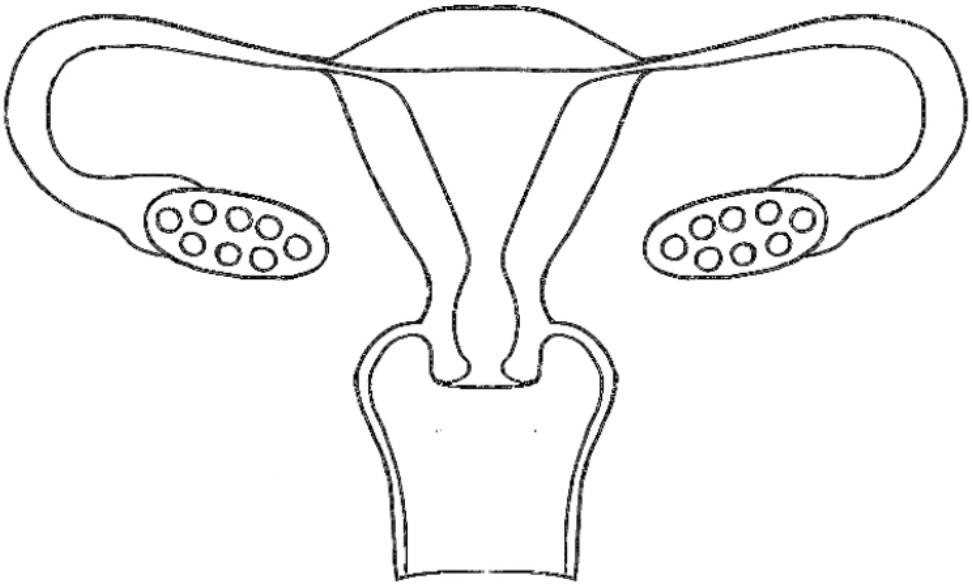
UPPSALA
UNIVERSITET

# Expansion of a previously developed real-time PCR system for detection of human papilloma virus to include additional virus types



Ivana Juko

---

Degree project in biology, Master of science (2 years), 2008
Examensarbete i biologi 30 hp till masterexamen, 2008
Biology Education Centre and Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University
Supervisors: Ulf Gyllensten and Inger Gustavsson

# Table of Contents

# Abbreviations

| | |
|---|---|
| ASC | Atypical Squamous Cells |
| Bp | Base pairs |
| CIN | Cervical Intraepithelial Neoplasm |
| dNTP | Deoxyribonucleoside triphosphatase |
| E1 ect. | Early genes |
| L1 ect. | Late genes |
| LCR | Long Control Region |
| LSIL | Low-grade squamous intraepithelial lesions of the cervix |
| LR | Low-risk |
| HPV | Human papillomavirus |
| HR | High-risk |
| HSIL | High-grade Squamous Intraepithelial Lesions |
| SDS | Sequence detection system |
| VIN | Vulval Intra-epithelial Neoplasia |
| Nt | Nucleotides |

# Summary

Cervical cancer is a known woman's disease, with 500 000 new cases reported annually worldwide. Only 20 years ago scientists confirmed that cervical cancer is caused by human papillomavirus (HPV). Today, in gynecological examinations Papanicolaou-stained (Pap) smears are used to detect cytological differences that are caused by the HPV infection. In Uppsala, 95% of the women do not show cytological differences in the primary screening. Among the positive 5%, 3% show atypical squamous cells (ASCUS) or cervical intraepithelial neoplasm (CIN) 1, and the remaining 2% CIN 2 or 3, which is moderate to severe cell abnormalities. Women in the positive group are invited to a second examination to monitor the cells and on that occasion they are also tested for HPV. Before 2007, the HPV high-risk types (16, 18, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73 and 82) were detected by a commercial kit (Digene Hybrid Capture 2), which detects high-risk HPVs in a combined measure. After 2007, the method for analyzing HPV has been changed to a real-time PCR based method, *hpVIR*, which identifies individual or groups of HPV types as well as viral load. The HPV types 16, 18, 31, 45, 33, 35, 39, 52 and 58 are detectable with *hpVIR* assay today.

In this study the focus has been on further developing *hpVIR* to detect also HPV types 51, 56 and 59 and if possible HPV6 and HPV11. Two different versions were compared, version 1 detecting HPV types 51, 56 and 59 and version 2 detecting HPV types 56 or 51, HPV59 and HPV6 or 11. DNA from the different HPV types were first PCR amplified and cloned, and plasmid DNA prepared. For each of the plasmids used a standard curve was made using reverse transcription PCR to determine the sensivity and specificity of the HPV assay. The standard curves for version 1 were sufficient and for version 2 the HPV59 had an adequate standard curve, but not HPV56, typed together with HPV51 and HPV6 typed together with HPV11. Upon testing the new developed *hpVIR* assays on 400 patient samples, I concluded that version 1 is a better HPV typing assay compared to version 2, because 12 patients were missed with version 2. Thus, *hpVIR* assay version 1 will be used in the clinical application, and further optimization of version 2 is needed. Among the tested patients HPV56 was most frequent, followed by HPV51 and HPV59. This suggests that this is the frequency found in the Swedish population, but more data are needed for a firm conclusion.

# 1. Introduction

## 1.1 Cervical cancer

Only 20 years ago scientist's confirmed that cervical cancer is caused by human papillomavirus (HPV), (zur Hausen 2002, Hazard 2007) but even today many people are still unaware of this fact. HPV infects 50% of women and men during their lifetime (Paavonen 2007). It is therefore not surprising that 500 000 new cases of cervical cancer are reported annually worldwide. Cervical cancer is also the cause of death for 250 000 women every year. Globally 80% of both new cervical cancer cases and deaths are found in the developing countries. (Parkin *et al* 2005)

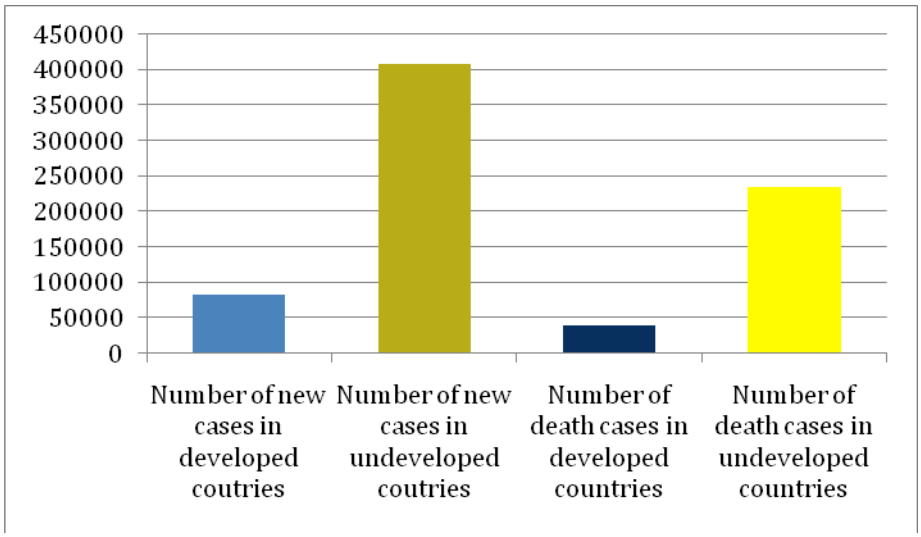



**Figure 1. Schematic diagram showing number of new cervical cancer and death cases reported each year around the world.**

In the early 1980's, scientists confirmed that cervical cancer is caused by HPV and is sexually transmitted (Hazard 2007). In many parts of the world, HPV is the most common sexually transmitted infection. About 20-46% of young sexually active women in ages 16-25 have been infected with at least one of the known HPV types (Schiffman *et al* 1995). In a study done in Finland it was discovered that the median time from sexual debut to first detection of HPV was 3 months and the 3-year cumulative risk of acquiring cervical HPV was 46% (Paavonen 2007). Even though the women were infected at an early age, they rarely developed invasive cervical cancer until their mid-thirties or even later. Given that the HPV infection is asymptomatic, most individuals are unaware about the infection until discovered during a gynecological examination.

