UNIVERSIDAD POPULAR AUTÓNOMA DEL ESTADO DE PUEBLA

Department of Biological Sciences
School of Biotechnology Engineering

ESTIMATE DETECTION LIMITS FOR DIFFERENT HPV TYPES USING THE MANOS ET AL., 1990 DETECTION SYSTEM

THESIS

FOR THE DEGREE OF

BIOTECHNOLOGY ENGINEER

Presents:

Salvador Aguilar Rosas

Director:

Javier Garcés Eisele, PhD.

Puebla, Mexico **June 2016**

UPAEP – Secretaría General
Dirección General de Apoyos Académicos
Dirección del Centro de Recursos para el Aprendizaje y la Investigación.
Biblioteca Central - **Karol Wojtyla**

Tesis Digitales Restricciones de uso:

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de textos, imágenes, gráficas, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente de donde la obtuvo mencionando el autor o autores involucrados en el documento.

Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Thesis Acceptance Letter

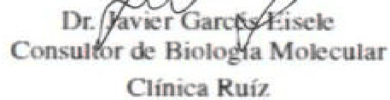
Heróica Puebla de Zaragoza a 7 de junio de 2016

Dra. Beatriz Pérez Armendáriz
Coordinadora de la Facultad Biotecnoambiental
PRESENTE:

Por este conducto me permito informar a usted que la tesis de Licenciatura en Ingeniería en Biotecnología del C. **Salvador Aguilar Rosas** con la matrícula **74600102** y ID **96293**, titulada "Estimate detection limits for different HPV types using the Manos et al., 1990 detection system" bajo mi dirección está terminada y cumple con los requisitos necesarios para obtener el título de Ingeniero en Biotecnología. Por lo anterior, espero que pueda presentar su examen profesional de licenciatura a la brevedad posible.

Agradezco su atención y quedo a sus órdenes para cualquier duda o aclaración.

Atentamente



Dr. Javier Garcés Eisele
Consultor de Biología Molecular
Clínica Ruíz

UPAEP
21 Sur 1103
Barrio de Santiago
Puebla, Pue. México
C.P. 72410

Tel: 01 (222) 229 9400
Fax: 01 (222) 232 5251
01 800 224 2200
www.upaep.mx

DEDICATION

To my beloved mother,

who has been an inspiration,

a role model who embody

hard work and success

and for making me

who I am today.

ACKNOWLEDGMENTS

I would like to gratefully thank my family and friends, who have supported me not just in this thesis but in all my endeavours. There are no words to describe how thankful I am, for all the moments we have spent together, I owe you all the best memories I have.

It has been a pleasure to have met my thesis advisor and teacher Dr. Javier Garcés Eisele, whom I admire and respect. There is no doubt that he is an extraordinary person and I want to express my deep sense of gratitude for his time and lessons in and outside the university.

It was a delight to do my research in the Clínica Ruiz laboratories, where I found many intelligent and wonderful people. I would like to thank them for the supplies and facilities for the realization of this thesis.

Finally, to UPAEP my university, a place where I was given the opportunity to study and to the professors that make a difference in the students life, I am deeply thankful especially to: Dr. Elie Girgis El Kassis and Dr. Patricia Eréndira Muratalla Guzmán.

ABSTRACT

Human papillomaviruses (HPVs) infect basal epithelial, such as skin and/or mucosal cells; it is transmitted primarily by skin-to-skin or mucosa-to-mucosa contact and due to insufficient data there is no probability reported of infection per sexual act. The cervical HPV infection is frequent among women with active sexual life under the age of 30 years old with the highest prevalence under 25 years old. A persistent infection may cause cervical cancer, specially when women are infected with high-risk genotypes. However, the analytical sensibility of the MY09/11 system, which is widely used because it amplifies a broad spectrum of HPVs, has not yet been determined. To explore this issue, biological samples were gathered and isolated by HPV type-specific primers. Subsequent to this, the HPV fragments were cloned and then amplified on bacterial cells. Dilution series were used to determine the sensitivity limit of the MY09/11 system for different HPVs, through a logit model. Taken together, the sensibilities for the HPV 31, 58, 62 and 66 were elucidated, 162, 117, 423 and 123 copies, respectively. Knowing the limitations of the MY09/11 system and its sensitivities of each HPV may help as an auxiliary parameter to corroborate a negative result.

RESUMEN

Los virus del papiloma humano (VPH) infectan epitelios basales, tales como la piel y/o células de la mucosa; se transmite principalmente por contacto de piel a piel o contacto de mucosa a mucosa y debido a la escasez de datos no hay ninguna probabilidad reportada de infección por acto sexual. La infección cervical del VPH es frecuente entre las mujeres con vida sexual activa con menos de 30 años de edad con una prevalencia más alta en mujeres con menos de 25 años. Una infección persistente puede causar cáncer cervicouterino, especialmente cuando las mujeres están infectadas con genotipos de alto riesgo. Sin embargo, la sensibilidad analítica del sistema MY09/11, que se utiliza ampliamente debido a que amplifica un amplio espectro de virus del papiloma humano, aún no ha sido determinada. Para explorar esta cuestión, se recolectaron muestras biológicas y se aislaron mediante cebadores específicos para cada tipo de VPH. Después de esto, los fragmentos de VPH se clonaron y amplificaron en células bacterianas. Se utilizaron diluciones seriadas para determinar el límite de sensibilidad del sistema de MY09/11 para diferentes virus del papiloma humano, a través de un modelo logit. En conjunto, las sensibilidades para el VPH 31, 58, 62 y 66 fueron encontradas, 162, 117, 423 y 123 copias, respectivamente. Conociendo las limitaciones del sistema MY09/11 y sus sensibilidades para cada VPH, la sensibilidad puede ayudar como un parámetro auxiliar para corroborar un resultado negativo.

