL
571-1	16 53	Eur/non P-L	LSIL
621-1	16	Afr2	LSIL
623-1	58	A2	LSIL
696-1	53	P-L	LSIL
727-1	16	Afr2	LSIL
811-1	16	Eur	LSIL
529-1	16 53 58	Eur /A2/non P-L	HSIL

3.5 Discussion

The analysis of sequence variants of HPV16 in this study revealed the presence of African 2a, Asian American and European variants in these South African men and women, with the African lineage, as expected, being the most prevalent. None of the HPV 16 variants belonged to the African1, Asian and North American lineages/sublineages. Since we only sequenced a

partial region of the LCR we were unable to precisely rule out the presence of the North American sublineage in our study. This is because in order to identify the North American sublineage we would also have to additionally include the E6 region, since the E6 is needed to identify North American sublineage (Tanzi *et al*., 2009, Yamada *et al*., 1997). This study is in agreement with the observation that, although HPV 16 phylogenetic analyses show a trend towards ethnogeographic separation of variants, some variants do exist in other regions; although not as prevalent as in their original geographic location (Tu *et al*., 2006). The observed phenomenon may be due to human migrations (Ho *et al*., 1993b; Stewart *et al*., 1996; Yamada *et al*., 1997).

HPV 58 phylogenetic tree revealed 4 distinct lineages (A-D) with some lineages further subdividing into smaller branches and similar results have been reported elsewhere (Chan *et al*., 2011). Phylogenetic tree of HPV 58 showed that 56.25% of HPV 58 LCR belonged to the A lineage, specific to A2 sublineage. Lineage A is has been shown to be the most prevalent lineage with the A2 sublineage being the most common (Chan *et al*., 2011).

HPV 53 variants have been found to cluster into two distinct branches (Cento *et al*., 2012; Kocjan *et al*., 2007; Oliveira *et al*., 2012; Prado *et al*., 2005; Wyant *et al*., 2011). However we identified a novel sublineage (from couple 669) with a unique five bases deletion at positions 7533-7537 that formed a new branch.

In the only published study to have determined HPV variant concordance in couples, Ho and co-workers (1993) revealed that HPV 16 variants were shared in four out of eight couples. Variants in the other four couples were discordant (Ho *et al*., 1993a). In this study a high level of variant specific concordance was observed between sexual partners due to the fact that HPV 16 and HPV 53 variants were shared in ten out of twelve couples. HPV 58 variants were shared in seven out of eight couples.

In couples that were discordant for HPV variants, there are several possible explanations as to why the partners in those couples have discordant variants. One possible explanation is that both partners could have been initially infected by different variants that have established themselves to exclude other late arriving variants. Another possible explanation is that the former or latter variant may be latent (Ho *et al*., 1993a). A similar phenomenon was noted in a study done by Jiang et al (2009) where the genomic diversity of HPV 16 variants was determined from two specimens obtained from one individual at different time intervals. At enrolment ten clones were analysed and sequencing results showed that eight clones contained identical sequences to the Eur variant. The remaining two clones contained different variations of the Eur-Afr2 sequence. Although the mutations in the two clones were similar and corresponded significantly to some base positions found in LCR of Afr2 variants. However the remaining E6 regions of the recombinants were identical to the Eur variant and therefore the Eur-Afr2 recombinants were considered to be Eur variants. At the second visit, 20 months later, neither the Eur variant nor the Eur-Afr2 recombinants were detected, instead only the Afr2 variant was detected. A possible explanation of the disappearance of the Eur variant as well as the recombinants may be due to clearance of the variants or suppression to a low level below detectable threshold. The detection of the Afr2 variant during the follow up sample may represent reactivation of a latent infection. Alternatively the infection may represent a separate infection that might have been acquired from the concurrent or different sex partner (Jiang *et al*., 2009).

The LCR contains transcription factor binding sites which activate transcription of the E6/E7 oncogenes (Hoppe-Seyler *et al*., 1991; Sibbet *et al*., 1995). Mutations occurring in the LCR may interfere with the activities of transcription factors which could potentially have significant role in tumorigenesis (Stephen *et al*., 2000). Therefore sequence variations in LCR could possibly enhance viral oncogenes expression and represent risk factors for development of cervical cancer. A link connecting transcriptional

factor binding site to virulence of the different HPV variants has been reported (Chansaenroj *et al*., 2012). It has been shown that overexpression of transcriptional factors such as octamer-1 (Oct-1) inhibit the promoter activities of HPV 16 and 18 (Hoppe-Seyler *et al*., 1991; Sibbet *et al*., 1995). HPV GRE has the ability to regulate the transcription of E6/E7 oncogenes (Chan *et al*., 1989). YY1 interferes and may prevent the LCR activity responsible for the expression of the E6/E7 oncogenes (Park *et al*., 1999; Stephen *et al*., 2000). NF1 plays a role in viral gene expression by activating viral regulator elements (Baldwin *et al*., 2007; Chen *et al*., 1996b). In essence mutation in the above mentioned transcriptional binding sites maybe the reason to why some variants have higher oncogenic potential compared to other variants. In this study mutations were found in the following transcriptional binding sites of HPV 16; GRE, YY1, NF1, TEF1, Oct1 and Sp1. Mutations were found in the following transcriptional binding sites of HPV 53; YY1, E2 binding site, NF-1binding site, Oct-1 and AP-1. Only one mutation was found in HPV 58 transcriptional binding site and that was at NF1. The effects of these mutations on the promoter activity of the LCR could be assessed using a luciferase reporter assay. By doing so, it would enable us to determine if these mutations result in different promoter activity. Differences in promoter activity may reflect differences in the transcription of the E6/E7 oncogenes and may be a factor in the oncogenic potential of the variant.

3.6 Conclusion

Nucleotide variations in the LCR of HPV types; 16, 58 and 53 can be used to investigate the transmission of HPV between sexual partners. This is because comparison of these HPV infections in couples confirms the efficient sexual exchange of variants. Ten of twelve couples analysed were concordant for HPV 16 lineages. Seven of eight couples were concordant for HPV 58 lineages and ten of twelve couples were concordant for HPV 53 lineages. HPV sharing between partners may be lower than reported due to the fact, that two out of twelve couples were discordant for HPV 16 lineages.

One out of eight couples was discordant for HPV 58 lineages and two out of twelve couples were discordant with HPV 53 lineages. Therefore future studies need to additionally focus on HPV variant concordance when assessing the transmission of HPV instead of only focussing on type-specific HPV concordance. In addition, more samples need to be analysed for each couple to detect rare/low copy number variants.

Further studies need to have a large enough sample size to further investigate whether the increase virulence observed with non-Eur variants in various populations is caused by differences in virulence of different HPV 16 variants or is a result of the genetic differences of different populations to mount an effective immunological response against a specific HPV 16 variant. This is because several studies have found the Eur variant to be more oncogenic compared to the non-Eur variant in various populations (Bokal, *et al*., 2010; Tu *et al*., 2006). We could not clarify such phenomenon in the current study due to several limitations such as small sample size and we did not clone the PCR product prior to sequencing therefore could not identify infections harbouring multiple variants of genotypes.

REFERENCES

1. **Abalos, A. T., R. B. Harris, A. G. Nyitray, A. Mitchell, P. A. Thompson, A. R. Giuliano, and F. Garcia**. 2012. Human papillomavirus type distribution among heterosexual couples. J. Low Genit. Tract. Dis. **16**:10-15.
2. **Adler, D., B. G. de, F. Laher, G. Gray, B. R. Allan, and A. L. Williamson**. 2008. Human papillomavirus genotype distribution among human immunodeficiency virus (HIV)-infected and non-HIV-infected women in Soweto, South Africa. J. Clin. Microbiol. **46**:4109-4110.
3. **Alam, S., M. J. Conway, H. S. Chen, and C. Meyers**. 2008. The cigarette smoke carcinogen benzo[a]pyrene enhances human papillomavirus synthesis. J. Virol. **82**:1053-1058.
4. **Alencar, T. R., D. M. Cerqueira, M. R. da Cruz, P. S. Wyant, E. D. Ramalho, and C. R. Martins**. 2007. New HPV-16 European and non-European variants in Central Brazil. Virus Genes **35**:1-4.
5. **Allan, B., D. J. Marais, M. Hoffman, S. Shapiro, and A. L. Williamson.** 2008. Cervical human papillomavirus (HPV) infection in South African women: implications for HPV screening and vaccine strategies. J. Clin. Microbiol. **46:**740-742.
6. **Ammatuna, P., L. Giovannelli, D. Matranga, S. Ciriminna, and A. Perino**. 2008. Prevalence of genital human papilloma virus infection and genotypes among young women in Sicily, South Italy. Cancer Epidemiol. Biomarkers Prev. **17**:2002-2006.
7. **Appleby, P., V. Beral, G. A. Berrington de, D. Colin, S. Franceschi, A. Goodill, J. Green, J. Peto, M. Plummer, and S. Sweetland**. 2006. Carcinoma of the cervix and tobacco smoking: collaborative reanalysis of individual data on 13,541 women with carcinoma of the cervix and 23,017 women without carcinoma of the cervix from 23 epidemiological studies. Int. J. Cancer **118**:1481-1495.
8. **Arnouk, H., M. A. Merkley, R. H. Podolsky, H. Stoppler, C. Santos, M. Alvarez, J. Mariategui, D. Ferris, J. R. Lee, and W. S. Dynan**. 2009. Characterization of Molecular Markers Indicative of Cervical Cancer Progression. Proteomics. Clin. Appl. **3** :516-527.
9. **Aubin, F., J. L. Pretet, A. C. Jacquard, M. Saunier, X. Carcopino, F. Jaroud, P. Pradat, B. Soubeyrand, Y. Leocmach, C. Mougin, and D. Riethmuller**. 2008. Human papillomavirus genotype distribution in external acuminata condylomata: a Large French National Study (EDiTH IV). Clin. Infect. Dis. **47**:610-615.
10. **Ault, K. A.** 2006. Epidemiology and natural history of human papillomavirus infections in the female genital tract. Infect. Dis. Obstet. Gynecol. **2006 Suppl**: 40470.
11. **Autier, P., M. Coibion, F. Huet, and A. R. Grivegnee**. 1996. Transformation zone location and intraepithelial neoplasia of the cervix uteri. Br. J. Cancer **74**:488-490.
12. **Baldwin, A., M. K. Hypes, L. Pirisi, and K. E. Creek**. 2007. NFI is an essential positive transcription factor for human papillomavirus type 16 early gene expression. Open. Virol. J. **1**:33-38.
13. **Benevolo, M., M. Mottolese, F. Marandino, M. Carosi, M. G. Diodoro, S. Sentinelli, P. Visca, F. Rollo, L. Mariani, G. Vocaturo, R. Sindico, I. Terrenato, R. P. Donnorso, and A. Vocaturo**. 2008. HPV prevalence among healthy Italian male sexual partners of women with cervical HPV infection. J. Med. Virol. **80**:1275-1281.
14. **Bernard, H. U., R. D. Burk, Z. Chen, D. K. Van, H. Hausen, and E. M. de Villiers**. 2010. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. Virology **401**:70-79.

15. **Bertram, C. C.** 2004. Evidence for practice: oral contraception and risk of cervical cancer. J. Am. Acad. Nurse Pract. **16**:455-461.

16. **Berumen, J., R. M. Ordonez, E. Lazcano, J. Salmeron, S. C. Galvan, R. A. Estrada, E. Yunes, A. Garcia-Carranca, G. Gonzalez-Lira, and C. A. Madrigal-de la**. 2001. Asian-American variants of human papillomavirus 16 and risk for cervical cancer: a case-control study. J. Natl. Cancer Inst. **93**:1325-1330.

17. **Bhattacharjee, B. and S. Sengupta**. 2006. HPV16 E2 gene disruption and polymorphisms of E2 and LCR: some significant associations with cervical cancer in Indian women. Gynecol. Oncol. **100**:372-378.

18. **Bleeker, M. C., C. J. Hogewoning, J. Berkhof, F. J. Voorhorst, A. T. Hesselink, P. M. van Diemen, A. J. van den Brule, P. J. Snijders, and C. J. Meijer**. 2005. Concordance of specific human papillomavirus types in sex partners is more prevalent than would be expected by chance and is associated with increased viral loads. Clin. Infect. Dis. **41**:612-620.

19. **Bokal. E. V., B. J. Kocjan, M. Poljak, Z. Bogovac, and N. Jancar**. 2010. Genomic variants of human papillomavirus genotypes 16, 18, and 33 in women with cervical cancer in Slovenia. J. Obstet. Gynaecol. Res. **36**:1204-1213.

20. **Bosch, F. X. and N. Munoz**. 2002. The viral etiology of cervical cancer. Virus Res. **89** :183-190.

21. **Bosch, F. X., A. N. Burchell, M. Schiffman, A. R. Giuliano, S. S. de, L. Bruni, G. Tortolero-Luna, S. K. Kjaer, and N. Munoz**. 2008. Epidemiology and natural history of human papillomavirus infections and type-specific implications in cervical neoplasia. Vaccine **26 Suppl 10**:K1-16.

22. **Botha, H., B. Cooreman, G. Dreyer, G. Lindeque, A. Mouton, F. Guidozzi, T. Koller, T. Smith, A. Hoosen, L. Marcus, N. Dyson M. Moodley, and R. Soeters**. 2010. Cervical cancer and human papillomavirus: South African guidelines for screening and testing. South Afr J Gynaecol Oncol. **2(1):**23-26.

23. **Botha, M. H.** 2009. Cervical cancer. Continuing Medical Education. **27(3)**: 444-449.

24. **Brown, D. R., S. K. Kjaer, K. Sigurdsson, O. E. Iversen, M. Hernandez-Avila, C. M. Wheeler, G. Perez, L. A. Koutsky, E. H. Tay, P. Garcia, K. A. Ault, S. M. Garland, S. Leodolter, S. E. Olsson, G. W. Tang, D. G. Ferris, J. Paavonen, M. Steben, F. X. Bosch, J. Dillner, E. A. Joura, R. J. Kurman, S. Majewski, N. Munoz, E. R. Myers, L. L. Villa, F. J. Taddeo, C. Roberts, A. Tadesse, J. Bryan, L. C. Lupinacci, K. E. Giacoletti, H. L. Sings, M. James, T. M. Hesley, and E. Barr**. 2009. The impact of quadrivalent human papillomavirus (HPV; types 6, 11, 16, and 18) L1 virus-like particle vaccine on infection and disease due to oncogenic nonvaccine HPV types in generally HPV-naive women aged 16-26 years. J. Infect. Dis. **199**:926-935.

25. **Bruni, L., M. Diaz, X. Castellsague, E. Ferrer, F. X. Bosch, and S. de Sanjose**. 2010. Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. J. Infect. Dis. **202**:1789-1799.

26. **Bulk, S., J. Berkhof, L. Rozendaal, N. C. Fransen Daalmeijer, M. Gok, F. A. de Schipper, F. J. van Kemenade, P. J. Snijders, and C. J. Meijer**. 2007. The contribution of HPV18 to cervical cancer is underestimated using high-grade CIN as a measure of screening efficiency. Br. J. Cancer **96**:1234-1236.

27. **Bulkmans, N. W., J. Berkhof, S. Bulk, M. C. Bleeker, F. J. van Kemenade, L. Rozendaal, P. J. Snijders, and C. J. Meijer**. 2007. High-risk HPV type-specific clearance rates in cervical screening. Br. J. Cancer **96**:1419-1424.