HPV-associated cervical cancer is rare in women. In fact, 70-90% of all women affected will clear the infection within 12 to 30 months (Cutts *et al* 2007). Cervical cancer arises in the epithelial cells of the cervix uterus that is the lower part of the uterus, where it forms a narrow canal between the vagina and the uterine body cavity. (Fig. 2)

Cervical cancer may be divided into two types: Cervical squamous cell carcinoma, derived from squamous cells and cervical adenocarcinomas, arising in the glandular cells of cervix. (Burghardt *et al* 1983)

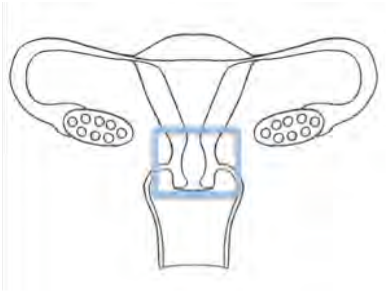


**Figure 2**. **A schematic picture of uterus.** The blue box shows the area where cervix cancer arises (Illustration by S. Stenquist).

## 1.2. Human papillomavirus

Papillomaviruses occur in humans as well as animals. Thus far, no cross infection cases have been reported, and viral infection is specific to the respective hosts. Today there are over 100 identified human papillomaviruses (HPV). About 40 of them infect the epithelial and mucosa lining of the anogenital tract and other areas (Steben 2007). Human papillomavirus types may be divided into high-risk (HR) (16, 18, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73 and 82) and low-risk (LR) types (6, 11, 40, 42, 43, 44, 54, 61, 70, 72, and 81). Thus, if you are infected with some of the HR types, you are at higher risk to develop cervical cancer compared to having an infection with a LR type (Munoz *et al* 2007).

Papillomaviruses were previously clustered together in the same family, the *Papovaviridae*. Eventually, scientists recognized that there are two viral groups that differ in genome size genome organization and also lack any major nucleotide (nt) similarities. These findings led to assignment of two families, *Papillomaviridae* and *Polyomaviridae*. Today, HPVs are classified according to sequence homology. HPVs with a similarity of 2-10% in the *L1* gene, are considered variants of the same type whereas those that differ by more that 10% are classified as separate HPV types. The *L1* gene is used to determine similarity between HPV types because it is the most conserved region within all HPV genomes (de Villiers *et al* 2004).

All of the HR and LR HPV types belong to the same genus, called alpha-papillomaviruses, and are divided into species. It is noteworthy that HPV6 and HPV11, which are low-risk types, are not associated with cervical cancer. They are found to be involved in the majority of benign lesions affecting the anogenital areas, such as genital warts (condyloma) and low-grade squamous intraepithelial lesions of the cervix (LSIL) and vulva (VIN1). They are also found in external genital warts. (Steben *et al* 2007)

### 1.2.1 Life cycle of Human papillomavirus

Human papillomavirus is a small, nonenveloped virus, with a DNA genome of 8000 base pairs (bp). The virus looks like a golf ball in electron microscope. The genome contains eight different genes, six of which code for nonstructural early proteins (E1, E2, E3, E4, E6, E7) and two code for structural late proteins (L1, L2) (Paavonen 2007). The long control region (LCR) is a 1000 bp region that separates the early and late genes, but does not code for any proteins. However, it contains *cis*-elements that are necessary for regulation of gene expression, replication of the genome and packaging of the genome to viral particles (Munoz *et al* 2006). Each individual protein has its own function and purpose in the life cycle of the virus.

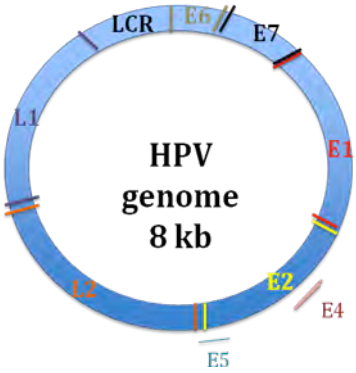



**Figure 3. The genomic organization of HPV.** The human papillomavirus genome contains 8 000 base pairs and eight different genes, which are divided into early (*E1, E2, E4, E5,* E6 and *E7*) and late genes (*L1* and *L2*). A long control region (LCR) separates these genes (modified from Syrjänen 2006).

Microtraumas, mild scratches in the epidermis, enable the virus to enter the cells of the basal layer in epithelium, which is the first step of the HPV replication cycle (Burd 2003). HPV uses the host cell machinery to replicate (Munoz *et al* 2006). In the differentiated keratinocytes of the suprabasal layers of the epithelium, the virus starts to increase its DNA production and also synthesize capsid proteins. The HPV encoded proteins make use of host cell factors to regulate viral transcription and replication. (Burd 2003)

Initially, host cell factors interact with the LCR region of the HPV genome, which leads to transcription of the *E6* and *E7* genes. The proteins E6 and E7 destabilize the cell growth-regulatory pathways by targeting two cellular proteins, tumor suppressor protein (p53) and the retinoblastoma gene product (pRB) that controls the cell growth. The E6 protein targets p53, to which it binds and marks it for fast degradation. As a result, p53 loses it normal functions to control G1 arrest, apoptosis, and DNA repair. E7 binds to the cellular protein pRB and destroys the complex formed between pRB and E2F-1, which is a cellular transcription factor. When the complex is dismantled, transcription of the genes whose products are necessary for the cell to enter the S phase of the cell cycle starts (zur Hausen 2002, Staubenrauch *et al* 1999). Thus, events caused by E6 and E7 proteins facilitate cellular DNA production and cell proliferation.

Products of the *E5* gene cause continuous host cell proliferation, but the differentiation is delayed. Next the E2 proteins bind to DNA and block the transcription of the *E6* and *E7* genes, leading to transcription of *E1* gene. E1 proteins bind to the viral origin of replication, which is found within the LCR region. The binding of E1 proteins initiates replication of the viral DNA as extrachromosomal elements in the S phase of the cell cycle. Because the E2 proteins down-regulate the production of E6 and E7 proteins, p53 and pRB will be released, leading to differentiation of the host cell and start of its normal function. A late promotor activates the late genes *L1* and *L2*. Viral particles are assembled in the nucleus, and complete virons are released from the mature squamous layer of the epithelium. The last of the eight proteins encoded by the HPV genome is E4, which plays an important role in the maturation and release of HPV particles (Burd 2003).

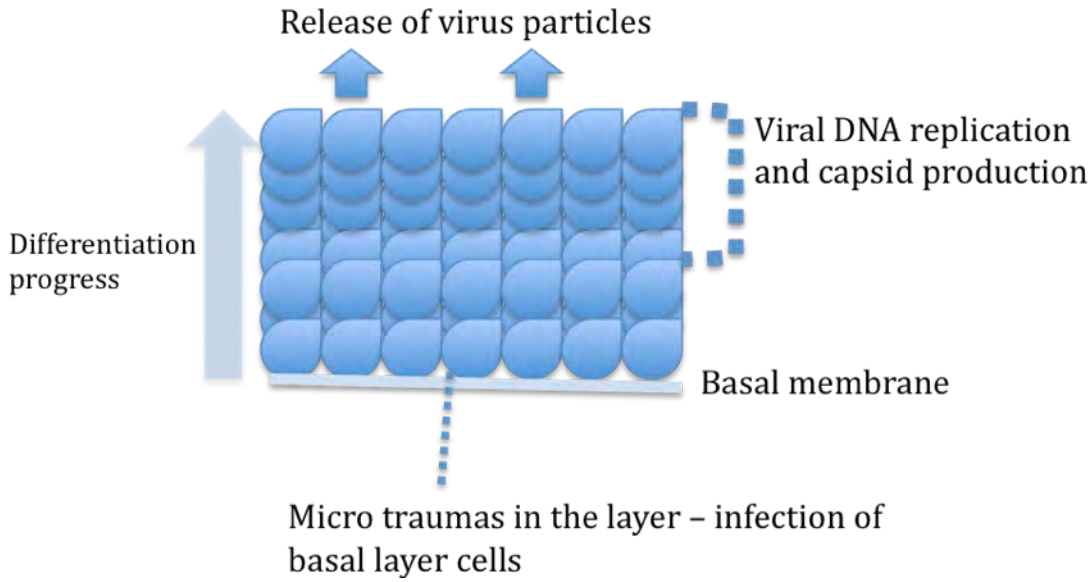



**Figure 4. HPV life cycle.** The basal layer cells need to be reached for successful infection. This is achieved through microlesions of the skin or mucosa. The infected cells spread into suprabasal differentiating cell layers, where viral genes are activated, viral DNA is replicated and capsid proteins are formed. Particles are released at the surface and might then infect additional tissues (modified from zur Hausen 2002).