CONTENTS

Thesis Acceptance Letter **ii**

Dedication **iii**

Acknowledgments **iv**

Abstract **iv**

Resumen **v**

List of Tables **ix**

List of Figures **xi**

List of Abbreviations and Symbols **xii**

1 Introduction **1**
- 1.1 Statement of the Problem . . . 2
- 1.2 Objectives . . . 4
  - 1.2.1 General Objective . . . 4
  - 1.2.2 Specific Objectives . . . 4
- 1.3 Hypothesis . . . 4
- 1.4 Justification . . . 4
- 1.5 Scope and Limitations . . . 6
  - 1.5.1 Scope . . . 6
  - 1.5.2 Limitations . . . 6

2 Literature Review **7**
- 2.1 Papillomavirus . . . 7
- 2.2 HPV . . . 8
  - 2.2.1 Transmission . . . 9
  - 2.2.2 Infection . . . 9
- 2.3 Genome . . . 9
  - 2.3.1 Oncogenic Activity . . . 10
- 2.4 Mechanism of infection . . . 11
  - 2.4.1 Replication . . . 12
  - 2.4.2 Function of E6 and E7 proteins . . . 13

2.5 Tropism 14
2.6 Cancer 16
2.7 Epidemiology 16
2.7.1 HPV-16 AA 19
2.8 Detection of Cervical Cancer 20
2.8.1 Papanicolaou 20
2.8.2 End-Point PCR 22
2.8.3 Hybrid Capture II 22
2.8.4 MY09/11 23
2.8.5 PGMY09/11 24
2.8.6 GP5+/GP6+ 24
2.9 Vaccine 25
2.10 Bioinformatics 26
2.10.1 C 26
2.10.2 Python 26
2.11 R Programming 27

3 Materials and Methods 28
3.1 Sequences 28
3.2 Primer Design 28
3.3 Samples 29
3.4 PCR amplification 29
3.5 Gel Electrophoresis 30
3.6 Fresh competent Bacteria 30
3.7 DNA Cloning 30
3.8 Plasmid Extraction 30
3.9 DNA Purification 30
3.10 DNA quantification 31
3.11 DNA copy number 31
3.12 Sensitivity 31
3.13 Logistic Regression 31

4 Results 32
4.1 Primer design 32
4.2 Type-specific Amplifications 34
4.3 Competent Bacteria and DNA Cloning 34
4.4 Cultivation and Plasmid Extraction 35
4.5 DNA Purification 36
4.6 DNA Quantification 37
4.7 DNA copy number and Sensitivity 38

5 Discussion 41
5.1 HPV samples 41
5.2 Type-specific primer design 41
5.3 Sensitivity and Detection limit 42

6 Conclusion **51**

REFERENCES **53**

Appendix A Tables **60**

LIST OF TABLES

2.1 Sensitivity of MY09/11 in PCR assays. 24
2.2 Characteristics of two HPV vaccines 25

3.1 Criteria for the HPV type-specific primers. 28

4.1 Calibration data from the standards 38
4.2 HPV DNA concentration 38
4.3 HPV copy number 39
4.4 PCRs of the intermediate dilutions 40
4.5 HPV sensitivity limit 40

5.1 Comparison of HPV sensitivities 42
5.2 Mismatches on HPV 31 and 35 46
5.3 Mismatches on HPV 35, 52, 53 and 56 46
5.4 Mismatches on HPV 51 and 68 46
5.5 Unique HPV sequences 49

A.1 Primer MY09 60
A.2 Primer MY11 62
A.3 HPV sequences and IDs 64
A.4 HPV forward and reverse primers 65
A.5 Primer MY09 vs HPV sequences 67
A.6 Primer MY11 vs HPV sequences 69
A.7 Unique MY09 sequences 71
A.8 Unique MY11 sequences 72
A.9 Variants of HPVs and their identifiers 73

LIST OF FIGURES

1.1 Conservation of the MY09 region . . . 3
1.2 Conservation of the MY11 region . . . 3

2.1 Electron micrograph of the HPV . . . 8
2.2 Human Papillomavirus 16 genome . . . 10
2.3 Binding of the HPV virion . . . 11
2.4 HPV virion within the cell . . . 12
2.5 Replication cycle of genital high-risk HPV . . . 13
2.6 Binding of E6 and E7 oncoproteins . . . 14
2.7 Tropism of HPV . . . 15
2.8 Cervical cancer incidence . . . 17
2.9 Cervical cancer mortality . . . 17
2.10 Incidence and mortality of cervical cancer by country . . . 18
2.11 Pap smear with koilocytic atypia . . . 21
2.12 ThinPrep pap smear . . . 21