28. **Buonaguro, F. M., M. L. Tornesello, I. Salatiello, P. Okong, L. Buonaguro, E. Beth-Giraldo, B. Biryahwaho, S. D. Sempala, and G. Giraldo**. 2000. The Uganda study on HPV variants and genital cancers. J. Clin. Virol. **19**:31-41.

29. **Burchell, A. N., F. Coutlee, P. P. Tellier, J. Hanley, and E. L. Franco**. 2011. Genital transmission of human papillomavirus in recently formed heterosexual couples. J. Infect. Dis. **204**:1723-1729.

30. **Burchell, A. N., P. P. Tellier, J. Hanley, F. Coutlee, and E. L. Franco**. 2010. Human papillomavirus infections among couples in new sexual relationships. Epidemiology **21**:31-37.

31. **Burchell, A. N., R. L. Winer, S. de Sanjose, and E. L. Franco**. 2006. Chapter 6: Epidemiology and transmission dynamics of genital HPV infection. Vaccine **24 Suppl 3**:S3-52-S3/61.

32. **Burk, R. D., M. Terai, P. E. Gravitt, L. A. Brinton, R. J. Kurman, W. A. Barnes, M. D. Greenberg, O. C. Hadjimichael, L. Fu, L. McGowan, R. Mortel, P. E. Schwartz, and A. Hildesheim**. 2003. Distribution of human papillomavirus types 16 and 18 variants in squamous cell carcinomas and adenocarcinomas of the cervix. Cancer Res. **63**:7215-7220.

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

34. **Calleja-Macias, I. E., L. L. Villa, J. C. Prado, M. Kalantari, B. Allan, A. L. Williamson, L. P. Chung, R. J. Collins, R. E. Zuna, S. T. Dunn, T. Y. Chu, H. A. Cubie, K. Cuschieri, M. von Knebel-Doeberitz, C. R. Martins, G. I. Sanchez, F. X. Bosch, N. Munoz, and H. U. Bernard**. 2005. Worldwide genomic diversity of the high-risk human papillomavirus types 31, 35, 52, and 58, four close relatives of human papillomavirus type 16. J. Virol. **79**:13630-13640.

35. **Canche, J. C., I. R. Lopez, N. G. Suarez, G. C. Acosta, L. Conde-Ferraez, T. C. Cetina, and M. R. Losa**. 2010. High prevalence and low E6 genetic variability of human papillomavirus 58 in women with cervical cancer and precursor lesions in Southeast Mexico. Mem. Inst. Oswaldo Cruz **105**:144-148.

36. **Castellsague, X.** 2008. Natural history and epidemiology of HPV infection and cervical cancer. Gynecol. Oncol. **110**:S4-S7.

37. **Castellsague, X. and N. Munoz**. 2003. Chapter 3: Cofactors in human papillomavirus carcinogenesis--role of parity, oral contraceptives, and tobacco smoking. J. Natl. Cancer Inst. Monogr20-28.

38. **Castellsague, X., A. Ghaffari, R. W. Daniel, F. X. Bosch, N. Munoz, and K. V. Shah**. 1997. Prevalence of penile human papillomavirus DNA in husbands of women with and without cervical neoplasia: a study in Spain and Colombia. J. Infect. Dis. **176**:353-361.

39. **Castellsague, X., C. Menendez, M. P. Loscertales, J. R. Kornegay, S. F. dos, F. X. Gomez-Olive, B. Lloveras, N. Abarca, N. Vaz, A. Barreto, F. X. Bosch, and P. Alonso**. 2001. Human papillomavirus genotypes in rural Mozambique. Lancet **358**:1429-1430.

40. **Castle, P. E., M. Schiffman, R. Herrero, A. Hildesheim, A. C. Rodriguez, M. C. Bratti, M. E. Sherman, S. Wacholder, R. Tarone, and R. D. Burk**. 2005. A prospective study of age trends in cervical human papillomavirus acquisition and persistence in Guanacaste, Costa Rica. J. Infect. Dis. **191**:1808-1816.

41. **Castresana, J.** 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. **17**:540-552.

42. **Centers for Disease Control (CDC)**. 2011. Recommendations on the use of quadrivalent human papillomavirus vaccine in males—Advisory Committee on Immunization Practices (ACIP). MMWR Morb Mortal Wkly Rep. **60:** 1705–08.

43. **Cento, V., N. Rahmatalla, M. Ciccozzi, C. F. Perno, and M. Ciotti**. 2011. Intratype variations of HPV 31 and 58 in Italian women with abnormal cervical cytology. J. Med. Virol. **83**:1752-1761.

44. **Cento, V., N. Rahmatalla, M. Ciccozzi, P. A. Lo, C. F. Perno, and M. Ciotti**. 2012. Human papillomaviruses 53 and 66: clinical aspects and genetic analysis. Virus Res. **163**:212-222.

45. **Chan, P. K.** 2012. Human papillomavirus type 58: the unique role in cervical cancers in East Asia. Cell Biosci. **2**:17.

46. **Chan, P. K., A. C. Luk, J. S. Park, K. K. Smith-McCune, J. M. Palefsky, R. Konno, L. Giovannelli, F. Coutlee, S. Hibbitts, T. Y. Chu, W. Settheetham-Ishida, M. A. Picconi, A. Ferrera, M. F. De, Y. L. Woo, T. Raiol, P. Pina-Sanchez, J. L. Cheung, J. H. Bae, M. Z. Chirenje, T. Magure, A. B. Moscicki, A. N. Fiander, S. R. Di, T. H. Cheung, M. M. Yu, S. K. Tsui, D. Pim, and L. Banks**. 2011. Identification of human papillomavirus type 58 lineages and the distribution worldwide. J. Infect. Dis. **203**:1565-1573.

47. **Chan, P. K., C. W. Lam, T. H. Cheung, W. W. Li, K. W. Lo, M. Y. Chan, J. L. Cheung, and A. F. Cheng**. 2002. Association of human papillomavirus type 58 variant with the risk of cervical cancer. J. Natl. Cancer Inst. **94**:1249-1253.

48. **Chan, P. K., C. Zhang, J. S. Park, K. K. Smith-McCune, J. M. Palefsky, L. Giovannelli, F. Coutlee, S. Hibbitts, R. Konno, W. Settheetham-Ishida, T. Y. Chu, A. Ferrera, P. M. Alejandra, M. F. De, Y. L. Woo, T. Raiol, P. Pina-Sanchez, J. H. Bae, M. C. Wong, M. Z. Chirenje, T. Magure, A. B. Moscicki, A. N. Fiander, G. Capra, K. E. Young, Y. Tan, Z. Chen, R. D. Burk, M. C. Chan, T. H. Cheung, D. Pim, and L. Banks**. 2012. Geographical distribution and oncogenic risk association of human papillomavirus type 58 E6 and E7 sequence variations. Int. J. Cancer. **132(11)**:2528-36.

49. **Chan, P. K., W. H. Li, M. Y. Chan, W. L. Ma, J. L. Cheung, and A. F. Cheng**. 1999. High prevalence of human papillomavirus type 58 in Chinese women with cervical cancer and precancerous lesions. J. Med. Virol. **59**:232-238.

50. **Chan, W. K., G. Klock, and H. U. Bernard**. 1989. Progesterone and glucocorticoid response elements occur in the long control regions of several human papillomaviruses involved in anogenital neoplasia. J. Virol. **63**:3261-3269.

51. **Chang, Y. J., H. C. Chen, B. H. Lee, S. L. You, C. Y. Lin, M. H. Pan, Y. C. Chou, C. Y. Hsieh, Y. M. Chen, Y. J. Cheng, and C. J. Chen**. 2011. Unique variants of human papillomavirus genotypes 52 and 58 and risk of cervical neoplasia. Int. J. Cancer **129**:965-973.

52. **Chansaenroj, J., A. Theamboonlers, P. Junyangdikul, S. Swangvaree, A. Karalak, and Y. Poovorawan**. 2012. Whole genome analysis of human papillomavirus type 16 multiple infection in cervical cancer patients. Asian Pac. J. Cancer Prev. **13**:599-606.

53. **Chaturvedi, A. K.** 2010. Beyond cervical cancer: burden of other HPV-related cancers among men and women. J. Adolesc. Health **46**:S20-S26.

54. **Chen, H. C., M. Schiffman, C. Y. Lin, M. H. Pan, S. L. You, L. C. Chuang, C. Y. Hsieh, K. L. Liaw, A. W. Hsing, and C. J. Chen**. 2011b. Persistence of type-specific human papillomavirus infection and increased long-term risk of cervical cancer. J. Natl. Cancer Inst. **103**:1387-1396.

55. **Chen, S. L., Y. K. Lin, L. Y. Li, Y. P. Tsao, H. Y. Lo, W. B. Wang, and T. C. Tsai**. 1996b. E5 proteins of human papillomavirus types 11 and 16 transactivate the c-fos promoter through the NF1 binding element. J. Virol. **70**:8558-8563.

56. **Chen, X. S., R. L. Garcea, I. Goldberg, G. Casini, and S. C. Harrison**. 2000. Structure of small virus-like particles assembled from the L1 protein of human papillomavirus 16. Mol. Cell **5**:557-567.

57. **Chen, Y. H., L. H. Huang, and T. M. Chen**. 1996a. Differential effects of progestins and estrogens on long control regions of human papillomavirus types 16 and 18. Biochem. Biophys. Res. Commun. **224**:651-659.

58. **Chen, Z., M. Schiffman, R. Herrero, R. Desalle, K. Anastos, M. Segondy, V. V. Sahasrabuddhe, P. E. Gravitt, A. W. Hsing, and R. D. Burk**. 2011a. Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. PLoS. One. **6**:e20183.

59. **Chen, Z., M. Terai, L. Fu, R. Herrero, R. DeSalle, and R. D. Burk**. 2005. Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. J. Virol. **79**:7014-7023.

60. **Cheng, S., D. C. Schmidt-Grimminger, T. Murant, T. R. Broker, and L. T. Chow**. 1995. Differentiation-dependent up-regulation of the human papillomavirus E7 gene reactivates cellular DNA replication in suprabasal differentiated keratinocytes. Genes Dev. **9**:2335-2349.

61. **Chouhy, D., M. Gorosito, A. Sanchez, E. C. Serra, A. Bergero, B. R. Fernandez, and A. A. Giri**. 2010. New generic primer system targeting mucosal/genital and cutaneous human papillomaviruses leads to the characterization of HPV 115, a novel Beta-papillomavirus species 3. Virology **397**:205-216.

62. **Chow, L. T., T. R. Broker, and B. M. Steinberg**. 2010. The natural history of human papillomavirus infections of the mucosal epithelia. APMIS **118**:422-449.

63. **Clifford, G. M., R. K. Rana, S. Franceschi, J. S. Smith, G. Gough, and J. M. Pimenta**. 2005a. Human papillomavirus genotype distribution in low-grade cervical lesions: comparison by geographic region and with cervical cancer. Cancer Epidemiol. Biomarkers Prev. **14**:1157-1164.

64. **Clifford, G. M., S. Gallus, R. Herrero, N. Munoz, P. J. Snijders, S. Vaccarella, P. T. Anh, C. Ferreccio, N. T. Hieu, E. Matos, M. Molano, R. Rajkumar, G. Ronco, S. S. de, H. R. Shin, S. Sukvirach, J. O. Thomas, S. Tunsakul, C. J. Meijer, and S. Franceschi**. 2005b. Worldwide distribution of human papillomavirus types in cytologically normal women in the International Agency for Research on Cancer HPV prevalence surveys: a pooled analysis. Lancet **366**:991-998.

65. **Clifford, G., S. Franceschi, M. Diaz, N. Munoz, and L. L. Villa**. 2006. Chapter 3: HPV type-distribution in women with and without cervical neoplastic diseases. Vaccine **24 Suppl 3**:S3-26-S3/34.

66. **Collins, S., T. P. Rollason, L. S. Young, and C. B. Woodman**. 2010. Cigarette smoking is an independent risk factor for cervical intraepithelial neoplasia in young women: a longitudinal study. Eur. J. Cancer **46**:405-411.

67. **Cornet, I., T. Gheit, S. Franceschi, J. Vignat, R. D. Burk, B. S. Sylla, M. Tommasino, and G. M. Clifford**. 2012. Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. J. Virol. **86**:6855-6861.

68. **Cripe, T. P., A. Alderborn, R. D. Anderson, S. Parkkinen, P. Bergman, T. H. Haugen, U. Pettersson, and L. P. Turek**. 1990. Transcriptional activation of the human papillomavirus-16 P97 promoter by an 88-nucleotide enhancer containing distinct cell-dependent and AP-1-responsive modules. New Biol. **2**:450-463.

69. **Culp, T. D., L. R. Budgeon, M. P. Marinkovich, G. Meneguzzi, and N. D. Christensen**. 2006. Keratinocyte-secreted laminin 5 can function as a transient receptor for human papillomaviruses by binding virions and transferring them to adjacent cells. J. Virol. **80**:8940-8950.

70. **Cuschieri, K. S., H. A. Cubie, M. W. Whitley, A. L. Seagar, M. J. Arends, C. Moore, G. Gilkisson, and E. McGoogan**. 2004. Multiple high risk HPV infections are common in cervical neoplasia and young women in a cervical screening population. J. Clin. Pathol. **57**:68-72.

71. **Dalstein, V., D. Riethmuller, J. L. Pretet, C. K. 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.

72. **Da Ros, C. T. and C. S. Schmitt**. 2008. Global epidemiology of sexually transmitted diseases. Asian J. Androl **10**:110-114.

73. **Dartell, M., V. Rasch, C. Kahesa, J. Mwaiselage, T. Ngoma, J. Junge, A. Gernow, S. F. Ejlersen, C. Munk, T. Iftner, and S. K. Kjaer**. 2012. Human papillomavirus prevalence and type distribution in 3603 HIV-positive and HIV-negative women in the general population of Tanzania: the PROTECT study. Sex Transm. Dis. **39**:201-208.

74. **Datta, P., N. Bhatla, R. M. Pandey, L. Dar, A. R. Patro, S. Vasisht, A. Kriplani, and N. Singh**. 2012. Type-specific incidence and persistence of HPV infection among young women: a prospective study in North India. Asian Pac. J. Cancer Prev. **13**:1019-1024.

75. **De la Rosa. G., A. Monroy-Garcia, M. L. Mora-Garcia, C. G. Pena, J. Hernandez-Montes, B. Weiss-Steider, and M. A. Gomez-Lim**. 2009. An HPV 16 L1-based chimeric human papilloma virus-like particles containing a string of epitopes produced in plants is able to elicit humoral and cytotoxic T-cell activity in mice. Virol. J. **6** :2.

76. **de Sanjose. S., M. Diaz, X. Castellsague, G. Clifford, L. Bruni, N. Munoz, and F. X. Bosch**. 2007. Worldwide prevalence and genotype distribution of cervical human papillomavirus DNA in women with normal cytology: a meta-analysis. Lancet Infect. Dis. **7**:453-459.