## 1.3 The immune response and human papillomavirus

The immune system plays an important role in protecting the human body from different infections and attacks. The immune system is divided into the adaptive and the innate response. It is difficult to study the natural response of the immune system to HPV, especially because the natural life cycle and the gene expression of the virus depend on the differentiation stages of keratinocytes and also because HPV is hard to cultivate *in vitro* (Syrjänen 2006). The HPV infection is localized to squamous epithelial sites of the cervix without systemic manifestation to study.

The innate immunity is the first line defense upon encountering foreign substances, and it is a rapid response. The major components of the innate immunity are (1) physical and chemical barriers like epithelial and antimicrobial substances, (2) phagocytic cells like macrophages and neutrophils, and natural killer (NK) cells, (3) blood proteins and different cytokines. Adaptive immunity is a more specific defense and includes both the humoral and the cell-mediated immunity (CMI). The humoral immunity is mediated by B-lymphocytes that produce antibodies (Abs). Antibodies may be located on a B-cell or freely circulating in the blood, protecting the body against extracellular microbes and their toxins, to which they bind, leading to destruction of the foreign agents. CMI is mediated by T-lymphocytes, which recognize intracellular microbes such as viruses and bacteria. There are two different types of T-cells, cluster of differentiation (CD) 4 helper T-lymphocytes and CD8 cytotoxic T-lymphocytes (CTL), both of which have different functions. Helper T-cells activate different lymphocytes and stimulate inflammation, by secreting cytokines. The main functions of CTLs are to bind and kill cells that are infected by virus or bacteria (Campbell *et al* 2002).

HPV is a sexually transmitted virus that mostly remains inside the host until it can be passed to the next individual. In order to succeed, HPV has developed strategies to avoid the immune system. HPV goes through its complete lifecycle in epithelium and in that way avoids the humoral immune response of the circulatory system. As long as HPV remains in the lower epithelial layer, the expression of the early proteins E1, E2, E5, E6 and E7 is restricted. The proteins are located in the nucleus to hide them from surrounding antigen presenting cells (APC). The proteins encoded by HPV also suppress the immune system. The E7 protein has been shown to down-regulate production of interferon (IFN) -$\alpha$ and -$\beta$ (Bernard *et al* 2000). Since HPV avoids lysing cells in the lower epithelial layers, the contact between HPV proteins and APCs is restricted to phagocytosis of dying kerationocytes. (Tindle 2002)

In conclusion, the human papillomavirus is a poor immunogen, which means that its ability to activate the immune system is low. This is mainly due to different physiological characteristics as follows: (1) HPV is doublestranded DNA virus with no RNA step to induce the innate immune response, (2) during the early phase of infection, HPV produces non-secreted nucleoproteins that are not cross-presented by infected cells, (3) HPV proteins are expressed at low levels in the basal epithelium, (4) HPV infects skin cells and does not induce apoptosis, which could lead to activation of innate and adaptive immunity and (5) systemic HPV antigen presentation is limited. Because of these physiological characteristics HPV is able to avoid and make use of the host immune response. (Frazer 2007)

## 1.4 Screening

In the 1960s the gynecological screening program was introduced in Sweden. Thereafter, the detection of early or more severe cytological difference has decreased the number of cervical cancer cases. The method used to detect cytological differences is the Papanicolaou-stained (Pap) smear. The method was introduced 1949 even before George Papanicolaou knew the cause of cervical cancer. The cells from cervix are collected with a cytobrush and then spread out on a glass slide. Thereafter, the cells are stained to enable examination by microscopy. The Pap-stains are graded depending on the appearance of the cells. The first system to grade Pap-smears was called cervical intraepithelial neoplasm (CIN) system, and was introduced 1973. Cell abnormalities are graded from 1 to 3, where CIN1 is a mild and CIN2 and CIN3 represent a moderate to severe abnormality. During 1991 the Bethesda system (Tab. 1) was developed to reflect an advanced understanding of cervical neoplasia and introduce uniform

descriptive diagnostic histologic terminology. It was modified in 1991 and also in 2001, and today classifies squamous cell abnormalities into four categories: (1) ASC (atypical squamous cells), (2) LSIL (low-grade squamous intraepithelial lesions) (3) HSIL (high-grade squamous intraepithelial lesions), and (4) squamous cell carcinoma. (Burd 2003)

Altough countries with a functioning screening program report a reduced number of patients who develop cervical cancer, unfortunately some parts of the world remain without this critical program. (Schiffman *et al 2007*) In Uppsala 95% of the women do not show cytological differences in the primary screening. Among the positive 5%, 3% are graded ASCUS or CIN 1, and the remaining 2% CIN2 or CIN3 (Sparen 2007). Women in the positive group are invited for a second examination to monitor the cells, and on that occasion the patient is also tested for HPV. Before 2007, the high-risk (HR) types of HPV were detected by a commercial kit Digene Hybrid Capture 2, which detects HR HPVs as a combined measure (I. Gustavsson, personal communication). This means that the test is not telling exactly which of the HPV high-risk types the patients is infected with. After 2007 the method for analyzing HPV has been changed to a real-time PCR-based method, *hpVIR*, which identifies individual or groups of HPV types as well as viral load (I. Gustavsson, personal communication).

**Table 1.The Bethesda and CIN Classification System for cervical squamous cell dysplasia.**

| CIN system | Bethesda System 1999 | Explanation |
|---|---|---|
| | **ASC**-US (atypical squamous cells) | The abnormalities of squamous cells that are greater than the normal changes but are not seen as squamous intraepithelial lesions. |
| **CIN1** | **LSIL** (low-grade squamous intraepithelial lesions) | Mild abnormalities seen in the cells. |
| **CIN2 or CIN3** | **HSIL** (high-grade squamous intraepithelial lesions) | Moderate to severe abnormalities are seen in the squamous cells. |

## 1.5 Quantitative Real-Time Polymerase Chain Reaction

The quantitative real-time polymerase chain reaction (QRT-PCR) follows the general principle of polymerase chain reaction (PCR). The main difference is that the amplified DNA is quantified as it accumulates in the reaction after each amplification cycle. As in the general PCR, forward and reverse primers are used. The difference is the probe, which is designed to match the target sequence between the forward and reverse primer. The probe has two

fluorophores attached, a reporter dye at the 5' end and a quencher dye at the 3' end. As long as the reporter is close to the quencher, the probe emits light of quencher wavelength. During the extension step in the PCR, the reporter will be chopped of and it will no longer be in the proximity of the quencher. This will lead to change to reporter specific wavelength. (Fig. 4) The emission of the reporter light during the each extension phase is recorded and a picture of how PCR accumulates over time is generated. This is used to extrapolate the initial amount of target DNA in the samples. Amount of DNA is plotted against $C_t$, which represents the PCR cycle number at which the signal exceeded a given baseline. This gives a standard curve that makes it possible to calculate the DNA concentration in an unknown sample (Strachan & Read 1999). In a QRT-PCR up to three different probes with different fluorophores attached, can be run in the same reaction, this is called a multiplex PCR reaction.

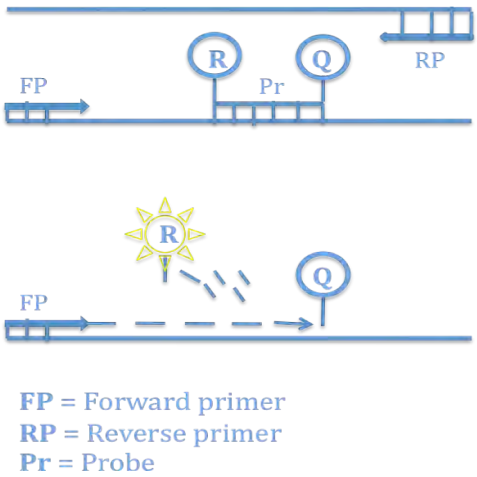



**Figure 5. Quantitative real-time PCR measures the amplified DNA as it accumulates in the reaction after each amplification cycle (modified from Strachan *et al* 2004).**

## 1.6 Aim

The aim of this project was to further develop the HPV-typing method named *hpVIR*, developed by the group of Ulf Gyllensten. The *hpVIR* assay was able to detect HPV16, 18, 31, 45, 33, 35, 39, 52 and 58. After the new development the *hpVIR* assay would also detect the additional types HPV51, 56, 59, and if possible also the low-risk types HPV6 and 11. The low-risk types are included on request from the clinicians. The new design was tested on a set of 400 clinical samples from women that previously participated in gynecological examination, and showed cytological differences. Examination of patient samples would also show the frequency of HPV51, 56, 59, 6 and 11 in the Swedish population.