3.1 Algorithm to find the HPV type-specific primers . . . 29

4.1 Primer-BLAST design . . . 32
4.2 Primer-BLAST generated primers . . . 33
4.3 Type-specific PCR amplification . . . 34
4.4 HPV Minilysates . . . 35
4.5 DNA’s large scale culture gel . . . 36
4.6 Digestions of purified plasmid DNAs . . . 37
4.7 Calibration curve of DNA quantification . . . 37

4.8 Initial dilution series of HPV 66 39
4.9 Regression curves of the HPV 66 40

5.1 Mismatches and sensitivity with the MY09 primer 43
5.2 Mismatches and sensitivity with the MY11 primer 44
5.3 Mismatches and sensitivity with the MY09/11 primers 44
5.4 Linear model of mismatches and sensitivity with the MY09/11 primers 45
5.5 Weighted logarithmic model of MY09 47
5.6 Weighted logarithmic model of MY11 47

LIST OF ABBREVIATIONS AND SYMBOLS

BLAST	Basic Local Alignment Search Tool
CIN	Cervical Intraepithelial Neoplasia
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleoside Triphosphate
EMBOSS	European Molecular Biology Open Software Suite
FNR	False Negative Rate
HPLC	High Performance Liquid Chromatography
HPV	Human Papilloma Virus
HSIL	High-grade Squamous Intraepithelial Lesions
HSPGs	Heparin Sulfate Proteoglycans
LB	Lysogeny Broth or Luria-Bertani
LCP	Laboratorios Clínicos de Puebla
LSIL	Low-grade Squamous Intraepithelial Lesions
MY09/11	Set of degenerated primers to detect the HPV
NCBI	National Center for Biotechnology Information
NGS	Next-Generation Sequencing
OD	Optical Density
PCR	Polymerase Chain Reaction
PV	Papillomavirus
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
SDS	Sodium Dodecyl Sulfate
SIL	Squamous Intraepithelial Lesions
T_m	Melting Temperature
TBE	Tris/Borate/EDTA
TE	Tris-EDTA
URR	Upstream Regulatory Region
VaIN	Vaginal Intraepithelial Neoplasia
VIN	Vulvar Intraepithelial Neoplasia
VLP	Virus-like Particle

CHAPTER 1

INTRODUCTION

The human papillomaviruses (HPVs) are double stranded DNA viruses that mainly infect basal epithelial, such as skin and/or mucosal cells. There exist three main classes of tropism: cutaneous, mucosal and mixed tropism (Eide & Debaque, 2012), with over 50 genotypes of the 220 currently classified. These viruses belong to the Papillomaviridae family and are characterized in “high-risk” i.e. 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66 and “low-risk” i.e. 6, 11, 40, 42, 43, 44, 53, 54, 61, 72, 73, 81 genotypes. Among the two categories the high-risk genotypes are associated with cervical cancer, they may however, generate cancers in other anogenital tissues. They are furthermore discussed as risk factor for cancers in other mucosal tissues (oral, bladder, etc.). Yet low-risk genotypes mainly cause benign cervical changes and genital warts, commonly considered a cosmetic problem that can be treated until warts disappear and do not come back again.

The cervical HPV infection is frequent among women with active sexual life under the age of 30 years old with the highest prevalence under 25 years old. Middle-aged and young women usually have transient HPV infections that can be resolved by the immune system in about 6 to 12 months. However, in presence of a persistent infection, within a period of 15 to 20 years, there is a high probability to develop a cervical neoplasia (Woodman, Collins, & Young, 2007).

As HPV infections are almost asymptomatic and occur frequently at the transition from cylindrical to squamous epithelium of the cervix, women have to go to the doctor to be examined. When infected mild changes can be seen in the epithelium and can be identified though virological or cytological techniques. Squamous intraepithelial lesions (SIL) can be screened through cytological examination of cervical smears, there are low or high SIL relying upon how

affected the epithelium is and how abnormal the cells appear. Nonetheless, when there is a confirmation of abnormal cells in the cervix detected by histological examination of cervical biopsies the term cervical intraepithelial neoplasia (CIN) is used, and is graded from 1 to 3; describing the morphology of the affected epithelial cells as well as the thickness of the cervical epithelium with abnormal cells. Similar lesions exists for vaginal (VaIN) and vulvar (VIN) tissues (Cutts et al., 2007).

As previously discussed, certain HPVs genomes can lead to cervical cancer, which is an important public health issue; it is the most frequent and mortal neoplastic disease worldwide in the female population. The Latin American female population is on the high-risk list to develop cervical cancer with up to 68,000 new cases every year. The highest mortality rates belong to Chile and Mexico, whereas the lowest to Cuba, Puerto Rico and Argentina. However, in Mexico the cervical cancer is the first cause of cancer related death in women over 25 years old. Although the Mexican National Health System provides medical assistance to roughly 9,000 cases of cervical cancer, annually there are about 4,000 deaths due to a late diagnosis. Hence, it is of utmost importance to prevent this type of cancer through education, vaccines and prevention campaigns (Hidalgo-Martínez, 2006).

1.1 STATEMENT OF THE PROBLEM

The primer system MY09/11 is widely used worldwide, because of its great coverage of different HPV types and its extensive characterization. However, the method was designed originally based on the L1 sequences of the only mucosal HPV sequences known at that time: 6, 11, 16, 18 and 33. Therefore, the amplified region in other HPVs may not be as efficient due to the variability of the viral sequences (see Figures 1.1 and 1.2). It is apparent that the variability is taken into account only partially by the primer design. Nonetheless, due to the high degree of conservation of these regions, amplification of a large number of HPVs has been reported in the literature.