77. **de Sanjose. S., W. G. Quint, L. Alemany, D. T. Geraets, J. E. Klaustermeier, B. Lloveras, S. Tous, A. Felix, L. E. Bravo, H. R. Shin, C. S. Vallejos, P. A. de Ruiz, M. A. Lima, N. Guimera, O. Clavero, M. Alejo, A. Llombart-Bosch, C. Cheng-Yang, S. A. Tatti, E. Kasamatsu, E. Iljazovic, M. Odida, R. Prado, M. Seoud, M. Grce, A. Usubutun, A. Jain, G. A. Suarez, L. E. Lombardi, A. Banjo, C. Menendez, E. J. Domingo, J. Velasco, A. Nessa, S. C. Chichareon, Y. L. Qiao, E. Lerma, S. M. Garland, T. Sasagawa, A. Ferrera, D. Hammouda, L. Mariani, A. Pelayo, I. Steiner, E. Oliva, C. J. Meijer, W. F. Al-Jassar, E. Cruz, T. C. Wright, A. Puras, C. L. Llave, M. Tzardi, T. Agorastos, V. Garcia-Barriola, C. Clavel, J. Ordi, M. Andujar, X. Castellsague, G. I. Sanchez, A. M. Nowakowski, J. Bornstein, N. Munoz, and F. X. Bosch**. 2010. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. Lancet Oncol. **11**:1048-1056.

78. **de Villiers, E. M., C. Fauquet, T. R. Broker, H. U. Bernard, and H. zur Hausen.** . 2004. Classification of papillomaviruses. Virology **324**:17-27.

79. **De Vuyst. H., M. R. Parisi, A. Karani, K. Mandaliya, L. Muchiri, S. Vaccarella, M. Temmerman, S. Franceschi, and F. Lillo**. 2010. The prevalence of human papillomavirus infection in Mombasa, Kenya. Cancer Causes Control **21**:2309-2313.

80. **Deluca, G. D., J. Basiletti, E. Schelover, N. D. Vasquez, J. M. Alonso, H. M. Marin, R. H. Lucero, and M. A. Picconi**. 2011. Chlamydia trachomatis as a probable cofactor in human papillomavirus infection in aboriginal women from northeastern Argentina. Braz. J. Infect. Dis. **15**:567-572.

81. **Demidov, V. V.** 2002. Rolling-circle amplification in DNA diagnostics: the power of simplicity. Expert. Rev. Mol. Diagn. **2**:542-548.

82. **Denny, L.** 2010. Cervical cancer in South Africa: An overview of current status and prevention strategies. Continuing Medical Education. **28(2)**: 70-73.

83. **Denny, L. A., S. Franceschi, S. S. de, I. Heard, A. B. Moscicki, and J. Palefsky**. 2012. Human papillomavirus, human immunodeficiency virus and immunosuppression. Vaccine **30 Suppl 5**:F168-F174.

84. **Deschuyteneer, M., A. Elouahabi, D. Plainchamp, M. Plisnier, D. Soete, Y. Corazza, L. Lockman, S. Giannini, and M. Deschamps**. 2010. Molecular and structural characterization of the L1 virus-like particles that are used as vaccine antigens in Cervarix, the AS04-adjuvanted HPV-16 and -18 cervical cancer vaccine. Hum. Vaccin. **6**:407-419.

85. **Dobec, M., F. Bannwart, F. Kaeppeli, and P. Cassinotti**. 2009. Automation of the linear array HPV genotyping test and its application for routine typing of human papillomaviruses in cervical specimens of women without cytological abnormalities in Switzerland. J. Clin. Virol. **45**:23-27.

86. **Dols, J. A., G. Reid, J. M. Brown, H. Tempelman, T. R. Bontekoe, W. G. Quint, and M. E. Boon**. 2012. HPV Type Distribution and Cervical Cytology among HIV-Positive Tanzanian and South African Women. ISRN. Obstet. Gynecol. **2012**:514146.

87. **Donne, A. J., L. Hampson, J. J. Homer, and I. N. Hampson**. 2010. The role of HPV type in Recurrent Respiratory Papillomatosis. Int. J. Pediatr. Otorhinolaryngol. **74**:7-14.

88. **Donovan, B., N. Franklin, R. Guy, A. E. Grulich, D. G. Regan, H. Ali, H. Wand, and C. K. Fairley**. 2011. Quadrivalent human papillomavirus vaccination and trends in genital warts in Australia: analysis of national sentinel surveillance data. Lancet Infect. Dis. **11**:39-44.

89. **Doorbar, J.** 2005. The papillomavirus life cycle. J. Clin. Virol. **32 Suppl 1**:S7-15.

90. **Doorbar, J.** 2006. Molecular biology of human papillomavirus infection and cervical cancer. Clin. Sci. (Lond) **110**:525-541.

91. **Doorbar, J., D. Campbell, R. J. Grand, and P. H. Gallimore**. 1986. Identification of the human papilloma virus-1a E4 gene products. EMBO J. **5**:355-362.

92. **Doorbar, J., W. Quint, L. Banks, I. G. Bravo, M. Stoler, T. R. Broker, and M. A. Stanley**. 2012. The biology and life-cycle of human papillomaviruses. Vaccine **30 Suppl 5**:F55-F70.

93. **Draper, E., S. L. Bissett, R. Howell-Jones, D. Edwards, G. Munslow, K. Soldan, and S. Beddows**. 2011. Neutralization of non-vaccine human papillomavirus pseudoviruses from the A7 and A9 species groups by bivalent HPV vaccine sera. Vaccine **29**:8585-8590.

94. **Dunne, E. F., E. R. Unger, M. Sternberg, G. McQuillan, D. C. Swan, S. S. Patel, and L. E. Markowitz**. 2007. Prevalence of HPV infection among females in the United States. JAMA **297**:813-819.

95. **Edgar, R. C.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**:1792-1797.

96. **Einstein, M. H., M. Baron, M. J. Levin, A. Chatterjee, R. P. Edwards, F. Zepp, I. Carletti, F. J. Dessy, A. F. Trofa, A. Schuind, and G. Dubin**. 2009. Comparison of the immunogenicity and safety of Cervarix and Gardasil human papillomavirus (HPV) cervical cancer vaccines in healthy women aged 18-45 years. Hum. Vaccin. **5**:705-719.

97. **Evander, M., I. H. Frazer, E. Payne, Y. M. Qi, K. Hengst, and N. A. McMillan**. 1997. Identification of the alpha6 integrin as a candidate receptor for papillomaviruses. J. Virol. **71**:2449-2456.

98. **Florin, L., C. Sapp, R. E. Streeck, and M. Sapp**. 2002. Assembly and translocation of papillomavirus capsid proteins. J. Virol. **76**:10009-10014.

99. **Fonn, S., B. Bloch, M. Mabina, S. Carpenter, H. Cronje, C. Maise, M. Bennun, T. G. du, J. E. de, I. Manana, and G. Lindeque**. 2002. Prevalence of pre-cancerous lesions and cervical cancer in South Africa--a multicentre study. S. Afr. Med. J. **92**:148-156.

100. **Forman, D., M. C. de, C. J. Lacey, I. Soerjomataram, J. Lortet-Tieulent, L. Bruni, J. Vignat, J. Ferlay, F. Bray, M. Plummer, and S. Franceschi**. 2012. Global burden of human papillomavirus and related diseases. Vaccine **30 Suppl 5**:F12-F23.

101. **Franco, E. L., L. L. Villa, J. P. Sobrinho, J. M. Prado, M. C. Rousseau, M. Desy, and T. E. Rohan**. 1999. Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer. J. Infect. Dis. **180**:1415-1423.

102. **Frazer, I. H., G. R. Leggatt, and S. R. Mattarollo**. 2011. Prevention and treatment of papillomavirus-related cancers through immunization. Annu. Rev. Immunol. **29**:111-138.

103. **Fukuchi, E., G. F. Sawaya, M. Chirenje, T. Magure, J. Tuveson, Y. Ma, S. Shiboski, C. M. Da, J. Palefsky, A. B. Moscicki, R. Makunike-Mutasa, T. Chipato, and K. K. Smith-McCune**. 2009. Cervical human papillomavirus incidence and persistence in a cohort of HIV-negative women in Zimbabwe. Sex Transm. Dis. **36**:305-311.

104. **Gage, J. C., K. O. Ajenifuja, N. A. Wentzensen, A. C. Adepiti, C. Eklund, M. Reilly, M. Hutchinson, S. Wacholder, J. Harford, A. S. Soliman, R. D. Burk, and M. Schiffman**. 2012. The age-specific prevalence of human papillomavirus and risk of cytologic abnormalities in rural Nigeria: implications for screen-and-treat strategies. Int. J. Cancer **130**:2111-2117.

105. **Gargiulo, F., M. A. De Francesco, C. Schreiber, G. Ciravolo, F. Salinaro, B. Valloncini, and N. Manca**. 2007. Prevalence and distribution of single and multiple HPV infections in cytologically abnormal cervical samples from Italian women. Virus Res. **125**:176-182.

106. **Gariglio, P., J. Gutierrez, E. Cortes, and J. Vazquez**. 2009. The role of retinoid deficiency and estrogens as cofactors in cervical cancer. Arch. Med. Res. **40**:449-465.

107. **Georgousakis, M., S. Jayasinghe, J. Brotherton, N. Gilroy, C. Chiu, and K. Macartney**. 2012. Population-wide vaccination against human papillomavirus in adolescent boys: Australia as a case study. Lancet Infect. Dis. **12**:627-634.

108. **Giovannelli, L., C. Bellavia, G. Capra, M. C. Migliore, M. Caleca, M. Giglio, A. Perino, D. Matranga, and P. Ammatuna**. 2007. HPV group- and type-specific concordance in HPV infected sexual couples. J. Med. Virol. **79**:1882-1888.

109. **Giuliano, A. R., J. M. Palefsky, S. Goldstone, E. D. Moreira, Jr., M. E. Penny, C. Aranda, E. Vardas, H. Moi, H. Jessen, R. Hillman, Y. H. Chang, D. Ferris, D. Rouleau, J. Bryan, J. B. Marshall, S. Vuocolo, E. Barr, D. Radley, R. M. Haupt, and D. Guris**. 2011. Efficacy of quadrivalent HPV vaccine against HPV Infection and disease in males. N. Engl. J. Med. **364**:401-411.

110. **GLOBOCAN** 2008. **Ferlay, J., H. R. Shin, F. Bray, D. Forman, C. Mathers, and D. M. Parkin**. 2010. Estimates of worldwide burden of cancer in 2008. Int. J. Cancer. **127(12)**:2893-917.

111. **Godinez, J. M., D. A. Heideman, T. Gheit, L. Alemany, P. J. Snijders, M. Tommasino, C. J. Meijer, S. S. de, F. X. Bosch, and I. G. Bravo**. 2013. Differential presence of Papillomavirus variants in cervical cancer: an analysis for HPV33, HPV45 and HPV58. Infect. Genet. Evol. **13**:96-104.

112. **Goodman, M. T., Y. B. Shvetsov, K. McDuffie, L. R. Wilkens, X. Zhu, P. J. Thompson, L. Ning, J. Killeen, L. Kamemoto, and B. Y. Hernandez**. 2008. Prevalence, acquisition, and clearance of cervical human papillomavirus infection among women with normal cytology: Hawaii Human Papillomavirus Cohort Study. Cancer Res. **68**:8813-8824.

113. **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.

114. **Guan, P., R. Howell-Jones, N. Li, L. Bruni, S. S. de, S. Franceschi, and G. M. Clifford**. 2012. Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer. Int. J. Cancer **131**:2349-2359.

115. **Guindon S., J. F. Dufayard , V. Lefort , M. Anisimova , W. Hordijk , O. Gascuel** . 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Systematic Biology, **59** (3):307-21.

116. **Healey, S. M., K. J. Aronson, Y. Mao, N. F. Schlecht, L. S. Mery, A. Ferenczy, and E. L. Franco**. 2001. Oncogenic human papillomavirus infection and cervical lesions in aboriginal women of Nunavut, Canada. Sex Transm. Dis. **28**:694-700.

117. **Hernandez, B. Y., L. R. Wilkens, X. Zhu, P. Thompson, K. McDuffie, Y. B. Shvetsov, L. E. Kamemoto, J. Killeen, L. Ning, and M. T. Goodman**. 2008. Transmission of human papillomavirus in heterosexual couples. Emerg. Infect. Dis. **14**:888-894.

118. **Hesselink, A. T., J. Berkhof, D. A. Heideman, N. W. Bulkmans, J. E. van Tellingen, C. J. Meijer, and P. J. Snijders**. 2009. High-risk human papillomavirus DNA load in a population-based cervical screening cohort in relation to the detection of high-grade cervical intraepithelial neoplasia and cervical cancer. Int. J. Cancer **124**:381-386.

119. **Hildesheim, A., M. Schiffman, C. Bromley, S. Wacholder, R. Herrero, A. Rodriguez, M. C. Bratti, M. E. Sherman, U. Scarpidis, Q. Q. Lin, M. Terai, R. L. Bromley, K. Buetow, R. J. Apple, and R. D. Burk**. 2001. Human papillomavirus type 16 variants and risk of cervical cancer. J. Natl. Cancer Inst. **93**:315-318.

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

121. **Ho, L., S. K. Tay, S. Y. Chan, and H. U. Bernard**. 1993a. Sequence variants of human papillomavirus type 16 from couples suggest sexual transmission with low infectivity and polyclonality in genital neoplasia. J. Infect. Dis. **168**:803-809.

122. **Ho, L., S. Y. Chan, R. D. Burk, B. C. Das, K. Fujinaga, J. P. Icenogle, T. Kahn, N. Kiviat, W. Lancaster, P. Mavromara-Nazos, . Labropoulou, S. Mitrani-Rosenbaum, B. Norrild, M. R. Pillai, J. Stoerker, K. Syrjaenen, S. Syrjaenen,S. K. Tay, L. L. Villa, C. M. Wheeler, A. L. Williamson, and H. U. Bernard.** 1993b. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. J. Virol. **67**:6413-6423.

123. **Hoppe-Seyler, F., K. Butz, and H. zur Hausen**. 1991. Repression of the human papillomavirus type 18 enhancer by the cellular transcription factor Oct-1. J. Virol. **65**:5613-5618.

124. **HPV Information Centre** (2009). Human papillomavirus and related cancers 2009: WHO/ICO Information Centre on HPV and Cervical Cancer (http://www.who.int/hpvcentre).

125. http://wenliang.myweb.uga.edu/mystudy/virology/ScienceOfVirology/DNAvirus.genomereplicationandtranscripti.html.

126. **Huang, L. W., S. L. Chao, and B. H. Lee**. 2008. Integration of human papillomavirus type-16 and type-18 is a very early event in cervical carcinogenesis. J. Clin. Pathol. **61**:627-631.

127. **Iftner, T. and L. L. Villa**. 2003. Chapter 12: Human papillomavirus technologies. J. Natl. Cancer Inst. Monogr80-88.

128. **Iftner, T., S. Eberle, A. Iftner, B. Holz, N. Banik, W. Quint, and A. N. Straube**. 2010. Prevalence of low-risk and high-risk types of human papillomavirus and other risk factors for HPV infection in Germany within different age groups in women up to 30 years of age: an epidemiological observational study. J. Med. Virol. **82**:1928-1939.

129. **Jamieson, D. J., A. Duerr, R. Burk, R. S. Klein, P. Paramsothy, P. Schuman, S. Cu-Uvin, and K. Shah**. 2002. Characterization of genital human papillomavirus infection in women who have or who are at risk of having HIV infection. Am. J. Obstet. Gynecol. **186**:21-27.

130. **Jiang, M., L. F. Xi, Z. R. Edelstein, D. A. Galloway, G. J. Olsem, W. C. Lin, and N. B. Kiviat**. 2009. Identification of recombinant human papillomavirus type 16 variants. Virology **394**:8-11.