# 2. Results

### 2.1 Amplification of Human papillomavirus fragment

Positive controls for the different HPV types detected by the *hpVIR* assay were generated by cloning. An old HPV plasmid (containing the entire HPV genome) or a positive sample from a patient with a known HPV infection was used as template in the polymerase chain reaction (PCR). HPV fragments were PCR amplified followed by analysis of the amplified fragments on an agarose gel. (Fig. 6) To show possible contaminations, a negative control was used. The expected size of HPV56 and HPV59 amplicons was around 65bp and 70bp respectively, and is seen in figure 6. Amplicons from HPV51, 6 and 11 were also cloned into a vector and analyzed similarly. Also here the amplification was successful, data not shown.

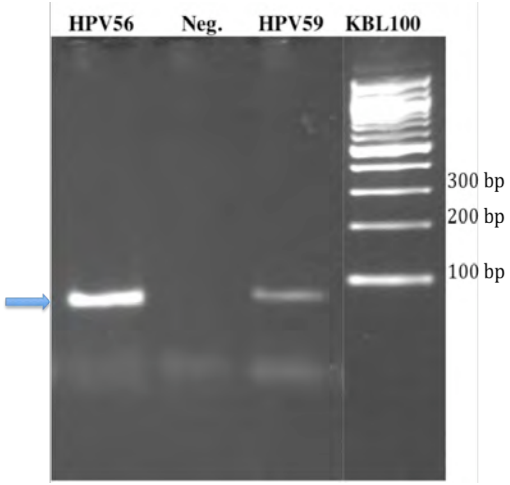



**Figure 6. PCR amplification of HPV56 and HPV59.** HPV 56 and HPV59 were amplified by PCR and amplicons analayzed by electrophoresis. The arrow is showing the expected fragment size of 65 bp and 70 bp for HPV56 and HPV59, respectively. Neg., negative control. KBL100, 100 bp ladder.

## 2.2 Cloning of the HPV amplified PCR products

PCR products were amplified from HPV51, HPV56, HPV59, HPV6 and HPV11. Thereafter each insert was cloned separately. Seven clones from each HPV were amplified and amplicons were analysed with electrophoresis. (Fig. 7)

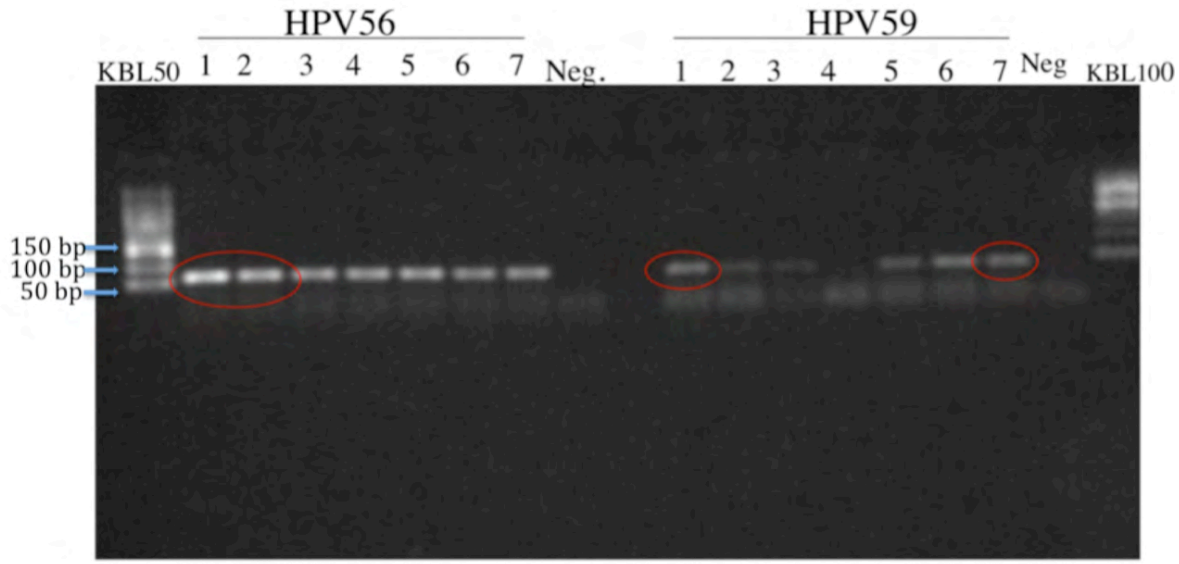



**Figure 7. PCR products from 7 colonies of HPV56 and HPV59 clones.** Clones were amplified by PCR and amplicons analyzed by electrophoresis. The red circles indicate what clones were used for plasmid purification.

All seven colonies from the HPV56 colonies carried an insert of the correct size, around 65 bp. The negative controls for both HPV56 and HPV59 showed no bands, confirming no contamination. Six of the colonies from the HPV59 cloning were positive, though the bands were weaker compared to the HPV56 bands. HPV 56 clones 1 and 2 and HPV59 clones 1 and 7 were used to prepare of plasmids.

## 2.3 Construction of standard curves in version 1 and version 2

Standard curves were generated with QRT-PCR and known concentrations of viral DNA. The standard curves were used to determine the sensitivity and specificity of the HPV assay for version 1 detecting HPV types 51, 56 and 59 (Fig. 8) and version 2 detecting HPV types 6, 11, 51, 56 and 59 (Fig. 9), using old or newly prepared HPV plasmids. The standard curve was made with diluted HPV plasmids containing $10^1$ to $10^4$ copies of the HPV genome. Standard curves were constructed for each of the HPV types in version 1, or groups of the HPV types in version 2. A linear regression between the HPV copy number and the threshold cycle (Ct) was seen for all of the HPV types.

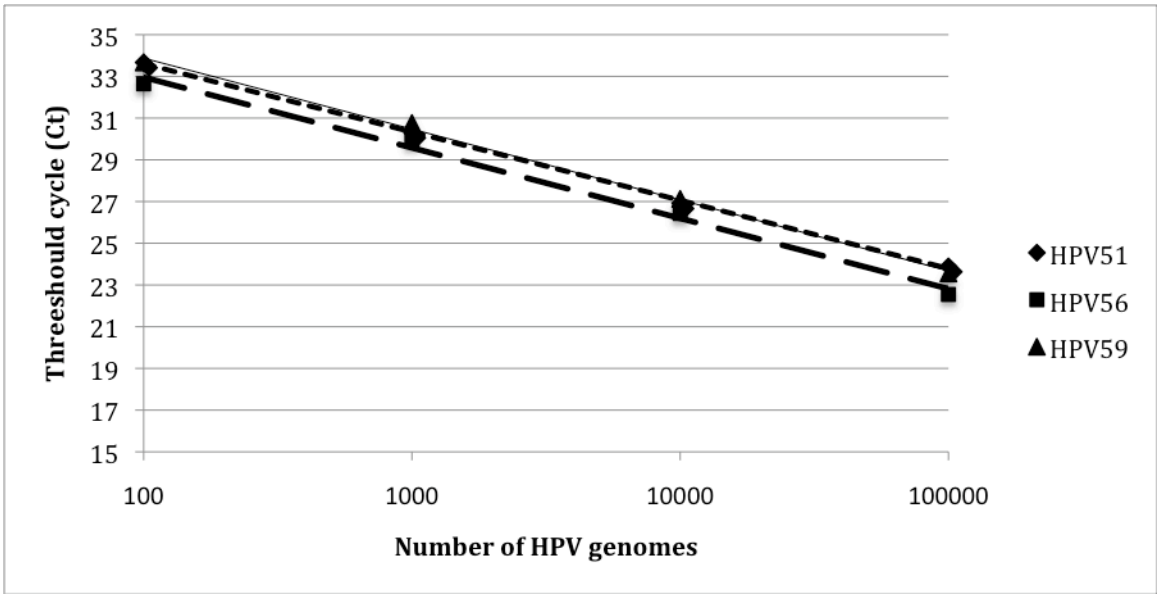



**Figure 8. Standard curves for HPV51, HPV56 and HPV59 in version 1.** The number of HPV genomes was plotted against threshold cycle (Ct) number. Each dot represents a replicate of the diluted copy number of HPV genome. $r^2$ values: HPV51= 0.98, HPV56= 0.99, HPV59= 0.99.