Surprisingly, despite its wide use, the information of the method's sensitivity is scarce. Ergo, it is not even clear which HPV types are detectable by the degenerated primer system MY09/11. Moreover, this could mean that the type specific prevalences that have been reported may be

biased by the sensitivity of MY09/11.

The conservation of the MY09/11 regions can be measured by a variability analyses, using the Shannon entropy (Ruiz, Pérez, & Bonev, 2009). See Figure 1.1 and Figure 1.2.

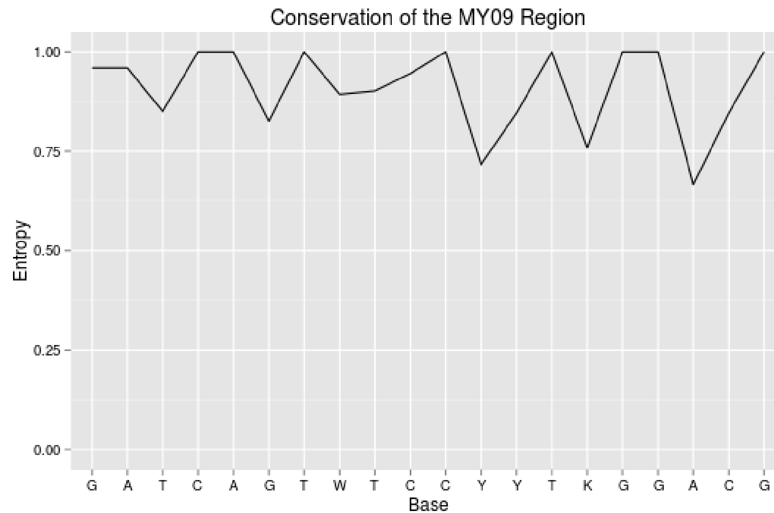


Figure 1.1: Conservation of the MY09 region, see Table A.1. The conservation of the MY09 region has been determined using the Shannon entropy of the sequence alignment. The sequence indicated corresponds to the complement of the MY09 primer sequence.

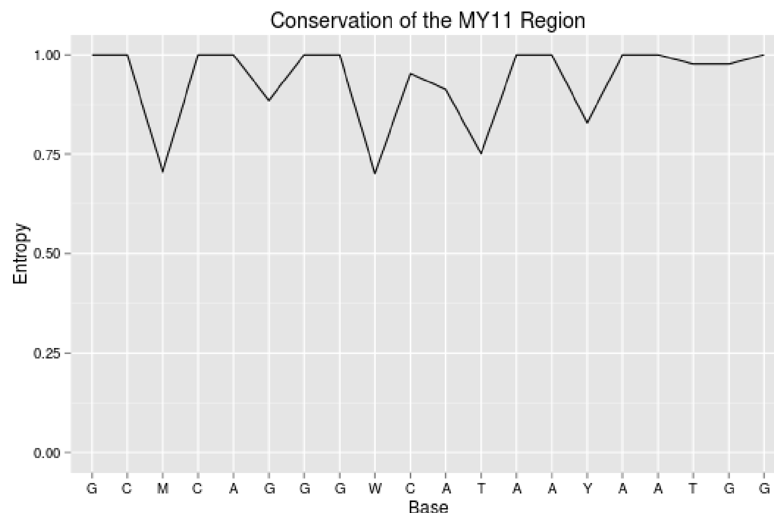


Figure 1.2: Conservation of the MY11 region, see Table A.2. The conservation of the MY11 region has been determined using the Shannon entropy of the sequence alignment. The sequence indicated corresponds to the complement of the MY11 primer sequence.

1.2 OBJECTIVES

1.2.1 General Objective

- Determinate the analytical sensibility for the different HPV types, using the MY09/11 primers.

1.2.2 Specific Objectives

- Obtain biologic samples from the Laboratorios Clínicos de Puebla (LCP)
- Using reference sequences for the mucosostropic HPV types obtained from National Center for Biotechnology Information (NCBI) design primers encompassing the MY09/11 region through Primer-BLAST
- Use these primers in order to clone different HPV types that are in the Clínica Ruiz gene library.
- Rigorously quantify the plasmids in order to obtain well defined reference DNAs.
- Prepare dilution series of the reference DNAs down to 1 copy/μl or less and amplify them repeatedly with the MY09/11 system.
- Determinate with the help of the logit model the sensibility limit of the MY09/11 primers.

1.3 HYPOTHESIS

It is possible to determine the type specific limits of detection for the MY09/11 system.

1.4 JUSTIFICATION

To achieve a more accurate a analysis, when detecting and typing the HPV, it is important to have information about the sensitivity. Because this information is not known for each type of the HPV when using the MY09/11 system. It arises an important problem when a doctor has to interpret a negative result. Usually a negative result is interpreted as if the patient has

normal cervical cells or no HPV infection. However, as there are differences of sensitivity for different HPVs when using the MY09/11 system a negative result should be interpreted taking into account these data. Therefore, when reporting a negative result the doctor and the patient have to understand the limitations of it and put this result into context with other laboratory or clinical tests.