131. **Jit, M., R. Chapman, O. Hughes, and Y. H. Choi**. 2011. Comparing bivalent and quadrivalent human papillomavirus vaccines: economic evaluation based on transmission model. BMJ **343**:d5775.

132. **Johne, R., H. Muller, A. Rector, R. M. van, and H. Stevens**. 2009. Rolling-circle amplification of viral DNA genomes using phi29 polymerase. Trends Microbiol. **17**:205-211.

133. **Junes-Gill, K., L. Sichero, P. C. Maciag, W. Mello, V. Noronha, and L. L. Villa**. 2008. Human papillomavirus type 16 variants in cervical cancer from an admixtured population in Brazil. J. Med. Virol. **80**:1639-1645.

134. **Jung, W. W., T. Chun, D. Sul, K. W. Hwang, H. S. Kang, D. J. Lee, and I. K. Han**. 2004. Strategies against human papillomavirus infection and cervical cancer. J. Microbiol. **42**:255-266.

135. **Kadaja, M., H. Isok-Paas, T. Laos, E. Ustav, and M. Ustav**. 2009. Mechanism of genomic instability in cells infected with the high-risk human papillomaviruses. PLoS. Pathog. **5**:e1000397.

136. **Kammer, C., M. Tommasino, S. Syrjanen, H. Delius, U. Hebling, U. Warthorst, H. Pfister, and I. Zehbe**. 2002. Variants of the long control region and the E6 oncogene in European human papillomavirus type 16 isolates: implications for cervical disease. Br. J. Cancer **86**:269-273.

137. **Kammer, C., U. Warthorst, N. Torrez-Martinez, C. M. Wheeler, and H. Pfister**. 2000. Sequence analysis of the long control region of human papillomavirus type 16 variants and functional consequences for P97 promoter activity. J. Gen. Virol. **81**:1975-1981.

138. **Khan, M. J., P. E. Castle, A. T. Lorincz, S. Wacholder, M. Sherman, D. R. Scott, B. B. Rush, A. G. Glass, and M. Schiffman**. 2005. The elevated 10-year risk of cervical precancer and cancer in women with human papillomavirus (HPV) type 16 or 18 and the possible utility of type-specific HPV testing in clinical practice. J. Natl. Cancer Inst. **97**:1072-1079.

139. **Kirnbauer, R., F. Booy, N. Cheng, D. R. Lowy, and J. T. Schiller**. 1992. Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. Proc. Natl. Acad. Sci. U. S. A **89**:12180-12184.

140. **Kitchener, H. C., P. E. Castle, and J. T. Cox**. 2006. Chapter 7: Achievements and limitations of cervical cytology screening. Vaccine **24 Suppl 3**:S3-63-S3/70.

141. **Kjaer, S. K., B. Chackerian, A. J. van den Brule, E. I. Svare, G. Paull, J. M. Walbomers, J. T. Schiller, J. E. Bock, M. E. Sherman, D. R. Lowy, and C. L. Meijer**. 2001. High-risk human papillomavirus is sexually transmitted: evidence from a follow-up study of virgins starting sexual activity (intercourse). Cancer Epidemiol. Biomarkers Prev. **10**:101-106.

142. **Kjaer, S. K., G. Breugelmans, C. Munk, J. Junge, M. Watson, and T. Iftner**. 2008. Population-based prevalence, type- and age-specific distribution of HPV in women before introduction of an HPV-vaccination program in Denmark. Int. J. Cancer **123**:1864-1870.

143. **Kocjan, B. J., K. Seme, T. Mocilnik, N. Jancar, E. Vrtacnik-Bokal, and M. Poljak**. 2007. Genomic diversity of human papillomavirus genotype 53 in an ethnogeographically closed cohort of white European women. J. Med. Virol. **79**:431-438.

144. **Konno, R., S. Tamura, K. Dobbelaere, and H. Yoshikawa**. 2011. Prevalence and type distribution of human papillomavirus in healthy Japanese women aged 20 to 25 years old enrolled in a clinical study. Cancer Sci. **102**:877-882.

145. **Kovacic, M. B., P. E. Castle, R. Herrero, M. Schiffman, M. E. Sherman, S. Wacholder, A. C. Rodriguez, M. L. Hutchinson, M. C. Bratti, A. Hildesheim, J. Morales, M. Alfaro, and R. D. Burk**. 2006. Relationships of human papillomavirus type, qualitative viral load, and age with cytologic abnormality. Cancer Res. **66**:10112-10119.

146. **Kozuka, T., Y. Aoki, K. Nakagawa, K. Ohtomo, H. Yoshikawa, K. Matsumoto, K. Yoshiike, and T. Kanda**. 2000. Enhancer-promoter activity of human papillomavirus type 16 long control regions isolated from cell lines SiHa and CaSki and cervical cancer biopsies. Jpn. J. Cancer Res. **91**:271-279.

147. **Kurvinen, K., M. Yliskoski, S. Saarikoski, K. Syrjanen, and S. Syrjanen**. 2000. Variants of the long control region of human papillomavirus type 16. Eur. J. Cancer **36** :1402-1410.

148. **Lace, M. J., C. Isacson, J. R. Anson, A. T. Lorincz, S. P. Wilczynski, T. H. Haugen, and L. P. Turek**. 2009. Upstream regulatory region alterations found in human papillomavirus type 16 (HPV-16) isolates from cervical carcinomas increase transcription, ori function, and HPV immortalization capacity in culture. J. Virol. **83**:7457-7466.

149. **Lee, S. H., V. S. Vigliotti, J. S. Vigliotti, and S. Pappu**. 2009. Validation of human papillomavirus genotyping by signature DNA sequence analysis. BMC. Clin. Pathol. **9**:3.

150. **Leggatt, G. R. and I. H. Frazer** . 2007. HPV vaccines: the beginning of the end for cervical cancer. Curr. Opin. Immunol. **19**:232-238.

151. **Lei, Y. J., K. Makhaola, D. Pittayakhajonwut, C. Wood, and P. C. Angeletti**. 2011. Human papillomavirus 16 variants from Zambian women with normal pap smears. J. Med. Virol. **83**:1230-1237.

152. **Lenselink, C. H., W. J. Melchers, W. G. Quint, A. M. Hoebers, J. C. Hendriks, L. F. Massuger, and R. L. Bekkers** . 2008. Sexual behaviour and HPV infections in 18 to 29 year old women in the pre-vaccine era in the Netherlands. PLoS. One. **3**:e3743.

153. **Liao, S. F., W. C. Lee, H. C. Chen, L. C. Chuang, M. H. Pan, and C. J. Chen**. 2012. Baseline human papillomavirus infection, high vaginal parity, and their interaction on cervical cancer risks after a follow-up of more than 10 years. Cancer Causes Control **23**:703-708.

154. **Liaw, K. L., A. Hildesheim, R. D. Burk, P. Gravitt, S. Wacholder, M. M. Manos, D. R. Scott, M. E. Sherman, R. J. Kurman, A. G. Glass, S. M. Anderson, and M. Schiffman**. 2001. A prospective study of human papillomavirus (HPV) type 16 DNA detection by polymerase chain reaction and its association with acquisition and persistence of other HPV types. J. Infect. Dis. **183**:8-15.

155. **Liu, J. H., G. L. Wang, W. Q. Zhou, C. Liu, L. X. Yang, Q. Ruan, and Z. R. Sun**. 2012. Variations of human papillomavirus type 58 E6, E7, L1 genes and long control region in strains from women with cervical lesions in Liaoning province, China. Infect. Genet. Evol. **12(7)**:1466-72.

156. **Lizano, M., J. Berumen, and A. Garcia-Carranca**. 2009. HPV-related carcinogenesis: basic concepts, viral types and variants. Arch. Med. Res. **40**:428-434.

157. **Lizano, M., l. C.-H. De, A. Carrillo-Garcia, A. Garcia-Carranca, d. L.-R. Ponce, A. Duenas-Gonzalez, D. M. Hernandez-Hernandez, and A. Mohar**. 2006. Distribution of HPV16 and 18 intratypic variants in normal cytology, intraepithelial lesions, and cervical cancer in a Mexican population. Gynecol. Oncol. **102**:230-235.

158. **Lopez-Saavedra, A., L. Gonzalez-Maya, S. Ponce-de-Leon, A. Garcia-Carranca, A. Mohar, and M. Lizano**. 2009. Functional implication of sequence variation in the long control region and E2 gene among human papillomavirus type 18 variants. Arch. Virol. **154**:747-754.

159. **Louie, K. S., S. S. de, and P. Mayaud**. 2009b. Epidemiology and prevention of human papillomavirus and cervical cancer in sub-Saharan Africa: a comprehensive review. Trop. Med. Int. Health **14**:1287-1302.

160. **Louie, K. S., S. S. de, M. Diaz, X. Castellsague, R. Herrero, C. J. Meijer, K. Shah, S. Franceschi, N. Munoz, and F. X. Bosch**. 2009a. Early age at first sexual intercourse and early pregnancy are risk factors for cervical cancer in developing countries. Br. J. Cancer **100**:1191-1197.

161. **Lowe, J., D. Panda, S. Rose, T. Jensen, W. A. Hughes, F. Y. Tso, and P. C. Angeletti**. 2008. Evolutionary and structural analyses of alpha-papillomavirus capsid proteins yields novel insights into L2 structure and interaction with L1. Virol. J. **5** :150.

162. **Lugo-Trampe, A., K. D. Trujillo-Murillo, I. P. Rodriguez-Sanchez, M. D. Barbosa-Cerda, J. D. Lugo-Trampe, L. C. Hernandez-Ramirez, L. M. Canseco-Avila, M. Espinoza-Ruiz, S. Dominguez-Arrevillaga, and I. gado-Enciso**. 2012. A PCR-RFLP method for typing human papillomavirus type 16 variants. J. Virol. Methods**. 187(2)**:338-44.

163. **Madeleine, M. M., T. Anttila, S. M. Schwartz, P. Saikku, M. Leinonen, J. J. Carter, M. Wurscher, L. G. Johnson, D. A. Galloway, and J. R. Daling**. 2007. Risk of cervical cancer associated with Chlamydia trachomatis antibodies by histology, HPV type and HPV cofactors. Int. J. Cancer **120**:650-655.

164. **Marais, D. J., D. Constant, B. Allan, H. Carrara, M. Hoffman, S. Shapiro, C. Morroni, and A. L. Williamson**. 2008. Cervical human papillomavirus (HPV) infection and HPV type 16 antibodies in South African women. J. Clin. Microbiol. **46**:732-739.

165. **Marincevic-Zuniga, Y., I. Gustavsson, and U. Gyllensten**. 2012. Multiply-primed rolling circle amplification of human papillomavirus using sequence-specific primers. Virology. **46(2)**:732-9.

166. **Marks, M., P. E. Gravitt, S. B. Gupta, K. L. Liaw, A. Tadesse, E. Kim, C. Phongnarisorn, V. Wootipoom, P. Yuenyao, C. Vipupinyo, S. Sriplienchan, and D. D. Celentano**. 2011. Combined oral contraceptive use increases HPV persistence but not new HPV detection in a cohort of women from Thailand. J. Infect. Dis. **204**:1505-1513.

167. **Matos, A., J. Moutinho, D. Pinto, and R. Medeiros**. 2005. The influence of smoking and other cofactors on the time to onset to cervical cancer in a southern European population. Eur. J. Cancer Prev. **14**:485-491.

168. **Matsumoto, K., A. Oki, R. Furuta, H. Maeda, T. Yasugi, N. Takatsuka, A. Mitsuhashi, T. Fujii, Y. Hirai, T. Iwasaka, N. Yaegashi, Y. Watanabe, Y. Nagai, T. Kitagawa, and H. Yoshikawa**. 2011. Predicting the progression of cervical precursor lesions by human papillomavirus genotyping: a prospective cohort study. Int. J. Cancer **128**:2898-2910.

169. **Maucort-Boulch, D., M. Plummer, P. E. Castle, F. Demuth, M. Safaeian, C. M. Wheeler, and M. Schiffman**. 2010. Predictors of human papillomavirus persistence among women with equivocal or mildly abnormal cytology. Int. J. Cancer **126**:684-691.

170. **Maufort, J. P., A. Shai, H. C. Pitot, and P. F. Lambert**. 2010. A role for HPV16 E5 in cervical carcinogenesis. Cancer Res**. 70**:2924-2931.

171. **Mayrand, M. H., F. Coutlee, C. Hankins, N. Lapointe, P. Forest, L. M. de, and M. Roger**. 2000. Detection of human papillomavirus type 16 DNA in consecutive genital samples does not always represent persistent infection as determined by molecular variant analysis. J. Clin. Microbiol. **38**:3388-3393.

172. **Mazumder, I. D., R. K. Singh, S. Mitra, S. Dutta, C. Chakraborty, P. S. Basu, R. K. Mondal, S. Roychoudhury, and C. K. Panda**. 2011. Genetic and epigenetic changes of HPV16 in cervical cancer differentially regulate E6/E7 expression and associate with disease progression. Gynecol. Oncol. **123**:597-604.

173. **Mbulawa, Z. Z., D. Coetzee, D. J. Marais, M. Kamupira, E. Zwane, B. Allan, D. Constant, J. R. Moodley, M. Hoffman, and A. L. Williamson**. 2009. Genital human papillomavirus prevalence and human papillomavirus concordance in heterosexual couples are positively associated with human immunodeficiency virus coinfection. J. Infect. Dis. **199**:1514-1524.

174. **Mbulawa, Z. Z., D. J. Marais, L. F. Johnson, A. Boulle, D. Coetzee, and A. L. Williamson**. 2010. Influence of human immunodeficiency virus and CD4 count on the prevalence of human papillomavirus in heterosexual couples. J. Gen. Virol. **91**:3023-3031.

175. **Mbulawa, Z. Z., D. J. Marais, L. F. Johnson, D. Coetzee, and A. L. Williamson**. 2012. Impact of human immunodeficiency virus on the natural history of human papillomavirus genital infection in South African men and women. J. Infect. Dis. **206**:15-27.

176. **McBride, A. A.** 2008. Replication and partitioning of papillomavirus genomes. Adv. Virus Res. **72**:155-205.

177. **McDonald, A. C., L. Denny, C. Wang, W. Y. Tsai, T. C. Wright, Jr., and L. Kuhn**. 2012. Distribution of High-Risk Human Papillomavirus Genotypes among HIV-Negative Women with and without Cervical Intraepithelial Neoplasia in South Africa. PLoS. One. **7**:e44332.

178. **Meijer, C. J., P. J. Snijders, and P. E. Castle**. 2006. Clinical utility of HPV genotyping. Gynecol. Oncol. **103**:12-17.

179. **Mejlhede, N., J. Bonde, and A. Fomsgaard**. 2009. High frequency of multiple HPV types in cervical specimens from Danish women. APMIS **117**:108-114.

180. **Memiah, P., W. Mbuthia, G. Kiiru, S. Agbor, F. Odhiambo, S. Ojoo, and S. Biadgilign**. 2012. Prevalence and Risk Factors Associated with Precancerous Cervical Cancer Lesions among HIV-Infected Women in Resource-Limited Settings. AIDS Res. Treat. **2012**:953743.