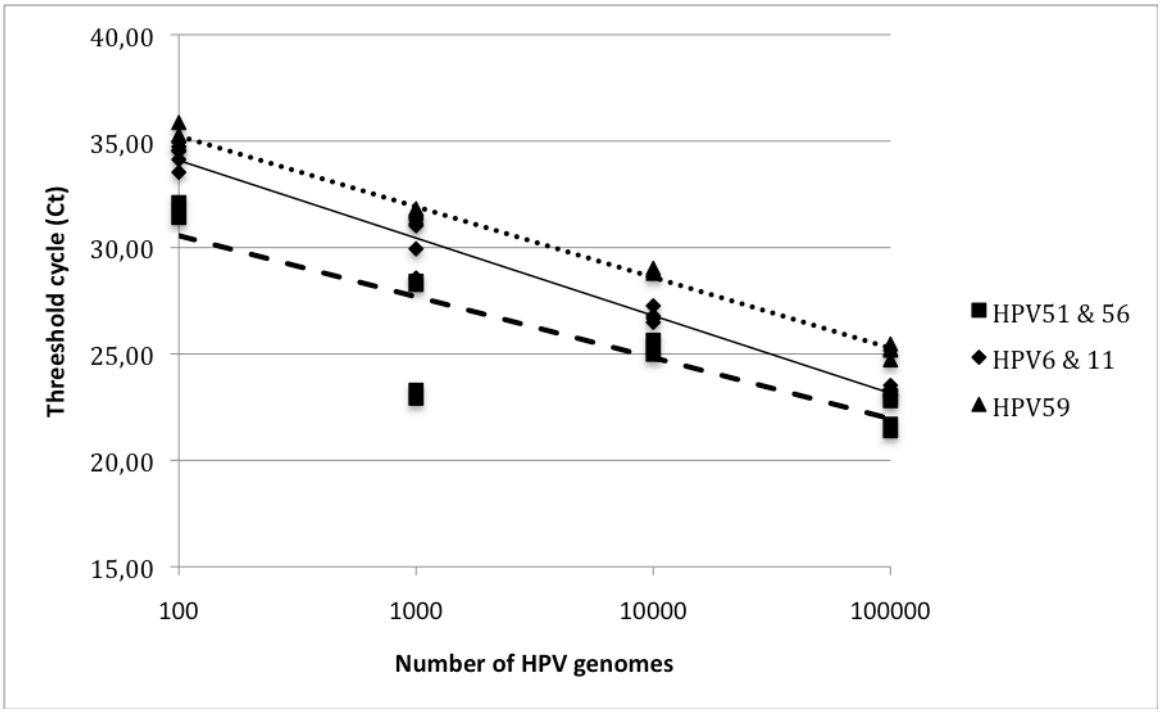



**Figure 9. Standard curves for HPV6 or 11, HPV56 or 51 and HPV59 in version 2.** The number of HPV genomes was plotted against threshold cycle (Ct) number. Each dot represents a replicate of the diluted copy number of HPV genome. $r^2$ values: HPV6 or11= 0.98, HPV56 or 51= 0.76, HPV59= 0.99.

From each standard curve, an equation was obtained that included a slope, a Y-intercept and $R^2$-value. The slope describes how much the $C_t$ value decreases for every 10 copies DNA that is added. In a 100% efficient PCR, the $C_t$ decrease with -3.3. The $r^2$- value is a measurement of how closely the data conform to a linear relationship. An $r^2$ value of 1 represents a perfect fit between the data and the line drawn through them.

To find out whether the obtained standard curves were acceptable, the slope and $r^2$-values were compared. For version 1 the slope value for HPV51 was -3.3 and -3.4 for HPV56 and HPV59. The $r^2$-value for the types detected with version 1 was 0.97 for HPV51, 0.99 for both HPV56 and HPV59. In version 2, HPV6 and HPV11 (Fig. 9a and Fig. 10) and HPV51 and HPV56 are detected together and are therefore presented in one standard curve together. The $r^2$-value for both HPV6 or 11 and HPV59 was good (0.98 and 0.99). The $r^2$-values for HPV56 or 51 was 0.76, indicating that the curve is not adequate. The reason could be an invalid dilution of the plasmid because of pipetting error when performing the PCR. The slope values for HPV6 or 11 and HPV59 were -3.6 and -3.3, while the value for HPV56 or 51 was only -2.8 indicating a less efficient PCR. (Fig.7) The points in the standard curves represent plasmid dilution series of $10^1$ to $10^4$. The different points should coincide if the dilution replicates in the series were good enough. When comparing the dots of all curves to each other inside the same standard curve, it was easy to see when they differ. The points representing 1000 copies of the HPV56 or 51genome showed most difference, followed by HPV6 or 11 also at 1000 copies.

The equations obtained from the standard curves for both versions (Tab. 2 and 3) were used to convert the $C_t$ values, obtained from patient samples, to copies of HPV genome. A sample was considered positive if it contained a HPV copy number equal to 10 or higher.

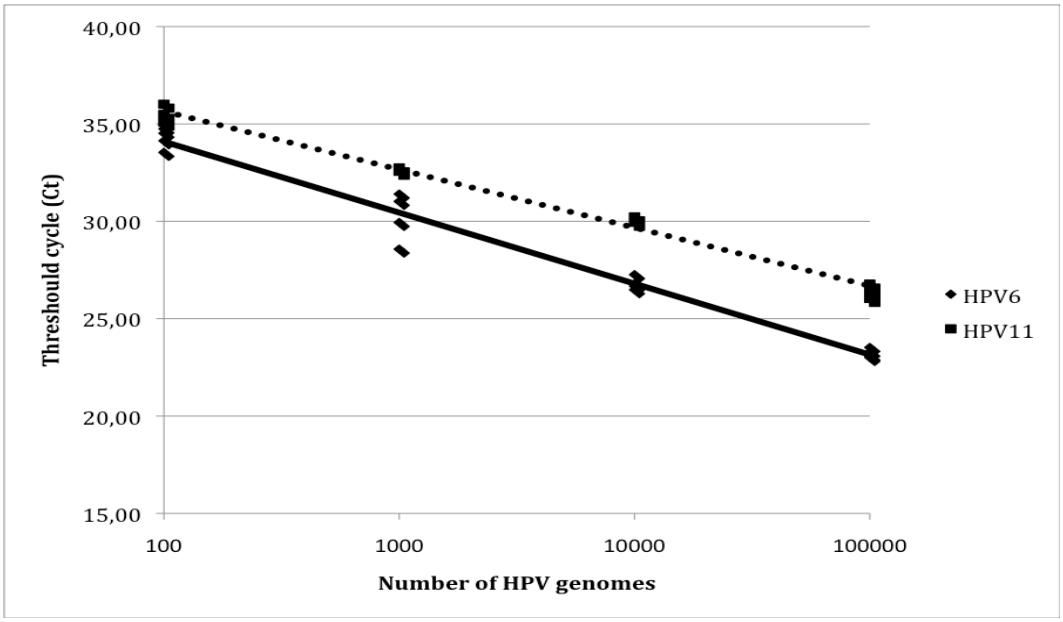



**Figure 10. HPV6 and HPV11 standard curve compared to each other.** The number of HPV genomes is plotted against threshold cycle ($C_t$) number. Each dot represents a replicate of the diluted copy number of HPV genome.