For example, with regards to a report of a normal cytology and a negative result from a molecular detection system caution should be taken. Although, the worldwide prevalence of the HPVs 16 and/or 18 in women with normal cytology is only 3.9%, these two viruses are responsible of over 70% of the cervical cancers. Therefore, even though a negative PCR result is reported, it should be kept in mind that the virus just may be not detectable due to a viral concentration in the sample below the detection limits (Bruni et al., 2015).

Nowadays it is relatively easy around the world to detect the HPV efficiently and systematically but one problem arise from the primers that are used: the Manos et al., 1990 degenerated primers. The primers are capable in theory to detect a vast number of the different HPV variants and types, many laboratories around the globe have been using this detection system since the 90s with good results so far, but what if we could improve the results that are given to the doctor by measuring the sensibility to each HPV variant, because up to now there is no information in regards to the sensibility of the Manos et al., 1990 detection system.

This work would help to establish the initial parameters to determine the true sensitivity of the degenerated primers to each of the different variants of the HPV. It will put into perspective the results that have been obtained along the last two decades. Furthermore, this work could be used as a complementary tool to evaluate how strong the reported prevalences of the mucosotropic HPVs could be biased. It is worth mentioning that the reported prevalences thus far have been used to assess the possible impact of vaccines against different sets of viruses. As far as the MY09/11 system has been used to obtain these data, defining the sensitivity of the method will offer a more accurate interpretation of the results and allow an improved knowledge based evaluation of vaccines on different populations. Finally, the MY09/11 system may be improved enhancing the detection of as yet difficult-to-detect types by modifying the primer set.

1.5 SCOPE AND LIMITATIONS

1.5.1 Scope

1. Detection of the sensitivity limits of MY09/11 for selected mucosotropic HPV types

1.5.2 Limitations

1. Significant variability among the HPV sequences retrieved from NCBI and the HPVs on the biological samples

CHAPTER 2

LITERATURE REVIEW

2.1 PAPILLOMAVIRUS

Papillomaviridae is the family of a whole set of viruses, who infect the epithelia of vertebrates, which are called papillomavirus (PV) (Fauquet, Mayo, Maniloff, Desselberger, & Ball, 2005). In some cases, the infection can lead to neoplastic disorders, such as, cervical cancer, whereas in other cases the virus in question can persist asymptomatically in the host "for an indeterminate period of time". Although PV are highly diverse, there exist more than 189 types (Bernard et al., 2010), the most common organisms that are infected are mammals and birds (De Villiers, Fauquet, Broker, Bernard, & zur Hausen, 2004). However, human beings are the priority in terms of research, because before 1970 papillomaviruses were thought to only cause warts, which back then was just a cosmetic problem; affecting mammals with cutaneous lesions. In other words, this was not a major threat to public health; it was until 1980s when using new protocols it was found out that some HPVs could induce neoplasms (Bernard, 2005).

Not so long ago the PVs had been considered members of the Papovaviridae, nowadays they are, however, separated into the Papillomaviridae. Originally, it was thought that the similarities between them was enough to group them in a single family, characterized by features such as circular double-stranded DNA genomes. But now this idea has been refuted, as the viruses have different genome sizes, genome organization and gene functions. PVs have no envelope and a diameter ranging from 50 to 55 nm, their structure is constituted by 72 capsomers in an icosaedric capsid (Longo et al., 2011).

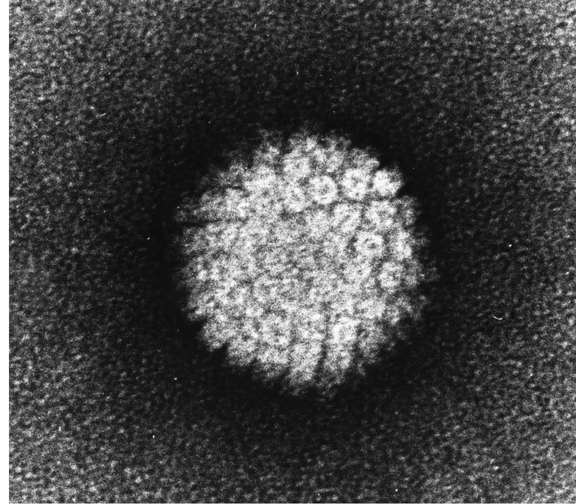


Figure 2.1: Electron micrograph of a negatively stained human papilloma virus (Laboratory of Tumor Virus Biology, 1986).

2.2 HUMAN PAPILLOMA VIRUS (HPV)

It is well known that there are many genera among the HPVs, however, the Alpha and Beta papillomaviruses are the most important as around 90% of currently characterized HPVs are found in these two genera (Doorbar, 2006).

Almost 4 decades ago, biochemical and immunological analysis were used as a primary tool to study biological processes, yet the characterization of viruses in skin warts, anal and genital condylomata and oral and laryngeal papillomas could not be determined by the lack of a cell system, where it could be possible to replicate the virus *in vitro*, this problem is due to the extensively variable viral content according to their clinical type (Orth, Favre, & Croissant, 1977).

Nowadays, with sequencing technologies and the improvement of computer algorithms in the last decade, phylogenetic trees have been generated based on DNA sequence and protein homologies. These data or information has been generated to better understand the classification and the pathogenic mechanism of the HPVs. However, it has been discovered that there are two main types that are responsible for 70% of cervical cancer, HPV16 and HPV18, these types are considered high risk HPVs. Among them other types such as, HPV31, and HPV45 are found (Longo et al., 2011), whereas the low risk HPVs are the HPV6 and HPV11 and are the ones

responsible for up to 90% of genital warts.