181. **Mendoza, L., M. A. Picconi, S. Mirazo, P. Mongelos, G. Gimenez, J. Basiletti, and J. Arbiza**. 2013. Distribution of HPV-16 variants among isolates from Paraguayan women with different grades of cervical lesion. Int. J. Gynaecol. Obstet. **122(1)**:44-7

182. **Menendez, C., X. Castellsague, M. Renom, J. Sacarlal, L. Quinto, B. Lloveras, J. Klaustermeier, J. R. Kornegay, B. Sigauque, F. X. Bosch, and P. L. Alonso**. 2010. Prevalence and risk factors of sexually transmitted infections and cervical neoplasia in women from a rural area of southern Mozambique. Infect. Dis. Obstet. Gynecol. **2010**:609315.

183. **Michala, L., E. Argyri, E. Tsimplaki, A. Tsitsika, C. Bakoula, A. Antsaklis, and E. Panotopoulou**. 2012. Human Papilloma Virus infection in sexually active adolescent girls. Gynecol. Oncol. **126**:207-210.

184. **Moodley, J. R., M. Hoffman, H. Carrara, B. R. Allan, D. D. Cooper, L. Rosenberg, L. E. Denny, S. Shapiro, and A. L. Williamson**. 2006. HIV and pre-neoplastic and neoplastic lesions of the cervix in South Africa: a case-control study. BMC. Cancer **6**:135.

185. **Moodley, M.** 2006. Reduction in prevalence of invasive cervical cancer in KwaZulu-Natal, South Africa: impact of the human immunodeficiency virus epidemic. Int. J. Gynecol. Cancer **16**:1036-1040.

186. **Moodley, M. and S. Mould**. 2005. Invasive cervical cancer and human immunodeficiency virus (HIV) infection in KwaZulu-Natal, South Africa. J. Obstet. Gynaecol. **25**:706-710.

187. **Moody, C. A. and L. A. Laimins**. 2010. Human papillomavirus oncoproteins: pathways to transformation. Nat. Rev. Cancer **10**:550-560.

188. **Moreno-Acosta, P., M. Molano, A. Huertas, M. S. de Gomez, A. Romero, M. Gonzalez, B. M. Mercedes, and A. Garcia-Carranca**. 2008. A non-radioactive PCR-SSCP analysis allows to distinguish between HPV 16 European and Asian-American variants in squamous cell carcinomas of the uterine cervix in Colombia. Virus Genes **37**:22-30.

189. **Moscicki, A. B.** 2008. HPV Vaccines: today and in the Future. J. Adolesc. Health **43**:S26-S40.

190. **Moscicki, A. B., M. Schiffman, S. Kjaer, and L. L. Villa**. 2006. Chapter 5: Updating the natural history of HPV and anogenital cancer. Vaccine **24 Suppl 3**:S3-42-S3/51.

191. **Mqoqi, N., P. Kellet, F. Sitas, and M. Jula**. Incidence of histologically diagnosed cancer in South Africa 1998 - 1999. 2004. National Cancer Registry of South Africa. Johannesburg: National Health Laboratory Service.

192. **Munoz, N., F. X. Bosch, S. S. de, R. Herrero, X. Castellsague, K. V. Shah, P. J. Snijders, and C. J. Meijer**. 2003. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N. Engl. J. Med. **348**:518-527.

193. **Munoz, N., F. X. Bosch, X. Castellsague, M. Diaz, S. S. de, D. Hammouda, K. V. Shah, and C. J. Meijer**. 2004. Against which human papillomavirus types shall we vaccinate and screen? The international perspective. Int. J. Cancer **111**:278-285.

194. **Munoz, N., S. Franceschi, C. Bosetti, V. Moreno, R. Herrero, J. S. Smith, K. V. Shah, C. J. Meijer, and F. X. Bosch**. 2002. Role of parity and human papillomavirus in cervical cancer: the IARC multicentric case-control study. Lancet **359**:1093-1101.

195. **Munoz, N., X. Castellsague, A. B. de Gonzalez, and L. Gissmann**. 2006. Chapter 1: HPV in the etiology of human cancer. Vaccine **24 Suppl 3**:S3-1-S310.

196. **Narisawa-Saito, M. and T. Kiyono**. 2007. Basic mechanisms of high-risk human papillomavirus-induced carcinogenesis: roles of E6 and E7 proteins. Cancer Sci. **98**:1505-1511.

197. **Naucler, P., W. Ryd, S. Tornberg, A. Strand, G. Wadell, B. G. Hansson, E. Rylander, and J. Dillner**. 2007. HPV type-specific risks of high-grade CIN during 4 years of follow-up: a population-based prospective study. Br. J. Cancer **97**:129-132.

198. **Nelson, J. R., Y. C. Cai, T. L. Giesler, J. W. Farchaus, S. T. Sundaram, M. Ortiz-Rivera, L. P. Hosta, P. L. Hewitt, J. A. Mamone, C. Palaniappan, and C. W. Fuller**. 2002. TempliPhi, phi29 DNA polymerase based rolling circle amplification of templates for DNA sequencing. Biotechniques **Suppl**:44-47.

199. **O'Connor, M., S. Y. Chan, and H. U. Bernard.** 1995. Transcription factor binding sites in the long control region of genital HPVs, p. 21–40. *In* G. Myers, H. U. Bernard, H. Delius, J. Icenogle, C. Baker, A. Halpern, and C. Wheeler (ed.), Human papillomaviruses. Los Alamos National Laboratory, Los Alamos N.Mex.

200. **Oh, J. K., S. Franceschi, B. K. Kim, J. Y. Kim, Y. H. Ju, E. K. Hong, Y. C. Chang, S. H. Rha, H. H. Kim, J. H. Kim, C. Y. Kim, and H. R. Shin**. 2009. Prevalence of human papillomavirus and Chlamydia trachomatis infection among women attending cervical cancer screening in the Republic of Korea. Eur. J. Cancer Prev. **18**:56-61.

201. **Oh, J. K., Y. H. Ju, S. Franceschi, W. Quint, and H. R. Shin**. 2008. Acquisition of new infection and clearance of type-specific human papillomavirus infections in female students in Busan, South Korea: a follow-up study. BMC. Infect. Dis. **8**:13.

202. **Oliveira, L. H., L. S. Santos, E. F. Augusto, and F. P. Neves**. 2012. Identification of human papillomavirus type 53 L1, E6 and E7 variants in isolates from Brazilian women. Infect. Genet. Evol. **12**:71-76.

203. **Ong, C. K., S. Y. Chan, M. S. Campo, K. Fujinaga, P. Mavromara-Nazos, V. Labropoulou, H. Pfister, S. K. Tay, M. J. ter, L. L. Villa, and .** 1993. Evolution of human papillomavirus type 18: an ancient phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic groups. J. Virol. **67**:6424-6431.

204. **Ordonez, R. M., A. M. Espinosa, D. J. Sanchez-Gonzalez, J. rmendariz-Borunda, and J. Berumen**. 2004. Enhanced oncogenicity of Asian-American human papillomavirus 16 is associated with impaired E2 repression of E6/E7 oncogene transcription. J. Gen. Virol. **85**:1433-1444.

205. **Palefsky, J.** 2006. Biology of HPV in HIV infection. Adv. Dent. Res. **19**:99-105.

206. **Palefsky, J. M.** 2007. HPV infection in men. Dis. Markers **23**:261-272.

207. **Pande, S., N. Jain, B. K. Prusty, S. Bhambhani, S. Gupta, R. Sharma, S. Batra, and B. C. Das**. 2008. Human papillomavirus type 16 variant analysis of E6, E7, and L1 genes and long control region in biopsy samples from cervical cancer patients in north India. J. Clin. Microbiol. **46**:1060-1066.

208. **Parada, R., R. Morales, A. R. Giuliano, A. Cruz, X. Castellsague, and E. Lazcano-Ponce**. 2010. Prevalence, concordance and determinants of human papillomavirus infection among heterosexual partners in a rural region in central Mexico. BMC. Infect. Dis. **10**:223.

209. **Park, J. S., E. S. Hwang, C. J. Lee, C. J. Kim, J. G. Rha, S. J. Kim, S. E. Namkoong, and S. J. Um**. 1999. Mutational and functional analysis of HPV-16 URR derived from Korean cervical neoplasia. Gynecol. Oncol. **74**:23-29.

210. **Peto, J., C. Gilham, O. Fletcher, and F. E. Matthews**. 2004. The cervical cancer epidemic that screening has prevented in the UK. Lancet **364**:249-256.

211. **Picconi, M. A., L. V. Alonio, L. Sichero, V. Mbayed, L. L. Villa, J. Gronda, R. Campos, and A. Teyssie**. 2003. Human papillomavirus type-16 variants in Quechua aboriginals from Argentina. J. Med. Virol. **69**:546-552.

212. **Pientong, C., P. Wongwarissara, T. Ekalaksananan, P. Swangphon, P. Kleebkaow, B. Kongyingyoes, S. Siriaunkgul, K. Tungsinmunkong, and C. Suthipintawong**. 2013. Association of human papillomavirus type 16 long control region mutation and cervical cancer. Virol. J. **10**:30.

213. **Pista, A., A. Oliveira, A. Barateiro, H. Costa, N. Verdasca, and M. T. Paixao**. 2007. Molecular variants of human papillomavirus type 16 and 18 and risk for cervical neoplasia in Portugal. J. Med. Virol. **79**:1889-1897.

214. **Pista, A., A. Oliveira, N. Verdasca, and F. Ribeiro**. 2011. Single and multiple human papillomavirus infections in cervical abnormalities in Portuguese women. Clin. Microbiol. Infect. **17**:941-946.

215. **Pittayakhajonwut, D. and P. C. Angeletti**. 2010. Viral trans-factor independent replication of human papillomavirus genomes. Virol. J. **7**:123.

216. **Plummer, M., M. Schiffman, P. E. Castle, D. Maucort-Boulch, and C. M. Wheeler**. 2007. A 2-year prospective study of human papillomavirus persistence among women with a cytological diagnosis of atypical squamous cells of undetermined significance or low-grade squamous intraepithelial lesion. J. Infect. Dis. **195**:1582-1589.

217. **Prado, J. C., I. E. Calleja-Macias, H. U. Bernard, M. Kalantari, S. A. Macay, B. Allan, A. L. Williamson, L. P. Chung, R. J. Collins, R. E. Zuna, S. T. Dunn, R. Ortiz-Lopez, H. A. Barrera-Saldana, H. A. Cubie, K. Cuschieri, M. von Knebel-Doeberitz, G. I. Sanchez, F. X. Bosch, and L. L. Villa**. 2005. Worldwide genomic diversity of the human papillomaviruses-53, 56, and 66, a group of high-risk HPVs unrelated to HPV-16 and HPV-18. Virology **340**:95-104.

218. **Pyeon, D., S. M. Pearce, S. M. Lank, P. Ahlquist, and P. F. Lambert**. 2009. Establishment of human papillomavirus infection requires cell cycle progression. PLoS. Pathog. **5**:e1000318.

219. **Raiol, T., P. S. Wyant, R. M. de Amorim, D. M. Cerqueira, N. G. Milanezi, M. M. Brigido, L. Sichero, and C. R. Martins**. 2009. Genetic variability and phylogeny of the high-risk HPV-31, -33, -35, -52, and -58 in central Brazil. J. Med. Virol. **81**:685-692.

220. **Rector, A., R. Tachezy, and R. M. Van**. 2004. A sequence-independent strategy for detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle amplification. J. Virol. **78**:4993-4998.

221. **Riva, E., D. Serraino, A. Pierangeli, F. Bambacioni, S. Zaniratti, C. Minosse, M. Selleri, M. Bucci, C. Scagnolari, A. M. Degener, M. R. Capobianchi, G. Antonelli, and F. Dianzani**. 2007. Markers of human papillomavirus infection and their correlation with cervical dysplasia in human immunodeficiency virus-positive women. Clin. Microbiol. Infect. 13:94-97.

222. **Rodriguez, A. C., R. Burk, R. Herrero, A. Hildesheim, C. Bratti, M. E. Sherman, D. Solomon, D. Guillen, M. Alfaro, R. Viscidi, J. Morales, M. Hutchinson, S. Wacholder, and M. Schiffman**. 2007. The natural history of human papillomavirus infection and cervical intraepithelial neoplasia among young women in the Guanacaste cohort shortly after initiation of sexual life. Sex Transm. Dis. **34**:494-502.

223. **Rositch, A. F., J. Koshiol, M. G. Hudgens, H. Razzaghi, D. M. Backes, J. M. Pimenta, E. L. Franco, C. Poole, and J. S. Smith**. 2012. Patterns of persistent genital human papillomavirus infection among women worldwide: A literature review and meta-analysis. Int. J. Cancer. **133(6)**:1271-85.

224. **Rousseau, M. C., J. S. Pereira, J. C. Prado, L. L. Villa, T. E. Rohan, and E. L. Franco**. 2001. Cervical coinfection with human papillomavirus (HPV) types as a predictor of acquisition and persistence of HPV infection. J. Infect. Dis. **184**:1508-1517.

225. **Safaeian, M., M. Kiddugavu, P. E. Gravitt, S. J. Gange, J. Ssekasanvu, D. Murokora, M. Sklar, D. Serwadda, M. J. Wawer, K. V. Shah, and R. Gray**. 2008. Determinants of incidence and clearance of high-risk human papillomavirus infections in rural Rakai, Uganda. Cancer Epidemiol. Biomarkers Prev. **17**:1300-1307.

226. **Sandri, M. T., D. Riggio, M. Salvatici, R. Passerini, L. Zorzino, S. Boveri, D. Radice, N. Spolti, and M. Sideri**. 2009. Typing of human papillomavirus in women with cervical lesions: prevalence and distribution of different genotypes. J. Med. Virol. **81**:271-277.

227. **Santos, A. L., S. F. Derchain, M. R. Martins, L. O. Sarian, E. Z. Martinez, and K. J. Syrjanen**. 2003. Human papillomavirus viral load in predicting high-grade CIN in women with cervical smears showing only atypical squamous cells or low-grade squamous intraepithelial lesion. Sao Paulo Med. J. **121**:238-243.

228. **Sasieni, P., A. Castanon, and J. Cuzick**. 2009. Effectiveness of cervical screening with age: population based case-control study of prospectively recorded data. BMJ **339**:b2968.

229. **Schiffman, M. and P. E. Castle**. 2005. The promise of global cervical-cancer prevention. N. Engl. J. Med. **353**:2101-2104

230. **Schiffman, M. and S. K. Kjaer**. 2003. Chapter 2: Natural history of anogenital human papillomavirus infection and neoplasia. J. Natl. Cancer Inst. Monogr14-19.

231. **Schiffman, M., A. C. Rodriguez, Z. Chen, S. Wacholder, R. Herrero, A. Hildesheim, R. Desalle, B. Befano, K. Yu, M. Safaeian, M. E. Sherman, J. Morales, D. Guillen, M. Alfaro, M. Hutchinson, D. Solomon, P. E. Castle, and R. D. Burk**. 2010. A population-based prospective study of carcinogenic human papillomavirus variant lineages, viral persistence, and cervical neoplasia. Cancer Res. **70**:3159-3169.

232. **Schiffman, M., A. G. Glass, N. Wentzensen, B. B. Rush, P. E. Castle, D. R. Scott, J. Buckland, M. E. Sherman, G. Rydzak, P. Kirk, A. T. Lorincz, S. Wacholder, and R. D. Burk**. 2011. A long-term prospective study of type-specific human papillomavirus infection and risk of cervical neoplasia among 20,000 women in the Portland Kaiser Cohort Study. Cancer Epidemiol. Biomarkers Prev. **20**:1398-1409.