**Table 2.** Equations for the different HPV types in version 1.

| Detector | Equation | Copies | $C_t$ |
|---|---|---|---|
| HPV51 | y= -3.265 Ln(x) + 40.132 | 10 | 36.9 |
| HPV56 | y= -3.379 Ln(x) + 39.714 | 10 | 36.3 |
| HPV59 | y= -3.398 Ln(x) + 40.671 | 10 | 37.3 |

**Table 3.** Equations for the different HPV types in version 2.

| Detector | Equation | Copies | $C_t$ (x) |
|---|---|---|---|
| HPV6 or 11 | y= -3.647 Ln(x) + 41.387 | 10 | 37.7 |
| HPV51 or 56 | y= -3.121 Ln(x) + 37.334 | 10 | 34.2 |
| HPV59 | y= -3.309 Ln(x) + 40.223 | 10 | 36.9 |

## 2.4 Detection of HPV with version 1 and version 2 in 400 patients samples

Sample material from 400 patients was tested with both versions. The same patient material had been used earlier in a study when comparing the two different HPV typing methods, *hpVIR* and Digene (Moberg *et al* 2003). The samples were analyzed in the same way as for the standard curve with template DNA from the patients. The $C_t$ was then used, to calculate the DNA copy number in the samples. From the 400 patients analyzed, 42 were positive for the types detected with version 1 and 50 for the types detectable by version 2. (Fig. 11)

To see if the design of version 2 was possible to use, a comparison was made between the two versions. It was important to see if infections were missed with version 2 and in that case which infections. I could compare both single and multiple infections, as the 400 patient samples earlier had been tested with *hp*VIR detecting HPV types HPV16, 18, 31, 45, 33, 35, 39, 52 and 58. Thus, I knew precisely which patient was affected and with what viruse type. This enabled me to see if each patient was infected with on or multiple HPV types. There were 16 single infections detectable with *hp*VIR version 1 and 14 with version 2. So, in conclusion 2 patients were missed with version 2. When comparing multiple infections between version 1 and version 2, 26 multiple infections were found with version 1 and 16 with version 2. Thus, 10 multiple infections was missed with version 2.

In version 2 also the two low risk HPV types 6 and 11 were detected, though together. In total 20 patients were positive for HPV6 or HPV11 (Fig. 11), of which 4 patients had a single infection and 16 multiple infections. To find out which of these two types are most frequent among the analyzed patients, another analysis with specific primers and probes for each of these types was performed. When comparing the patients who were positive for HPV6 or

HPV11, 16 of them were positive for HPV6 while only 1 was positive for HPV11. Three of the patients were negative for both HPV types.

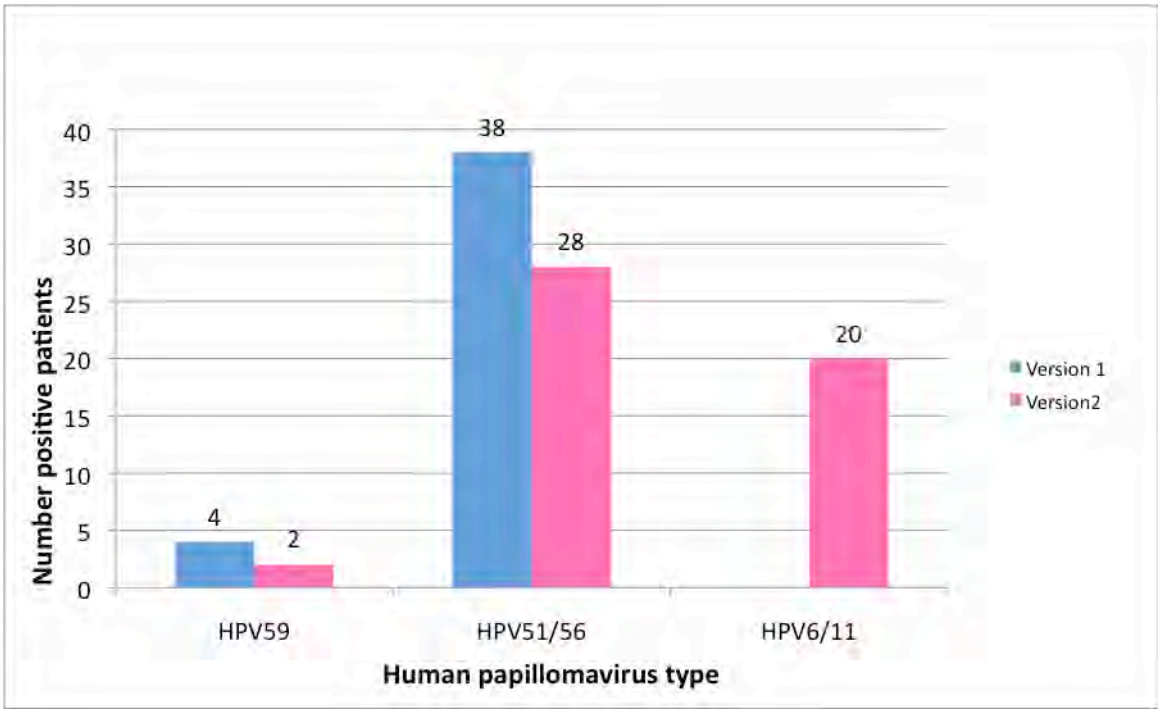



**Figure 11. Patients positive for the different HPV types.** Using the *hp*VIR version 1 and version 2 a calculation was made for number of positive patients for each HPV type.

## 2.5 Frequency of the analyzed HPV types in the patients samples

All 400 samples were analyzed and collected in Sweden. The frequency of HPV types 51, 56, 59, 6 and 11 in my study was compared to a population study from 2002 (Forslund *et al* 2002) on subjects in Sweden. The samples in my study were derived from the followup samples, while the other study measured HPV in the primary screening samples. Using the result of *hpVIR* assay version 1, I found that 1% of the tested patients carried a HPV59 infection, 6% carried a HPV56 infection and 3.5% a HPV51 infection. Results from *hpVIR* assay version 2 were used to calculate the frequency of HPV6 and 11 that were 4% and 0.25%.

**Table 4. Frequency of different HPV types among Swedish women.**

| Human papillomavirus type | Number | Frequency among my 400 patient samples | Frequency among Swedish women, age 32-38 (Forslund e*t al* 2002) |
|---|---|---|---|
| HPV59 | 4 | 1 % | 0.15 % |
| HPV56 | 24 | 6 % | 0.5 % |
| HPV51 | 14 | 3.5 % | 0.4 % |
| HPV6 | 16 | 4 % | - |
| HPV11 | 1 | 0.25 % | - |

# 3. Discussion

## 3.1 Construction and evaluation of standard curves using human papillomavirus plasmids

When constructing standard curves, a number of replicates for known copy number of HPV genomes are added to the PCR reaction. When all replicates are clustered together, a sufficient standard curve is obtained. In figure 8 standard curves of HPV types 59, 56 and 51 in version 1 are seen. Comparing their $r^2$-value it is easy to see that the standard curves are adequate. Most, but not all standard curves made with both versions for each HPV or a group of HPVs, had $r^2$-values close to 1. The $r^2$-value for the standard curve of HPV56 or 51 in version 2 was low, because of incorrect dilution of the 1000 copy replicates of the plasmid DNA. Unfortunately, there was no time left during my exam work to finish the optimization of the dilution series for HPV56 and 51.

Standard curves for single HPV types were more precise than those for HPV types typed together, since the efficiency in the PCR could differ slightly between the two types. This was the case for HPV6 or 11 where the PCR of HPV6 was more efficient than that of HPV11. (Fig. 10) To be able to detect these two types together the PCR has to be further optimized to achieve a better efficiency. Otherwise, HPV11 may be underestimated and weak infection will be missed. The expected $C_t$ values with a valid standard curve are around 36-37 (I. Gustavsson, personal communication). The $C_t$ value for 10 copies DNA detected for HPV56/51 was only 34.2. When the $C_t$ value is that low, there is a risk that positive samples from a weak infection are missed. This is because; if the standard curve were adequate, samples below 36 or 37 would be positive. With the wrong standard curve, samples above 34 will not be detected as positive, leading to that positive patients are missed.