2.2.1 Transmission

The HPV transmission occurs primarily by skin-to-skin or mucosa-to-mucosa contact, thus care must be taken in order to prevent infection, especially as there are insufficient data to compute the probability of infection per sexual act, but due to the high incidence of infections for sure that the probability is not very low. Unfortunately, different types of HPVs may be transmitted simultaneously by the same transmission route, such as, sexual activity. This is critical as double to multiple infections seem to increase the probability of cervical cancer development (Schiffman, Castle, Jeronimo, Rodriguez, & Wacholder, 2007).

2.2.2 Infection

In order to establish an infection, the HPVs have to reach through cuts or abrasions the epithelial stem cells localized on the basal membrane, after this step, the viral DNA may stay inactive (latent) or become active. Within the natural infection of the HPV there are four steps: the latent phase, in which the viral genome stays in a plasmid state, the activation of the viral gene expressing condylomas or papillomas, this phase may not contain a strong viral DNA replication, given that the vast part of the epithelial lesions produce few virions. The transition to a dysplasia if the infection is persistent and the last one, the evolution to a carcinoma (Fuchsmann, Ayari-Khalfallah, Coulombeau, & Froehlich, 2011).

2.3 GENOME

The genome of the PVs is formed by circular double-stranded DNA (Bernard et al., 2010) of approximately 7,900 base pairs. On the other hand the genome organization of the whole set of virus that integrates the Papillomaviridae family has an early region (E), a late region (L) and a non-coding regulatory region (URR, upstream regulatory region). Some types of oncogenic HPVs may immortalize human keratinocytes, which are highly specialized epithelial cells (skin cells) that separate an organism from the outside world (Eckert & Rorke, 1989).

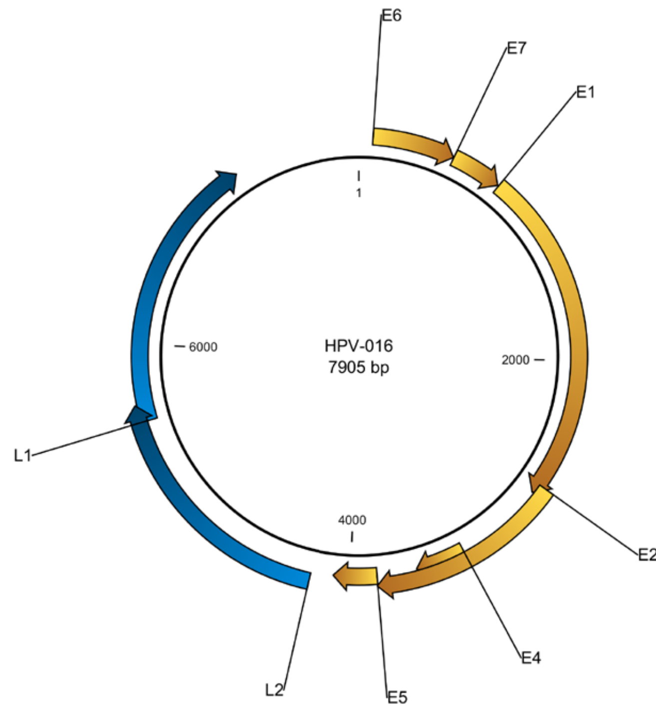


Figure 2.2: Human Papillomavirus 16 genome. Circular representation of the HPV genomes wherein blue arrows represent late genes and yellow early genes.

2.3.1 Oncogenic Activity

The oncogenic activity has been reported in the products of E6 and E7 genes, in which the E6 protein regulates the degradation of the p53 protein that suppresses tumors, whereas the E7 protein binds the product of the retinoblastoma gene and related proteins, a second important anti-oncogene. The proteins E1 and E2 regulate the viral DNA replication and gene expression. The L1 gene codes for the mayor protein of the capsid, it constitutes up to 80% of the virion's mass. L2 codes for a secondary protein of the capsid. The degree of homology between different isolates is used to define types, subtypes, and variants of the virus. If two HPVs differ in more than 10% of their L1 region sequence they belong to two different types. Nowadays, more than 200 HPV types have been described (Bernard et al., 2010) and some of them are associated with specific clinical manifestations, such as, squamous cell carcinomas, epithelial tumors, carcinoids, small cell carcinomas, sarcomas and lymphomas (Longo et al., 2011).

2.4 MECHANISM OF INFECTION

The human papillomavirus (HPV) can only reproduce itself within a terminally differentiating pluristratified squamous epithelium (Fauquet et al., 2005). Virus replication is limited to basal cells of a stratified epithelium, which is the only tissue type in which they replicate (Schiller, Day, & Kines, 2010). The mechanism starts when the virus particle binds to heparin sulfate proteoglycans (HSPGs) on the basement membrane that has been exposed by an injury, this binding induces a conformational change in which the site of the L2 segment of the virion is cleaved by furin or a proprotein convertase (PC 5/6). It is suspected that an epitope is exposed after the L2 cleavage and an unexposed region of the L1 segment is now available to bind to an unidentified secondary receptor localized on the epithelial cells (Figure 2.3). It is mainly believed that the virus is internalized in the epithelial cell through the endocytic pathway which is a slow process in comparison to the infectious pathways of other viruses: it takes 4 hours to become an early endosome, then after about 12 hours the virus uncoats while the endosome becomes a late one, and finally the L2 protein with the genome is released. During this step the L2-genome complex is transported through the cytoplasm, perhaps via microtubules, and enters the nucleus by 24 hours. After entering the nucleus, gene transcription may start (Schiller et al., 2010).