233. **Schiffman, M., G. Clifford, and F. M. Buonaguro**. 2009. Classification of weakly carcinogenic human papillomavirus types: addressing the limits of epidemiology at the borderline. Infect. Agent. Cancer **4**:8.

234. **Schiffman, M., P. E. Castle, J. Jeronimo, A. C. Rodriguez, and S. Wacholder**. 2007. Human papillomavirus and cervical cancer. Lancet **370**:890-907.

235. **Schlecht, N. F., R. W. Platt, E. Duarte-Franco, M. C. Costa, J. P. Sobrinho, J. C. Prado, A. Ferenczy, T. E. Rohan, L. L. Villa, and E. L. Franco**. 2003. Human papillomavirus infection and time to progression and regression of cervical intraepithelial neoplasia. J. Natl. Cancer Inst. **95**:1336-1343.

236. **Schmeink, C. E., W. J. Melchers, A. G. Siebers, W. G. Quint, L. F. Massuger, and R. L. Bekkers**. 2011. Human papillomavirus persistence in young unscreened women, a prospective cohort study. PLoS. One. **6**:e27937.

237. **Schmeink, C., L. Massuger, C. Lenselink, W. Quint, B. Witte, J. Berkhof, W. Melchers, and R. Bekkers**. 2013. Prospective follow-up of 2065 young unscreened women to study human papillomavirus incidence and clearance. Int. J. Cancer. **133(1)**:172-81.

238. **Schmidt, M. T., A. K. Olejnik, and A. Gozdzicka-Jozefiak**. 2005. The HPV16 E2 transcriptional regulator mode of action depends on the physical state of the viral genome. Acta Biochim. Pol. **52**:823-832.

239. **Schmidt, M., W. Kedzia, and A. Gozdzicka-Jozefiak**. 2001. Intratype HPV16 sequence variation within LCR of isolates from asymptomatic carriers and cervical cancers. J. Clin. Virol. **23**:65-77.

240. **Selva, L., E. Gonzalez-Bosquet, M. T. Rodriguez-Plata, C. Esteva, M. Sunol, and C. Munoz-Almagro**. 2009. Detection of human papillomavirus infection in women attending a colposcopy clinic. Diagn. Microbiol. Infect. Dis. **64**:416-421.

241. **Shang, Q., Y. Wang, Y. Fang, L. Wei, S. Chen, Y. Sun, B. Li, F. Zhang, and H. Gu**. 2011. Human papillomavirus type 16 variant analysis of E6, E7, and L1 [corrected] genes and long control region in [corrected] cervical carcinomas in patients in northeast China. J. Clin. Microbiol. **49**:2656-2663.

242. **Shank-Retzlaff, M. L., Q. Zhao, C. Anderson, M. Hamm, K. High, M. Nguyen, F. Wang, N. Wang, B. Wang, Y. Wang, M. Washabaugh, R. Sitrin, and L. Shi**. 2006. Evaluation of the thermal stability of Gardasil. Hum. Vaccin. **2**:147-154.

243. **Shapiro, S., L. Rosenberg, M. Hoffman, J. P. Kelly, D. D. Cooper, H. Carrara, L. E. Denny, T. G. du, B. R. Allan, I. A. Stander, and A. L. Williamson**. 2003. Risk of invasive cancer of the cervix in relation to the use of injectable progestogen contraceptives and combined estrogen/progestogen oral contraceptives (South Africa). Cancer Causes Control **14**:485-495.

244. **Sibbet, G. J., S. Cuthill, and M. S. Campo**. 1995. The enhancer in the long control region of human papillomavirus type 16 is up-regulated by PEF-1 and down-regulated by Oct-1. J. Virol. **69**:4006-4011.

245. **Sichero, L., J. S. Sobrinho, and L. L. Villa**. 2012. Oncogenic potential diverge among human papillomavirus type 16 natural variants. Virology **432**:127-132.

246. **Sichero, L., S. Ferreira, H. Trottier, E. Duarte-Franco, A. Ferenczy, E. L. Franco, and L. L. Villa**. 2007. High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18. Int. J. Cancer **120**:1763-1768.

247. **Silva, J., J. Ribeiro, H. Sousa, F. Cerqueira, A. L. Teixeira, I. Baldaque, T. Osorio, and R. Medeiros**. 2011. Oncogenic HPV Types Infection in Adolescents and University Women from North Portugal: From Self-Sampling to Cancer Prevention. J. Oncol. **2011**:953469.

248. **Simonetti, A. C., J. H. Melo, P. R. de Souza, D. Bruneska, and J. L. de Lima Filho**. 2009. Immunological's host profile for HPV and Chlamydia trachomatis, a cervical cancer cofactor. Microbes. Infect. **11**:435-442.

249. **Sitas, F., J. Madhoo, and J. Wessie**. 1998. Incidence of histological diagnosed cancer in South Africa 1993 - 1995. National Cancer Registry of South Africa. Johannesburg: South African Institute of Medical Research.

250. **Smith, J. S., A. Melendy, R. K. Rana, and J. M. Pimenta**. 2008. Age-specific prevalence of infection with human papillomavirus in females: a global review. J. Adolesc. Health **43**:S5-25, S25.

251. **Smith, J. S., R. Herrero, C. Bosetti, N. Munoz, F. X. Bosch, J. Eluf-Neto, X. Castellsague, C. J. Meijer, A. J. Van den Brule, S. Franceschi, and R. Ashley**. 2002. Herpes simplex virus-2 as a human papillomavirus cofactor in the etiology of invasive cervical cancer. J. Natl. Cancer Inst. **94**:1604-1613.

252. **Snijders, P. J., R. D. Steenbergen, D. A. Heideman, and C. J. Meijer**. 2006. HPV-mediated cervical carcinogenesis: concepts and clinical implications. J. Pathol. **208**:152-164.

253. **Spinillo, A., B. Dal, Bello. Gardella, M. Roccio, M. D. Dacco', and E. M. Silini**. 2009. Multiple human papillomavirus infection and high grade cervical intraepithelial neoplasia among women with cytological diagnosis of atypical squamous cells of undetermined significance or low grade squamous intraepithelial lesions. Gynecol. Oncol. **113**:115-119.

254. **Stanley, M., D. R. Lowy, and I. Frazer**. 2006. Chapter 12: Prophylactic HPV vaccines: underlying mechanisms. Vaccine **24 Suppl 3**:S3-106-S3/113.

255. **Stephen, A. L., C. H. Thompson, M. H. Tattersall, Y. E. Cossart, and B. R. Rose**. 2000. Analysis of mutations in the URR and E6/E7 oncogenes of HPV 16 cervical cancer isolates from central China. Int. J. Cancer **86**:695-701.

256. **Stewart, A. C., A. M. Eriksson, M. M. Manos, N. Munoz, F. X. Bosch, J. Peto, and C.M. Wheeler**. 1996. Intratype variation in 12 human papillomavirus types: a worldwide perspective. J. Virol. **70**:3127-3136.

257. **Strickler, H. D., J. M. Palefsky, K. V. Shah, K. Anastos, R. S. Klein, H. Minkoff, A. Duerr, L. S. Massad, D. D. Celentano, C. Hall, M. Fazzari, S. Cu-Uvin, M. Bacon, P. Schuman, A. M. Levine, A. J. Durante, S. Gange, S. Melnick, and R. D. Burk**. 2003. Human papillomavirus type 16 and immune status in human immunodeficiency virus-seropositive women. J. Natl. Cancer Inst. **95**:1062-1071.

258. **Strickler, H. D., R. D. Burk, M. Fazzari, K. Anastos, H. Minkoff, L. S. Massad, C. Hall, M. Bacon, A. M. Levine, D. H. Watts, M. J. Silverberg, X. Xue, N. F. Schlecht, S. Melnick, and J. M. Palefsky**. 2005. Natural history and possible reactivation of human papillomavirus in human immunodeficiency virus-positive women. J. Natl. Cancer Inst. **97**:577-586.

259. **Stunkel, W. and H. U. Bernard**. 1999. The chromatin structure of the long control region of human papillomavirus type 16 represses viral oncoprotein expression. J. Virol. **73**:1918-1930.

260. **Sun, M., L. Gao, Y. Liu, Y. Zhao, X. Wang, Y. Pan, T. Ning, H. Cai,H. Yang , W. Zhai W, Y. Ke.** 2012. Whole Genome Sequencing and Evolutionary Analysis of Human Papillomavirus Type 16 in Central China. PLoS ONE **7** (5): e36577.

261. **Swangvaree, S. S., P. Kongkaew, P. Rugsuj, and O. Saruk**. 2010. Prevalence of high-risk human papillomavirus infection and cytologic results in Thailand. Asian Pac. J. Cancer Prev. **11**:1465-1468.

262. **Syrjanen, S., I. Shabalova, N. Petrovichev, V. Kozachenko, T. Zakharova, J. Pajanidi, J. Podistov, G. Chemeris, L. Sozaeva, E. Lipova, I. Tsidaeva, O. Ivanchenko, A. Pshepurko, S. Zakharenko, R. Nerovjna, L. Kljukina, O. Erokhina, M. Branovskaja, M. Nikitina, V. Grunberga, A. Grunberg, A. Juschenko, P. Tosi, M. Cintorino, R. Santopietro, and K. Syrjanen**. 2004. Acquisition of high-risk human papillomavirus infections and pap smear abnormalities among women in the New Independent States of the Former Soviet Union. J. Clin. Microbiol. **42**:505-511.

263. **Szostek, S., M. Klimek, B. Zawilinska, and M. Kosz-Vnenchak**. 2008. Genotype-specific human papillomavirus detection in cervical smears. Acta Biochim. Pol. **55**:687-692.

264. **Tanzi, E., A. Amendola, S. Bianchi, M. M. Fasolo, R. Beretta, E. Pariani, A. Zappa, E. Frati, and G. Orlando**. 2009. Human papillomavirus genotypes and phylogenetic analysis of HPV-16 variants in HIV-1 infected subjects in Italy. Vaccine **27 Suppl 1**:A17-A23.

265. **Thomas, J. O., R. Herrero, A. A. Omigbodun, K. Ojemakinde, I. O. Ajayi, A. Fawole, O. Oladepo, J. S. Smith, A. Arslan, N. Munoz, P. J. Snijders, C. J. Meijer, and S. Franceschi**. 2004. Prevalence of papillomavirus infection in women in Ibadan, Nigeria: a population-based study. Br. J. Cancer **90**:638-645.

266. **Thomas, K. K., J. P. Hughes, J. M. Kuypers, N. B. Kiviat, S. K. Lee, D. E. Adam, and L. A. Koutsky**. 2000. Concurrent and sequential acquisition of different genital human papillomavirus types. J. Infect. Dis. **182**:1097-1102.

267. **Ting, J., D. T. Kruzikas, and J. S. Smith**. 2010. A global review of age-specific and overall prevalence of cervical lesions. Int. J. Gynecol. Cancer **20**:1244-1249.

268. **Tornesello, M. L., F. M. Buonaguro, A. Meglio, L. Buonaguro, E. Beth-Giraldo, and G. Giraldo**. 1997. Sequence variations and viral genomic state of human papillomavirus type 16 in penile carcinomas from Ugandan patients. J. Gen. Virol. **78 (Pt 9)** :2199-2208.

269. **Tornesello, M. L., F. M. Buonaguro, L. Buonaguro, I. Salatiello, E. Beth-Giraldo, and G. Giraldo**. 2000. Identification and functional analysis of sequence rearrangements in the long control region of human papillomavirus type 16 Af-1 variants isolated from Ugandan penile carcinomas. J. Gen. Virol. **81**:2969-2982.

270. **Tornesello, M. L., M. L. Duraturo, I. Salatiello, L. Buonaguro, S. Losito, G. Botti, G. Stellato, S. Greggi, R. Piccoli, S. Pilotti, B. Stefanon, P. G. De, S. Franceschi, and F. M. Buonaguro**. 2004. Analysis of human papillomavirus type-16 variants in Italian women with cervical intraepithelial neoplasia and cervical cancer. J. Med. Virol. **74**:117-126.

271. **Tornesello, M. L., M. L. Duraturo, S. Losito, G. Botti, S. Pilotti, B. Stefanon, P. G. De, A. Gallo, L. Buonaguro, and F. M. Buonaguro**. 2008. Human papillomavirus genotypes and HPV16 variants in penile carcinoma. Int. J. Cancer **122**:132-137.

272. **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.

273. **Tu, J. J., L. Kuhn, L. Denny, K. J. Beattie, A. Lorincz, and T. C. Wright, Jr.** 2006. Molecular variants of human papillomavirus type 16 and risk for cervical neoplasia in South Africa. Int. J. Gynecol. Cancer **16**:736-742.

274. **Vaccarella, S., R. Herrero, M. Dai, P. J. Snijders, C. J. Meijer, J. O. Thomas, P. T. Hoang Anh, C. Ferreccio, E. Matos, H. Posso, S. S. de, H. R. Shin, S. Sukvirach, E. Lazcano-Ponce, G. Ronco, R. Rajkumar, Y. L. Qiao, N. Munoz, and S. Franceschi**. 2006. Reproductive factors, oral contraceptive use, and human papillomavirus infection: pooled analysis of the IARC HPV prevalence surveys. Cancer Epidemiol. Biomarkers Prev. **15**:2148-2153.

275. **van Doorslaer. K., Q. Tan, S. Xirasagar, S. Bandaru, V. Gopalan, Y. Mohamoud, Y. Huyen, and A. A. McBride**. 2013. The Papillomavirus Episteme: a central resource for papillomavirus sequence data and analysis. Nucleic Acids Res. **41**:D571-D578.

276. **Veldhuijzen, N. J., N. Dhont, J. Vyankandondera, A. Gasarabwe, R. Busasa, T. Crucitti, and J. H. van de Wijgert** . 2012. Prevalence and concordance of HPV, HIV, and HSV-2 in heterosexual couples in Kigali, Rwanda. Sex Transm. Dis. **39**:128-135.

277. **Venturoli, S., S. Ambretti, M. Cricca, E. Leo, S. Costa, M. Musiani, and M. Zerbini**. 2008. Correlation of high-risk human papillomavirus genotypes persistence and risk of residual or recurrent cervical disease after surgical treatment. J. Med. Virol. **80**:1434-1440.

278. **Veress, G., K. Szarka, X. P. Dong, L. Gergely, and H. Pfister**. 1999. Functional significance of sequence variation in the E2 gene and the long control region of human papillomavirus type 16. J. Gen. Virol. **80 (Pt 4)**:1035-1043.

279. **Veress, G., M. Murvai, K. Szarka, A. Juhasz, J. Konya, and L. Gergely**. 2001. Transcriptional activity of human papillomavirus type 16 variants having deletions in the long control region. Eur. J. Cancer **37**:1946-1952.

280. **Verheijen, R. H.** 2011. Comparing bivalent and quadrivalent HPV vaccines. BMJ **343**:d5720.

281. **Vinodhini, K., S. Shanmughapriya, B. C. Das, and K. Natarajaseenivasan**. 2012. Prevalence and risk factors of HPV infection among women from various provinces of the world. Arch. Gynecol. Obstet. **285**:771-777.

282. **Vinokurova, S. and D. M. von Knebel**. 2011. Differential methylation of the HPV 16 upstream regulatory region during epithelial differentiation and neoplastic transformation. PLoS. One. **6**:e24451.

283. **von Knebel, D. M.** 2002. New markers for cervical dysplasia to visualise the genomic chaos created by aberrant oncogenic papillomavirus infections. Eur. J. Cancer **38**:2229-2242.