## 3.2 Comparison of *hpVIR* result with version 1 and 2 in 400 patient samples

About 42 of the 400 patient samples were positive with version 1 compared to 30 with version 2, which is 10.5% and 7.5%, respectively, of the tested patients. To understand why version 2 failed to type 12 patients, it is important to know which patients were missed. Two groups of patients were identified, patients infected with one HPV type and patients infected with several HPV types. In my study, 10 patients with multiple HPV types and 2 patients with a single infection were missed by version 2. Altogether 16 of the patients in whom version 1 detected HPV types 51, 56 and 59 had a single HPV type infection, while 26 carried multiple infections. Version 2, detected 14 positive patients carrying a single infection with HPV51, 56 or 59, compared to 16 with a multiple infections. With version 2 HPV types 6 and 11 were also detectable, and 4 of the patients tested carried a single infection and 16 a multiple infections.

In a study performed by Forslund and co-workers, about 6100 Swedish women aged 32-38 were screened for HPV. In their study 62 women were positive for any of the HPV types 51, 56 and 59 (Forslund *et al* 2002). Their screening included more patients and the women were younger than in my study. HPV infection is more frequent at younger age, which could explain the higher frequency of these types. HPV51, 56 and 59 are normally of low frequencies in the population and a few extra positive samples have a great effect of the final result.

In conclusion *hpVIR* assay version 1 is a better HPV typing assay compared to version 2, which missed 12 women who carried a single or multiple infection. Version 2 needs more optimization to be useful in the clinical application.

## 3.3 Frequency of analyzed human papillomavirus types in the Swedish population

The frequencies of HPV56, HPV51 and HPV59 among the patients tested were 6%, 3.5% and 1% when typed with *hpVIR* (version 1) (Tab. 4). The frequency of the different HPV types typed with *hpVIR* assay was not exactly the same as the frequency found by Forslund and his co-workers (Forslund *et al* 2002) among the Swedish women. However, in both studies HPV56 was most common followed by HPV51 and HPV59. The low-risk HPV types, HPV6 and HPV11, were found in 4% and 0.25% respectively among the 400 typed patients (version 2) (Tab. 4), but no study on the Swedish population was found where these types had been analyzed.

In a global study performed by Munoz and his co-workers (Munoz *et al* 2006), about 1700 women with squamous-cell cervical cancer HPV were typed. In that study 1.2% were positive for HPV59, while 0.7% and 0.6% were positive for HPV 51 and HPV56 respectively. In their study HPV59 was more frequent than HPV51 and 56 in populations around world. On the other hand, in the Swedish population HPV56 seems to be the most frequent. Another difference is that the global study involved patients with developed squamous-cell cervical cancer. Could it be that there is a higher risk that HPV59 progresses to cancer than HPV56 or HPV51? If it was possible to follow the patients tested in the study done by Forslund and his co-workers (Forslund *et al* 2002), would we see in the future that HPV59 has become more prevalent compare to HPV56 and 51 in patients with progression to cervical cancer? The most likely explanation however is that there is a difference between the populations.

## 3.4 Future perspectives

The *hpVIR* assay that detects HPV types 16, 18, 31, 45, 33, 35, 39, and 58 is used today in gynecological examinations, in Uppsala. It was important to have an efficient assay, which will detect all positive patients. In nearest future the clinicians in Uppsala are going to add version 1 to the *hpVIR* assay for use in clinical HPV typing. In cases when the clinicians asking for the low-risk types 6 and 11 they will be tested with two separately QRT-PCR reactions. Optimization of version 2 will be continued until sufficient standard curves are obtained. In case of a successful optimization of version 2 the clinicians will be able to use that assay in the future.

# 4. Materials and methods

## 4.1 Strains

### 4.1.1 Bacteria and plasmids

Chemically competent TOP 10 *Escherichia coli* (*E. coli*) cells (Invitrogen) were used for high-efficiency cloning and plasmid propagation. Old plasmids from Dr. Martin Moberg (Tab. 5) were used as positive controls with version 1 and they contained the entire genome of HPV56 and HPV59. Plasmids used in version 2 and for HPV51 in version 1 contained short fragments, cloned from DNA extracted from patient samples with a known infection. (Tab. 5) The concentration of plasmid DNA was measured with a NanoDrop ND-100 Spectrophotometer (NanoDrop technologies) using 2 µl of each plasmid stock for analysis. The sequences of the plasmids and HPV types were obtained from the National Center for Biotechnology and Information (NCBI 2008, http://www.ncbi.nlm.nih.gov/). Using Oligo Calc: Oligonucleotide Properties Calculator, molecular weight (MW) was calculated for the plasmids and HPV types (Oligo Calc 2008, http://www.basic.northwestern.edu/biotools/oligocalc.html).

**Table 5. Plasmids containing HPV insert**.

**A.** Version 1

| HPV type | Vector |
|---|---|
| HPV51 | pCR2.1-TOPO (Invitrogen) |
| HPV56 | pT713 (Dr. Moberg) |
| HPV59 | pUC9 (Dr. Moberg) |

**B.** Version 2

| HPV type | Vector |
|---|---|
| HPV51 | pCR2.1-TOPO (Invitrogen) |
| HPV56 | pCR2.1-TOPO (Invitrogen) |
| HPV59 | pCR2.1-TOPO (Invitrogen) |
| HPV6 | pCR2.1-TOPO (Invitrogen) |
| HPV11 | pCR2.1-TOPO (Invitrogen) |

## 4.2 Patient samples

Four hundred patient samples were used, the same samples that had been used earlier used to compare (Moberg *et al* 2003) the two HPV typing methods *hpVIR* and Digene. The patient samples were collected in Sweden, from follow up samples of women who had been positive for ASCUS or CIN1 at there first screening occasion.

## 4.3 Polymerase chain reactions

### 4.3.1 Primers and probes

Oligonucleotides were designed using Primer Express 2.0 software. (Tab. 6) PCR primers that had a melting temperature around 57 °C and probes that had a melting temperature around 70 °C were preferred. To be sure that each primer and probe annealed only to the HPV type it was designed for, the sequences was analyzed in BLAST (NCBI http://www.ncbi.nlm.nih.gov/) and compared to all other HPV types. Each probe was labeled with one of the fluorophores, NED, FAM or VIC.

**Table 6. Oligonucleotides**

| Oligo-nucleotide | Specificity | Sequence 5'->3' | Fluorophore | Use |
|---|---|---|---|---|
| F51E6 | HPV51 | GCAAAAATTGGTGG ACGAAAA | - | Amplification |
| R51E6 | HPV51 | ACGTTGGACGGGGC AAT | - | Amplification |
| Pb51E6 | HPV51 | AAAGGTTCCATGAA ATAG | NED (V1)<br>FAM (V2) | Probe |
| F56E6 | HPV56 | CTCCGGAGGAAAAG CAATTG | - | Amplification |
| R56E6 | HPV56 | GATTTCATCTAATAG CACATGGTTGG | - | Amplification |
| Pb56E6 | HPV56 | ATTGTGACAGAAAA AG | FAM (V1)<br>FAM (V2) | Probe |
| F59E6 | HPV59 | TCCTACACAACGACC ATACAAACTG | - | Amplification |
| R59E6 | HPV59 | GAATATTCCTCTGCA TGATATTCGC | - | Amplification |
| Pb59E6 | HPV59 | CTGATTTGAGCACAA CAT | VIC (V1)<br>NED (V1) | Probe |
| F6E6 | HPV6 | CGGTGCTACCTGTGT CACAAA | - | Amplification |
| R6E6 | HPV6 | GGTAAAACATATACT | - | Amplification |

| | | | | |
|---|---|---|---|---|
| | | AACCAAGGCAC | | |
| Pb6E6 | HPV6 | CGCTGTGTGAAGTAG AAA | VIC (V2) | Probe |
| F11E6 | HPV11 | CTAAAGGTTGTGTGG CGAGACA | - | Amplification |
| R11E6 | HPV11 | CTAAAGGTTGTGTGG CGAGACA | - | Amplification |
| Pb11E6 | HPV11 | CCCTTTGCAGCGTGT GCC | VIC (2) | Probe |

### 4.3.2 Quantitative polymerase chain reaction

A standard curve was prepared for each HPV type individually and thereafter for the types combined in the multiplex assay. The PCR amplifications were performed in 25 µl volumes containing 5 $\mu$l template DNA (plasmid), 1.5 mM $MgCl_2$, 1 x Taqman universal mastermix (Invitrogen), 3.1 $\mu$g bovine serum albumin (BSA), 200 nM of each primers and probe (Applied Biosystems). Amplification and detection was performed using a 7900 HT Sequence Detection System (Applied Biosystems, Inc. Foster City, CA, USA). The PCR program used started with a hold step for 10 min at 95 °C, followed by a two step-cycle of 15 s at 95 °C and 60 s at 57 °C, repeated 40 times. The real-time PCR data was analyzed using the software SDS 2.0 (Applied Biosystems). The SDS program measured the threshold cycle number value ($C_t$) for each HPV type detected. The equation obtained from the standard curve was used to calculate the amount of HPV copies in sample.