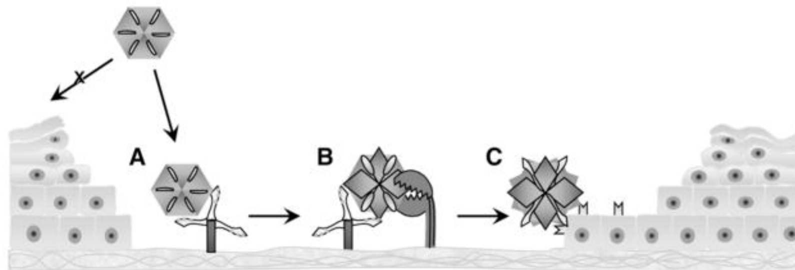


Figure 2.3: Binding of the HPV virion. (A) The virion encounters HSPGs on the basement of the membrane. (B) Conformational change in the virion structure exposing a site on L2 susceptible to furin or PC 5/6 cleavage. (C) After the L2 cleavage, a unexposed region of L1 binds to a unknown secondary receptor localized on the epithelial cells (Schiller et al., 2010).

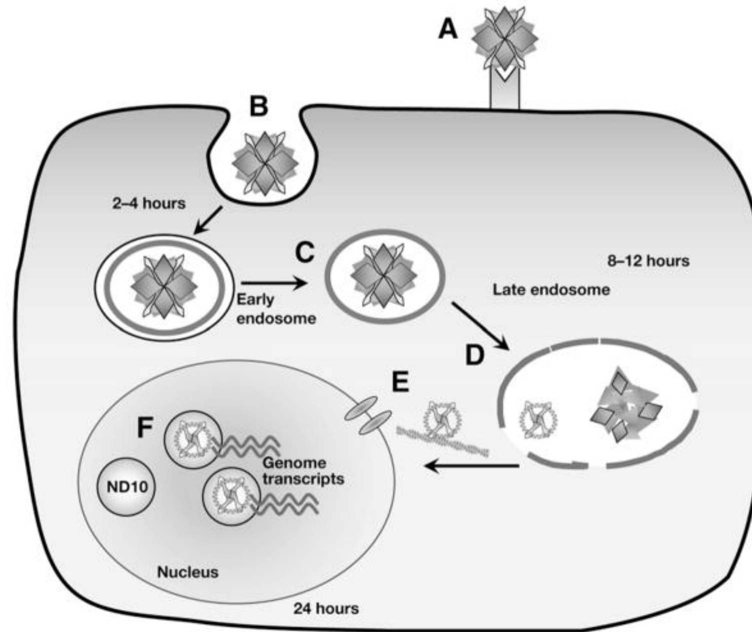


Figure 2.4: HPV virion within the cell. (A) The virus binds to an unknown receptor on the cell surface. (B) The virus is internalized in the epithelial cell via endocytic pathway. (C) After 4 hours the virus localize itself in an early endosome. (D) Between 8 and 12 hours the virus uncoats within a late endosome and a L2-genome complex is released. (E) The L2-genome complex travels through the cytoplasm, perhaps via microtubules and enters the nucleus by 24 h. (F) After nuclear entry, the L2-genome complex co-localizes with ND10 (Nuclear Domain 10) and RNA transcription begins (Schiller et al., 2010).

2.4.1 Replication

The replication of the HPVs occurs exclusively in keratinocytes, in the first stage the virus infects keratinocyte stem cells that are localized in the basal layer of the epithelium. As it does not code for an own DNA polymerase, after the initial infection the virus starts using the cell's machinery to replicate its genome. Upon continuing differentiation, the cells move towards the outer layers of the epithelium and the virus keep replicating without increasing their viral load, this is to remain unnoticed by the immune system (Stanley, 2006), see Figure 2.5.

However, when the infected keratinocyte enters the S-phase in the differentiation compartment of the epithelium, the virus genome will be replicated and reaches a high copy number of around 1000 per cell. This is when growth promoters E6 and E7 are liberated, which are the oncogenes of the virus. Expression of the capsid producing late genes L1 and L2 requires advanced to terminal differentiation. Finally when the keratinocyte dies at the superficial epithe-

lium, the HPV genomes are packaged into capsids and liberated (Doorbar, 2005).