284. **von Knebel, D. M., T. Bauknecht, D. Bartsch, and H. zur Hausen**. 1991. Influence of chromosomal integration on glucocorticoid-regulated transcription of growth-stimulating papillomavirus genes E6 and E7 in cervical carcinoma cells. Proc. Natl. Acad. Sci. U. S. A **88**:1411-1415.

285. **Walboomers, J. M., M. V. Jacobs, M. M. Manos, F. X. Bosch, J. A. Kummer, K. V. Shah, P. J. Snijders, J. Peto, C. J. Meijer, and N. Munoz**. 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J. Pathol. **189**:12-19.

286. **Wang, S. S., R. E. Zuna, N. Wentzensen, S. T. Dunn, M. E. Sherman, M. A. Gold, M. Schiffman, S. Wacholder, R. A. Allen, I. Block, K. Downing, J. Jeronimo, J. D. Carreon, M. Safaeian, D. Brown, and J. L. Walker**. 2009. Human papillomavirus cofactors by disease progression and human papillomavirus types in the study to understand cervical cancer early endpoints and determinants. Cancer Epidemiol. Biomarkers Prev. **18**:113-120.

287. **Watts, K. J., C. H. Thompson, Y. E. Cossart, and B. R. Rose**. 2002. Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia. Int. J. Cancer **97**:868-874.

288. **Wheeler, C. M., S. K. Kjaer, K. Sigurdsson, O. E. Iversen, M. Hernandez-Avila, G. Perez, D. R. Brown, L. A. Koutsky, E. H. Tay, P. Garcia, K. A. Ault, S. M. Garland, S. Leodolter, S. E. Olsson, G. W. Tang, D. G. Ferris, J. Paavonen, M. Steben, F. X. Bosch, J. Dillner, E. A. Joura, R. J. Kurman, S. Majewski, N. Munoz, E. R. Myers, L. L. Villa, F. J. Taddeo, C. Roberts, A. Tadesse, J. Bryan, L. C. Lupinacci, K. E. Giacoletti, M. James, S. Vuocolo, T. M. Hesley, and E. Barr**. 2009. The impact of quadrivalent human papillomavirus (HPV; types 6, 11, 16, and 18) L1 virus-like particle vaccine on infection and disease due to oncogenic nonvaccine HPV types in sexually active women aged 16-26 years. J. Infect. Dis. **199**:936-944.

289. **Wheeler, C. M., W. C. Hunt, M. Schiffman, and P. E. Castle**. 2006. Human papillomavirus genotypes and the cumulative 2-year risk of cervical precancer. J. Infect. Dis. **194**:1291-1299.

290. **WHO/ICO HPV information centre**. 2010. Human papillomavirus and related cancer. Summary report update September 15, 2010.

291. **Widdice, L. E., D. J. Breland, J. Jonte, S. Farhat, Y. Ma, A. C. Leonard, and A. B. Moscicki**. 2010. Human papillomavirus concordance in heterosexual couples. J. Adolesc. Health **47**:151-159.

292. **Widdice, L., Y. Ma, J. Jonte, S. Farhat, D. Breland, S. Shiboski, and A. B. Moscicki**. 2013. Concordance and Transmission of Human Papillomavirus within Heterosexual Couples Observed over Short Intervals. J. Infect. Dis. **207(8)**:1286-94.

293. **Wilson, R., G. B. Ryan, G. L. Knight, L. A. Laimins, and S. Roberts**. 2007. The full-length E1E4 protein of human papillomavirus type 18 modulates differentiation-dependent viral DNA amplification and late gene expression. Virology **362**:453-460.

294. **Wilson, V. G., M. West, K. Woytek, and D. Rangasamy**. 2002. Papillomavirus E1 proteins: form, function, and features. Virus Genes **24**:275-290.

295. **Winer, R. L., N. B. Kiviat, J. P. Hughes, D. E. Adam, S. K. Lee, J. M. Kuypers, and L. A. Koutsky**. 2005. Development and duration of human papillomavirus lesions, after initial infection. J. Infect. Dis. **191**:731-738.

296. **Wright, T. C., Jr., L. S. Massad, C. J. Dunton, M. Spitzer, E. J. Wilkinson, and D. Solomon**. 2007. 2006 consensus guidelines for the management of women with abnormal cervical cancer screening tests. Am. J. Obstet. Gynecol. **197**:346-355.

297. **Wu, E. Q., X. Zha, X. H. Yu, G. N. Zhang, Y. G. Wu, Y. Fan, Y. Ren, L. Q. Kong, and W. Kong**. 2009. Profile of physical status and gene variation of human papillomavirus 58 genome in cervical cancer. J. Gen. Virol. **90**:1229-1237.

298. **Wyant, P. S., D. M. Cerqueira, D. S. Moraes, J. P. Leite, C. R. Martins, B. M. de Macedo, and T. Raiol**. 2011. Phylogeny and polymorphism in the long control region, E6, and L1 of human papillomavirus types 53, 56, and 66 in central Brazil. Int. J. Gynecol. Cancer **21**:222-229.

299. **Xi, L. F., J. J. Carter, D. A. Galloway, J. Kuypers, J. P. Hughes, S. K. Lee, D. E. Adam, N. B. Kiviat, and L. A. Koutsky**. 2002. Acquisition and natural history of human papillomavirus type 16 variant infection among a cohort of female university students. Cancer Epidemiol. Biomarkers Prev. **11**:343-351.

300. **Xi, L. F., J. P. Hughes, P. E. Castle, Z. R. Edelstein, C. Wang, D. A. Galloway, L. A. Koutsky, N. B. Kiviat, and M. Schiffman**. 2011. Viral load in the natural history of human papillomavirus type 16 infection: a nested case-control study. J. Infect. Dis. **203**:1425-1433.

301. **Yamada, T., C. M. Wheeler, A. L. Halpern, A. C. Stewart, A. Hildesheim, and S. A. Jenison**. 1995. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments. J. Virol. **69**:7743-7753.

302. **Yamada, T., M. M. Manos, J. Peto, C. E. Greer, N. Munoz, F. X. Bosch, and C. M. Wheeler**. 1997. Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. J. Virol. **71**:2463-2472.

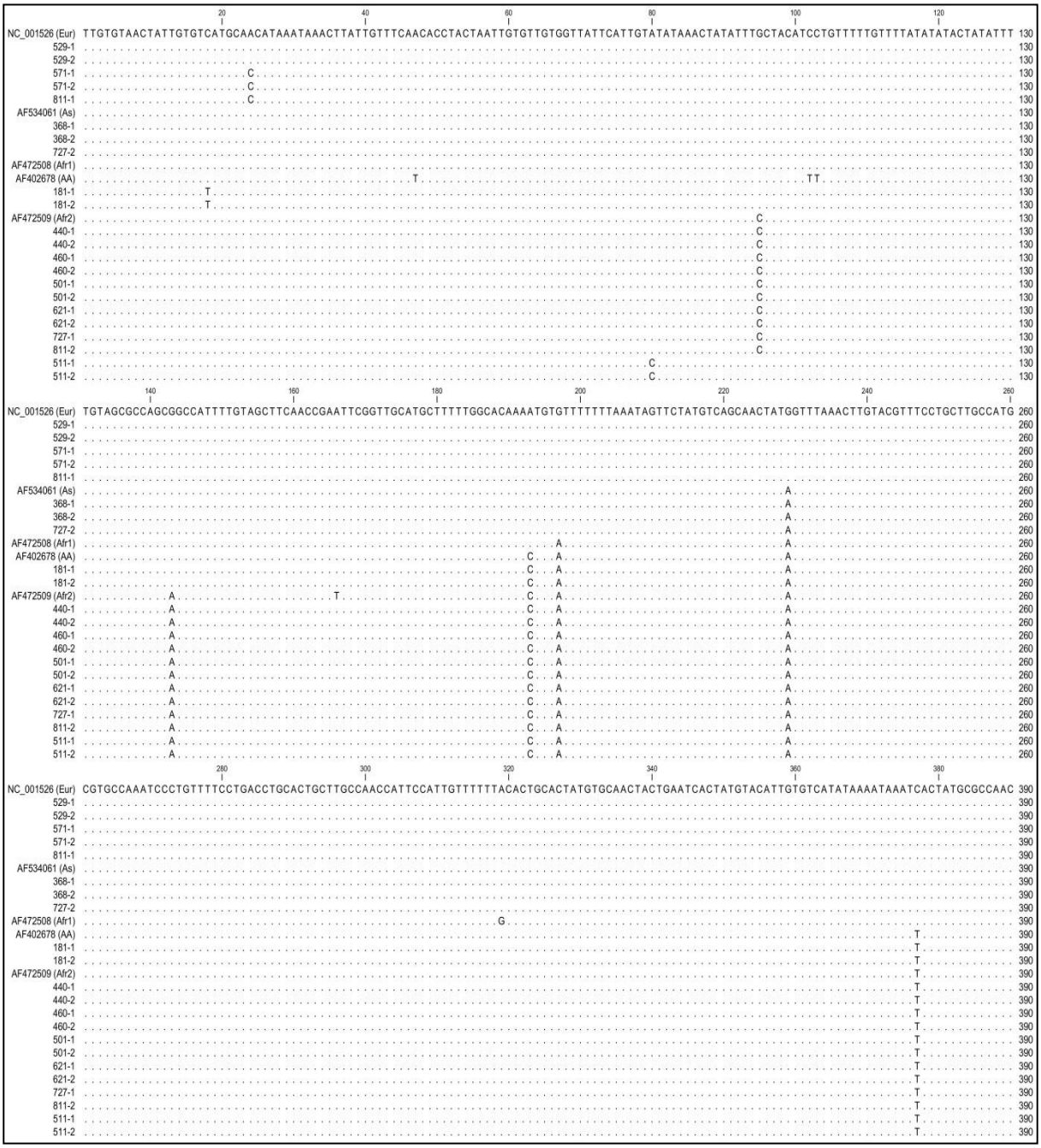
303. **Zhang, T., Y. Xu, L. Qiao, Y. Wang, X. Wu, D. Fan, Q. Peng, and X. Xu**. 2010. Trivalent Human Papillomavirus (HPV) VLP vaccine covering HPV type 58 can elicit high level of humoral immunity but also induce immune interference among component types. Vaccine **28**:3479-3487.

304. **H. zur Hausen.** 2002. Papillomaviruses and cancer: from basic studies to clinical application. Nat. Rev. Cancer **2**:342-350.

APPENDIX

HPV 16, 58 and 53 alignments

HPV 16



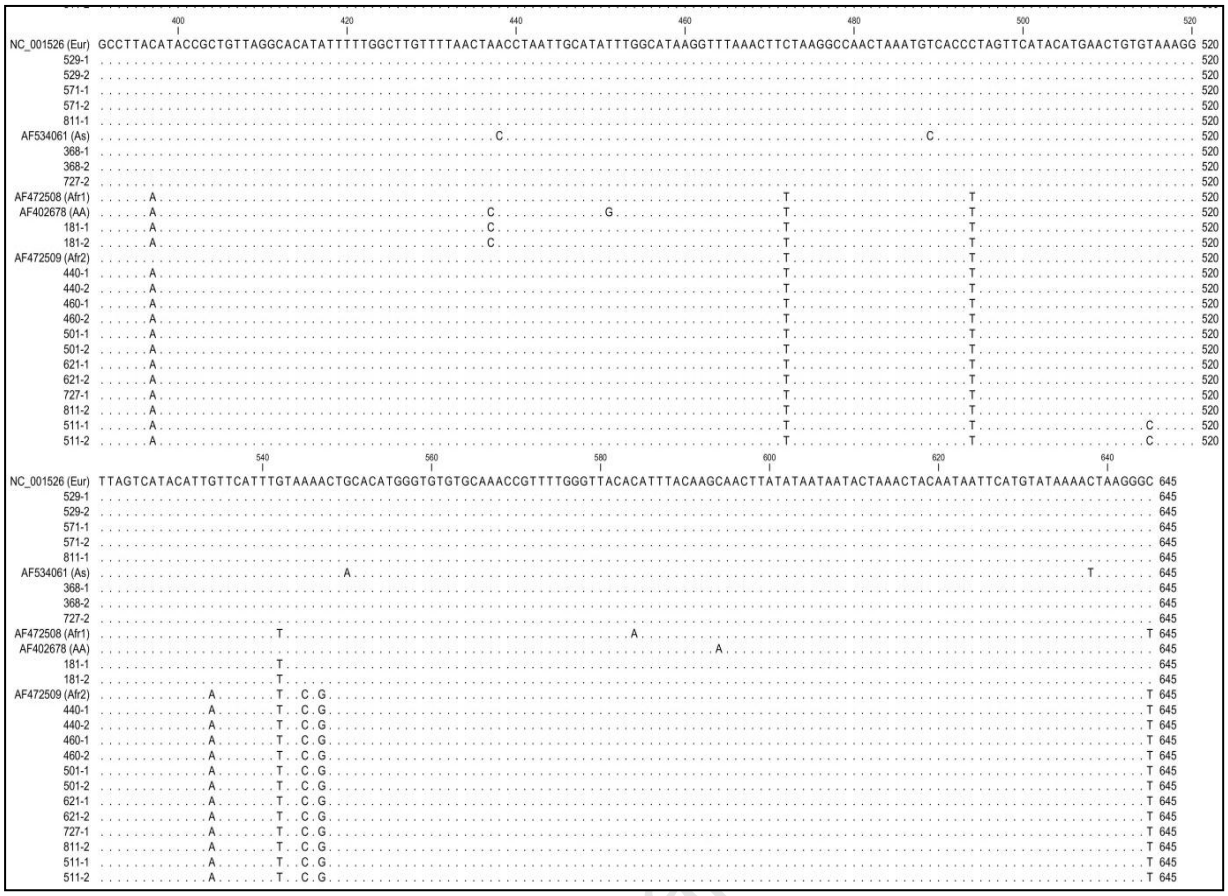
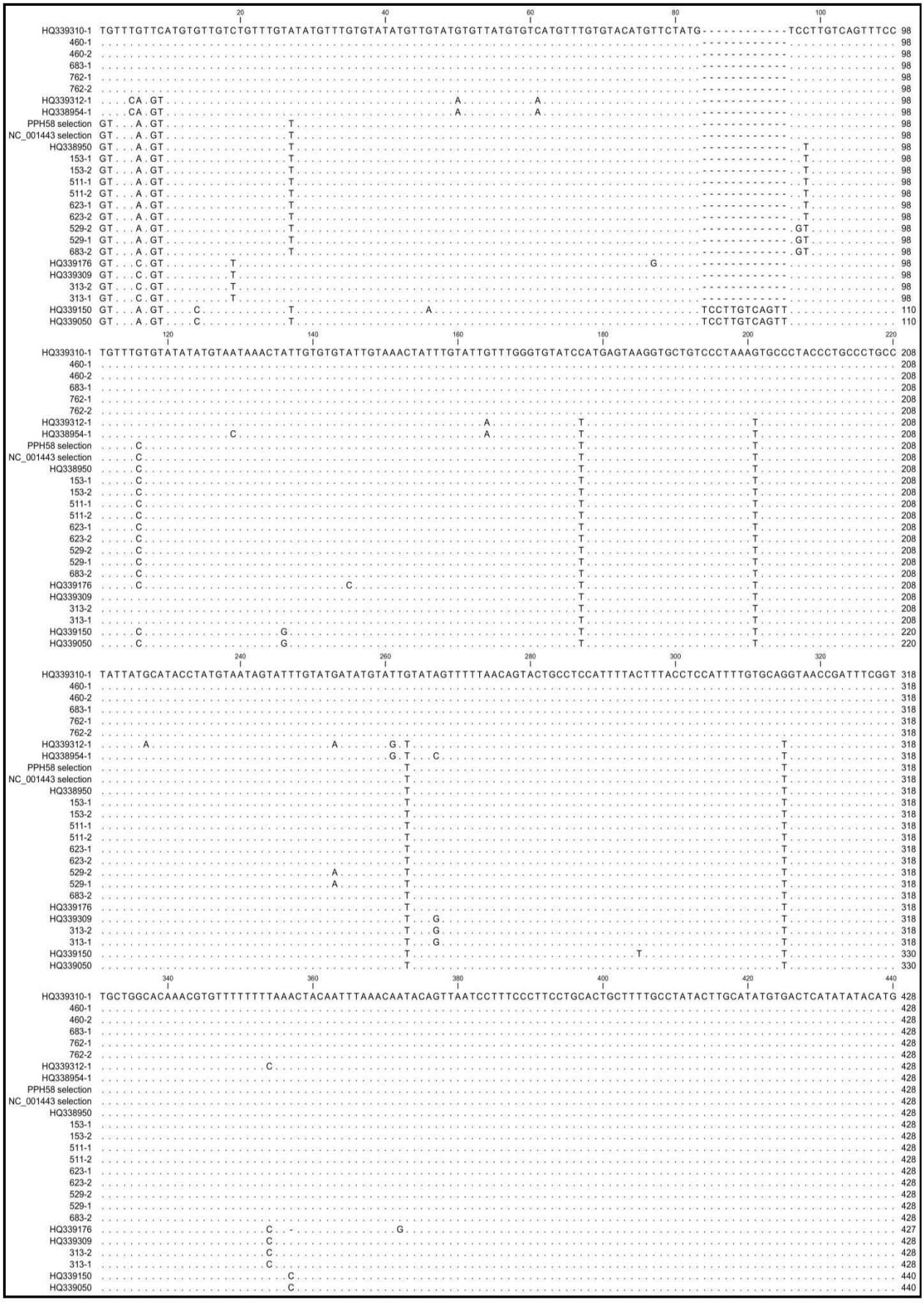