### 4.3.3 Polymerase chain reaction for cloning

A specific PCR was first set up for the HPV infection found in the patient. The amplification was performed in 50 µl volume containing 3 $\mu$l template DNA, 1x PCR-buffer (Invitrogen), 1.5 mM $MgCl_2$, 1 U Taq platinum polymerase (Invitrogen), 3.1 $\mu$g BSA, 1µM of each primer (Tab. 7) (Applied Biosystems), 500 µM deoxyribonucleoside triphosphates (dNTP). Cycling conditions were 95°C 10 min, 40 x (95 °C 30s, 57 °C 1 min) and 72 °C 10 min.

### 4.3.4 Gel electrophoresis

PCR products were run on a 2% agarose gel (2 g agarose, 100 ml 0.5 x TEB), at 120 V, in 0.5 x TEB (1x TEB is 89 mM Tris (pH 8.3), 0.25 mM EDTA, 99 mM boric acid) containing 0.625 µg/µL ethidium bromide, together with a 50 and 100 bp molecular size ladder (Invitrogen).

## 4.4 Cloning

The PCR product was cloned into TOP10 cells (Invitrogen) using the TOPO TA Cloning kit (Invitrogen). To the PCR product, salt solution and TOPO vector were added; the mixture was then incubated in 5 min at room temperature (RT). The TOPO cloning reaction was added to a tube of One Shot chemically competent *E. coli* TOP 10 cells (Invitrogen). The tube was gently moved up and down and then incubated on ice for 5 min. The cells were heat shocked for 30 s at 42 °C and then kept on ice. S.O.C medium (Invitrogen) was added to the cells and the tubes were incubated on a shaker for 1 hour (h) at 37 °C. After incubation, 40 µl of the cells, now containing the vector with the inserted HPV fragment was spread on an agar (LB (1.0% Tryptone, 0.5% yeast Extract, 1%NaCl pH 7.0) and 15 g/l agar) plate containing 50 µg/ml ampicillin and incubated over night (ON).

Seven colonies from each plate were inoculated into 1 ml LB medium (1.0% Tryptone, 0.5% yeast Extract, 1%NaCl pH 7.0), which contained 50 µg/ml ampicillin in a 15-ml falcon tube, until a visible bacterial growth was seen. From the bacterial cultures 10 µl was used in a control PCR, to see which colonies contained the wanted fragment. To the rest of the cultures new LB medium was added and the tubes were incubated on a shaker at 37 °C, ON. Two of the colonies with a band of correct size were used further.

### 4.4.4 Plasmid Mini Prep kit

To purificate plasmid DNA, a PureLink™ Quick Plasmid Miniprep kit (Invitrogen) was used as described by the manufacturer.

# 5. Acknowledgment

I would like to thank Ulf Gyllensten for giving me an opportunity to do my master thesis within such an interesting field. Special thanks to Inger Gustavsson for being so helpful, understanding and always having time to answer my questions, but most of all thanks for being such a nice and open person. Thanks to all people I have meet on Rudbeck, for making it a very pleasant place to work at. Also a great thank to Enzo Shifo for reading my thesis. Finally, thanks to my wonderful family *"Zahvaljujuci vama, sam to sto jesam danas"*.

# 6. References


Barnard P., Payne E., McMillian N. A. 2000. The Human papillomavirus E7 protein is able to inhibit the antiviral and anti-growth functions of nterferon-alpha. *Viriology* 277:411-9.

Burd M. E. 2003. Human papillomavirus and cervical cancer. *Clinical Microbiology Reviews* 16:1

Burghardt E., Ostor A. G. 1983. Site and orgin of squamous cervical cancer: a histomorphologic study. *Obstet* 62: 117-27

Campbell A. N., Reece B. J. 2002. Biology. 3rd edition. Pearson Education Inc. San Francisco.

Cutts F. T., Franceschi S., Goldie S., Castellsague X., de Sanjose S., Garnett G., Edmunds W. J., Claeys P., Goldenthal K. L., Harper D. M., Markowitz L. 2007. Human papillomavirus and HPV vaccines: a review. *Bulletin of the World Health Organization* 85:719-726

de Villiers E. M., Fauquet C., Broker R. T., Bernard H. U., zur Hausen H. 2004. Classification of papillomaviruses. *Virology* 324:17-25.

Forslund O., Antosson A., Edlund K., van der Brule J.C. A., Hansson B.G., Meijer J.L.M C., Ryd W., Rylander E., Strand A., Wadell G., Dillner J., Johansson B. 2002. Population-Based Type-Specific Prevalence of High-Risk Human Pappillomavirus Infection in Middle-Aged Swedish Women. *Journal of Medical Virology* 66:535-541.

Frazer I. 2007. Correlating immunity with protection for HPV infections. *International Journal of Infectious Diseases* 11:S10-S16.

Hazard K. 2007. Humant papillomavirus och kampen mot cancer. *Laboratoriet* 2007:4.

Munoz N., Bosch X., de Sanjose S., Herrero R., Castellsague X., ShahV. K., Snijders J.F. P., Meijer J.L.M C. 2003. Epidemiologic Classification of Human Papillomavirus Types Associated with Cervical Cancer. *New England Journal of Medicine* 348:518-27.

Munoz N., Castellsague X., Berrington de Gonzalez A., Gissmann L. 2006. Chapter 1: HPV in the etiology of the human cancer. *Vaccine* 243:1-10.

Moberg M., Gustavsson I., Gyllensten U. 2003. Real-Time PCR-Based System for Simultaneous Quantification of Human Papillomavirus Types Associated with High Risk of Cervical Cancer. *Journal of Clinical Microbiology* 41:3221-8

National Center for Biotechnology and Information (NCBI). 2008. BLAST. WWW-document: http://blast.ncbi.nlm.nih.gov/Blast.cgi Used: 2008-04-17

Oligo Calc: Oligonucleotide Properities Calculator. 2008. WWW-document: http://biotech.about.com/od/technicaltheory/tp/Ecoli.htm Used: 2008-04-17

Paavonen J. 2007. Human papillomavirus infection and the development of cervical cancer and related genetial neoplasia. *International Journal of Infectious Diseases* 11:S3-9.

Parkin M. D., Bray F., Ferlay J., Pisani P. 2005. A cancer journal for clinicians. *CA Cancer J Clin* 55:74-108.

Phillips T. 2008. Top 5 Reasons E. coli is used for Gene Cloning. WWW-document: http://biotech.about.com/od/technicaltheory/tp/Ecoli.htm Download 2008-06-19

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

Schiffman M. H., Brinton L. A. The epidemiology of cervical carcinogenesis. 1995. *Cancer 15;76(10 Suppl):1888-901*.

Staubenrauch F., Laimins A. L. 1999. Human papillomavirus life cycle: active and latent phases. *Cancer biology* 9:379-386.

Steben M., Duarte-Franco E. 2007. Human papillomavirus infection: Epidemiology and pathophysiology. *Gynecological Oncology* 107:S2-S5.

Strachan T. and Read P. A. 2004. Human molecular genetics. 3rd edition. Garland Science. New York.

Sparen Pär. 2007. Gynekologisk cellprovskontroll i Sverige. WWW- document: http://ki.se/ki/jsp/polopoly.jsp?d=20173&a=50503&l=sv Download 2008-07-11

Syrjänen S. 2006. Humana papillomavirus och orala infektioner. *Tandläkartidningen* 98:40-50

Tindle R. W. 2002. Immune evasion in human papillomavirus-associated cervical cancer. *Nature Review Cancer* 2:59-65.

Zur Hausen H. 2002. Papillomaviruses and cancer: from basic studies to clinical application. *Nature reviews 2:342-350*.