Figure 1 Alignment of HPV 16R (Eur), Asian American, Asian, African1 and African2, GenBank accession numbers; NC_001526 and AF402678 (AA), AF534061 (As), AF472508 (Afr1) and AF472509 (Afr2) respectively aligned with LCR sequences from couples; 181, 368, 440, 460, 501, 511, 529, 571, 621, 626, 727 and 811.

HPV 58


HQ339310-1 TGTTTGTTCATGTGTTGTCTGTTTGTATATGTTTGTGTATATGTTGTATGTGTTATGTGTCATGTTTGTGTACATGTTCTATG----------TCCTTGTCAGTTTCC 98
460-1 98
460-2 98
683-1 98
762-1 98
762-2 98
HQ339312-1 ....CA.GT..........A.........A..........98
HQ338954-1 ....CA.GT..........A.........A..........98
PPH58 selection GT...A.GT..........T..........98
NC_001443 selection GT...A.GT..........T..........98
HQ338950 GT...A.GT..........T..........T98
153-1 GT...A.GT..........T..........T98
153-2 GT...A.GT..........T..........T98
511-1 GT...A.GT..........T..........T98
511-2 GT...A.GT..........T..........T98
623-1 GT...A.GT..........T..........T98
623-2 GT...A.GT..........T..........T98
529-2 GT...A.GT..........T..........GT98
529-1 GT...A.GT..........T..........GT98
683-2 GT...A.GT..........T..........GT98
HQ339176 GT...C.GT.....T..........G..........98
HQ339309 GT...C.GT.....T..........98
313-2 GT...C.GT.....T..........98
313-1 GT...C.GT.....T..........98
HQ339150 GT...A.GT...C.....T.....A..........TCCTTGTCAGTT..110
HQ339050 GT...A.GT...C.....T..........TCCTTGTCAGTT..110
HQ339310-1 TGTTTGTGTATATATGTAATAAACTATTGTGTGTATTGTAAACTATTTGTATTGTTTGGGTGTATCCATGAGTAAGGTGCTGTCCCTAAAGTGCCCTACCCTGCCCTGCC 208
460-1 208
460-2 208
683-1 208
762-1 208
762-2 208
HQ339312-1 ..........A.....T.....T..........208
HQ338954-1 .....C.....A.....T.....T..........208
PPH58 selection ....C.....T.....T..........208
NC_001443 selection ....C.....T.....T..........208
HQ338950 ....C.....T.....T..........208
153-1 ....C.....T.....T..........208
153-2 ....C.....T.....T..........208
511-1 ....C.....T.....T..........208
511-2 ....C.....T.....T..........208
623-1 ....C.....T.....T..........208
623-2 ....C.....T.....T..........208
529-2 ....C.....T.....T..........208
529-1 ....C.....T.....T..........208
683-2 ....C.....T.....T..........208
HQ339176 ....C.....C.....T.....T..........208
HQ339309 ..........T.....T..........208
313-2 ..........T.....T..........208
313-1 ..........T.....T..........208
HQ339150 ....C.....G.....T.....T..........220
HQ339050 ....C.....G.....T.....T..........220
HQ339310-1 TATTATGCATACCTATGTAATAGTATTTGTATGATATGTATTGTATAGTTTTTAACAGTACTGCCTCCATTTTACTTTACCTCCATTTTGTGCAGGTAACCGATTTCGGT 318
460-1 318
460-2 318
683-1 318
762-1 318
762-2 318
HQ339312-1 .....A.....A.....G.T.....T..........318
HQ338954-1 .....G.T...C.....T..........318
PPH58 selection .....T.....T..........318
NC_001443 selection .....T.....T..........318
HQ338950 .....T.....T..........318
153-1 .....T.....T..........318
153-2 .....T.....T..........318
511-1 .....T.....T..........318
511-2 .....T.....T..........318
623-1 .....T.....T..........318
623-2 .....T.....T..........318
529-2 .....A.....T.....T..........318
529-1 .....A.....T.....T..........318
683-2 .....T.....T..........318
HQ339176 .....T.....T..........318
HQ339309 .....T...G.....T..........318
313-2 .....T...G.....T..........318
313-1 .....T...G.....T..........318
HQ339150 .....T.....T.....T..........330
HQ339050 .....T.....T..........330
HQ339310-1 TGCTGGCACAAACGTGTTTTTTTTAAACTACAATTTAAACAATACAGTTAATCCTTTCCCTTCCTGCACTGCTTTTGCCTATACTTGCATATGTGACTCATATATACATG 428
460-1 428
460-2 428
683-1 428
762-1 428
762-2 428
HQ339312-1 ....C.....428
HQ338954-1 428
PPH58 selection 428
NC_001443 selection 428
HQ338950 428
153-1 428
153-2 428
511-1 428
511-2 428
623-1 428
623-2 428
529-2 428
529-1 428
683-2 428
HQ339176 ..........C..-.....G.....427
HQ339309 ....C.....428
313-2 ....C.....428
313-1 ....C.....428
HQ339150 ....C.....440
HQ339050 ....C.....440

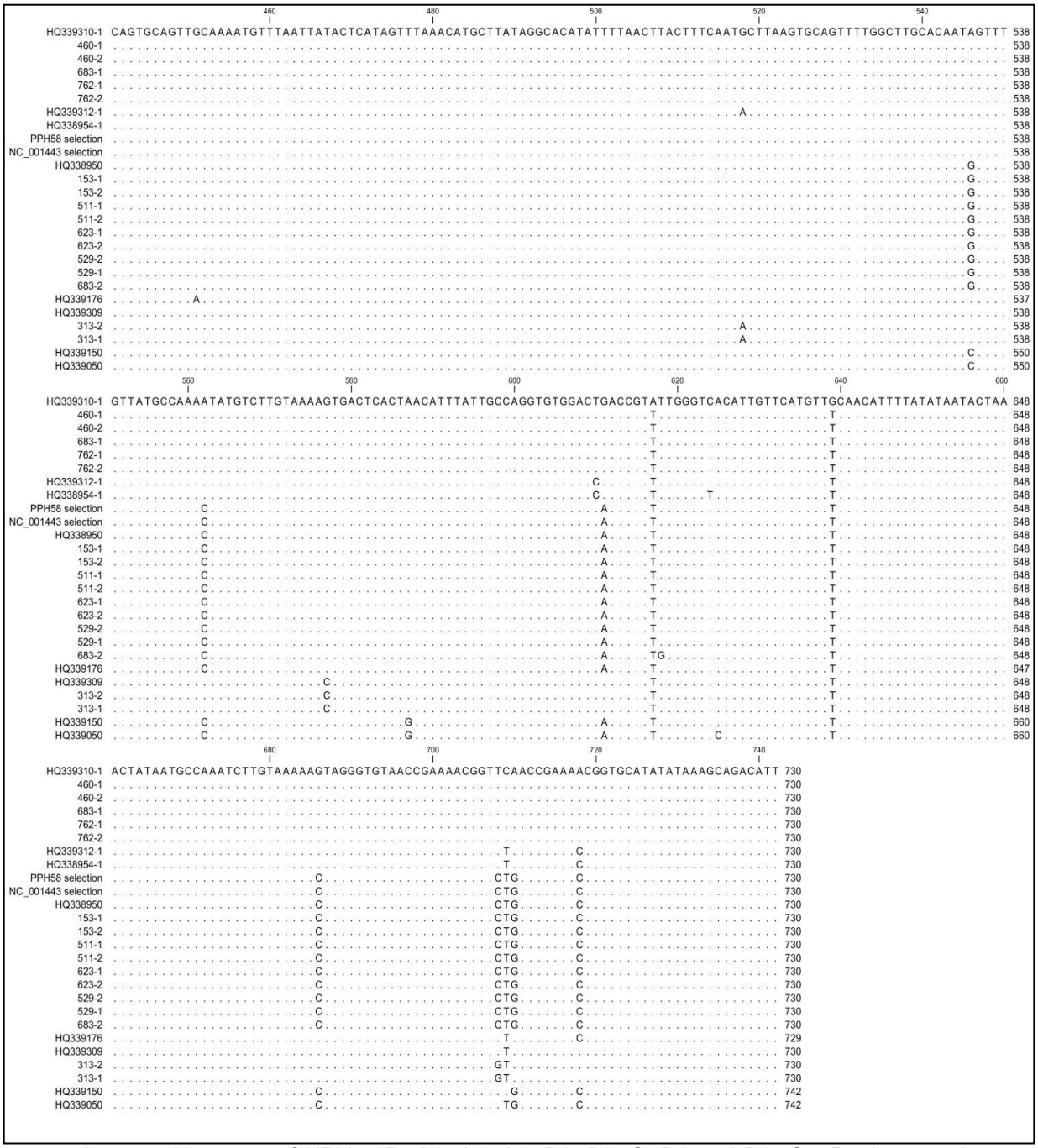


Figure 2 Alignment of HPV 58R, A1, A2, A3, B1, B2, C, D1 and D2, GenBank accession number; NC_001443/D90400, HQ339045 (A1), HQ338950 (A2), QH339150 (A3), HQ339176 (B1), HQ339309 (B2), HQ339310 (C), HQ338954 (D1) and HQ339312 (D2), aligned with LCR sequences from couples; 153, 313, 460, 511, 529, 623, 683 and 762.

HPV 53

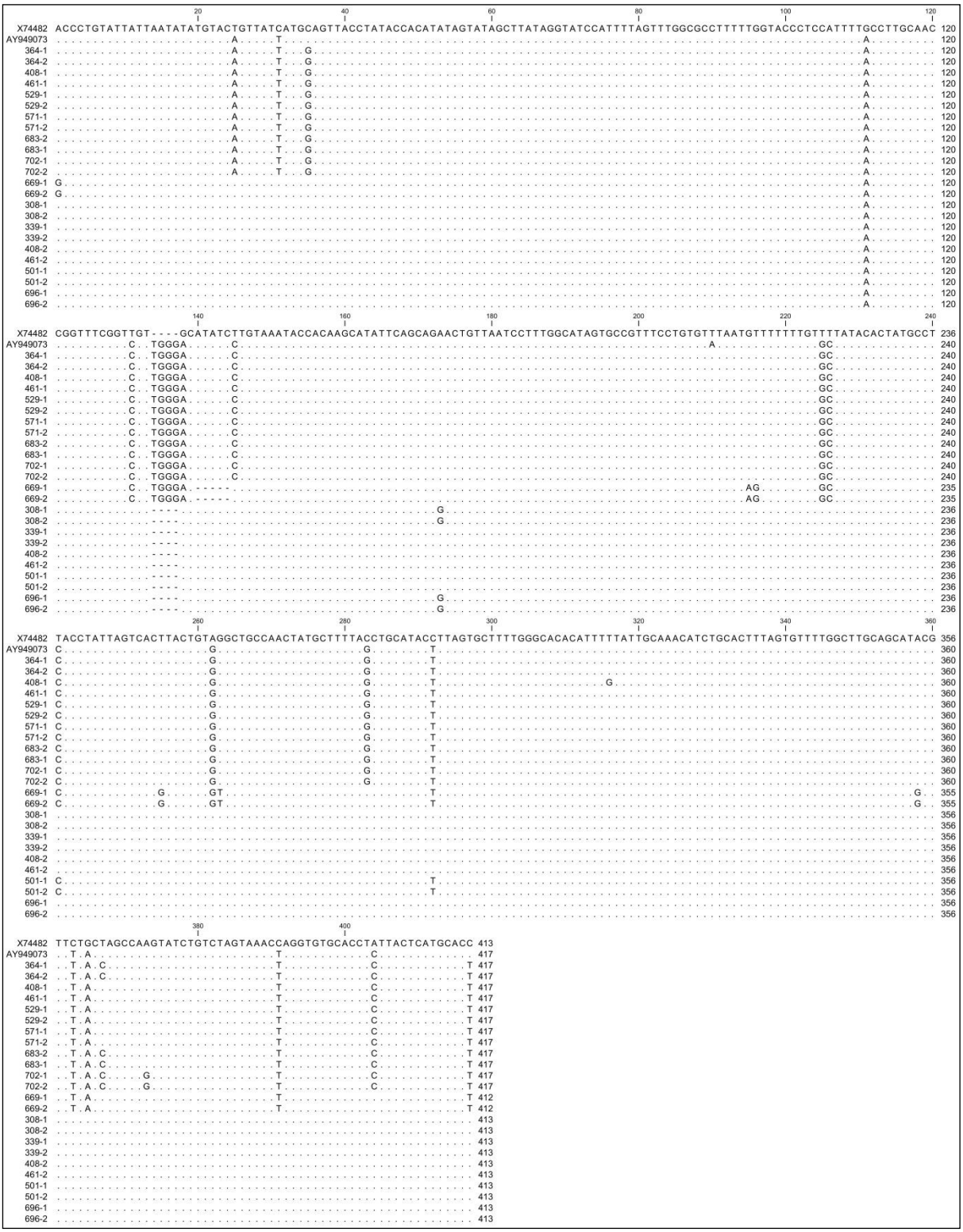


Figure 3 Alignment of HPV 53R (P-L) and non prototype-like (non P-L) variants, GenBank accession number X74482 and AY949073, alighned with LCR sequences from couples; 308, 339, 364, 408, 461, 501, 529, 571, 669, 683, 696 and 702.