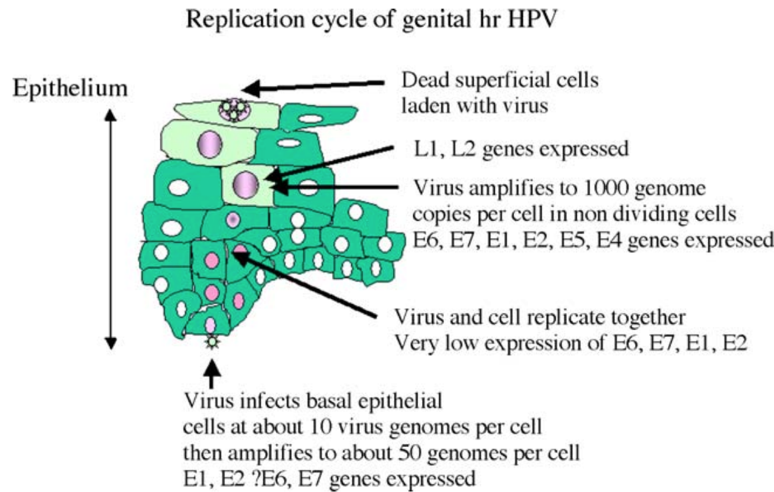


Figure 2.5: Replication cycle of genital high-risk HPV. The infection starts by a microabrasion of the mucosal epithelium. However, the infected cells move towards the outer layers of the epithelium. The evolution of the infection shows the expression of different genes and the copy number of the virus (Stanley, 2006).

2.4.2 Function of E6 and E7 proteins

E6 and E7 are pleiotropic and have different functions, for instance they interfere with trans-membrane signaling, regulation of the cell cycle, transformation of established cell lines, immortalization of primary cell line and regulation of chromosomal stability. A relationship between the E6 and E7 proteins and tumors has been associated because those proteins interact with and inactivate the main cell cycle regulating tumor suppressors p53 and pRB, respectively.

One of the functions of the E6 protein is to promote cell proliferation by stimulating degradation of the tumor suppressor p53 protein creating a trimeric complex between E6, p53 and the cellular ubiquitination enzyme E6-AP, which may be related to an increase in tumor cell growth. On the other hand the E7 protein binds to the retinoblastoma protein in a region so-called “pocket domains”, which are essential for tumor suppressor activity. However, one of the main functions of RB is to bind E2F-family transcription factors and repress the expressions of replication enzyme genes, this activity correlates with the tumor suppression function of RB, which is responsible for the liberation of E2Fs in their active forms that stimulates replication and cell-division (Yim & Park, 2005), see the figure below.

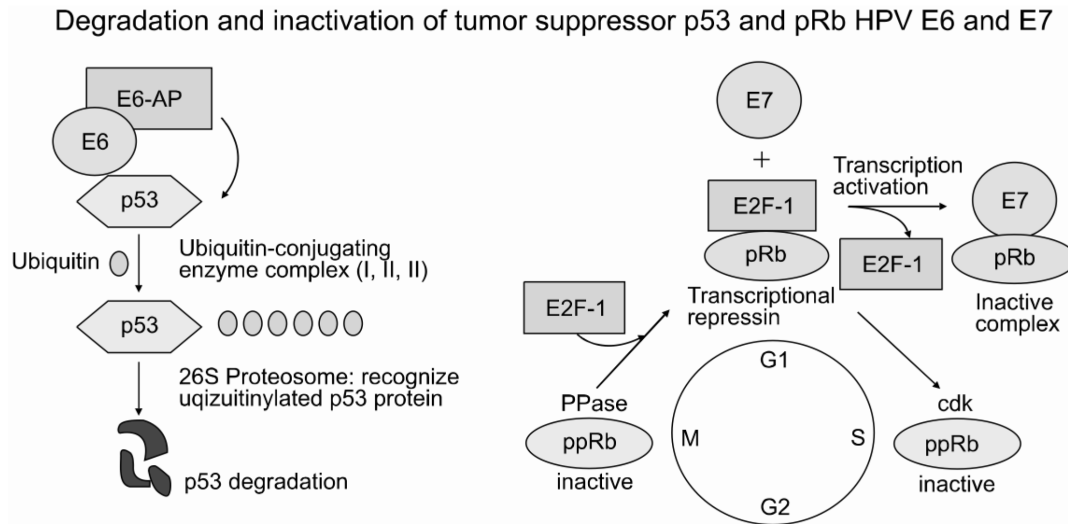


Figure 2.6: Binding of E6 and E7 oncoproteins. E6 binds to E6-AP and p53 creating a trimeric complex, whereas E7 binds to pRB (Yim & Park, 2005).

2.5 TROPISM

Papillomaviruses are species specific and prefer, depending on the genus and type different epithelial tissues, for example, sole of foot, non-genital skin, anogenital skin or anogenital/oropharyngeal mucosa. This is due to the evolution of the papilloma viruses over millions of years with different animal hosts (Schiffman et al., 2007). The tropism of the PVs is associated with its classification and can be used to divide them broadly into two groups: those who infect cutaneous epithelia and those who infect mucosal epithelia (Figure 2.7), it is well known that the mucosotropic viruses rarely infect cutaneous regions and vice versa (Bernard, 2005). Only a few viral types are regularly found in mucosal as well as cutaneous tissues as for example HPV2, however, even in this case HPV2 is preferentially found in cutaneous lesions.

There are five genera that affect the human beings named with Greek letters: Alpha, Beta, Gamma, Mu, and Nu (Bernard et al., 2010). The most interesting genus for humans are the Alpha and Beta PVs, within this two genera roughly 90% of all HPVs are found (Doorbar, 2006).

HPV6 and HPV11 are strongly related with respiratory papillomas, although they are frequently found in the genital area as well, it is worth to mention that these types are the most

frequent among the HPV-caused viral infections (Fuchsmann et al., 2011).

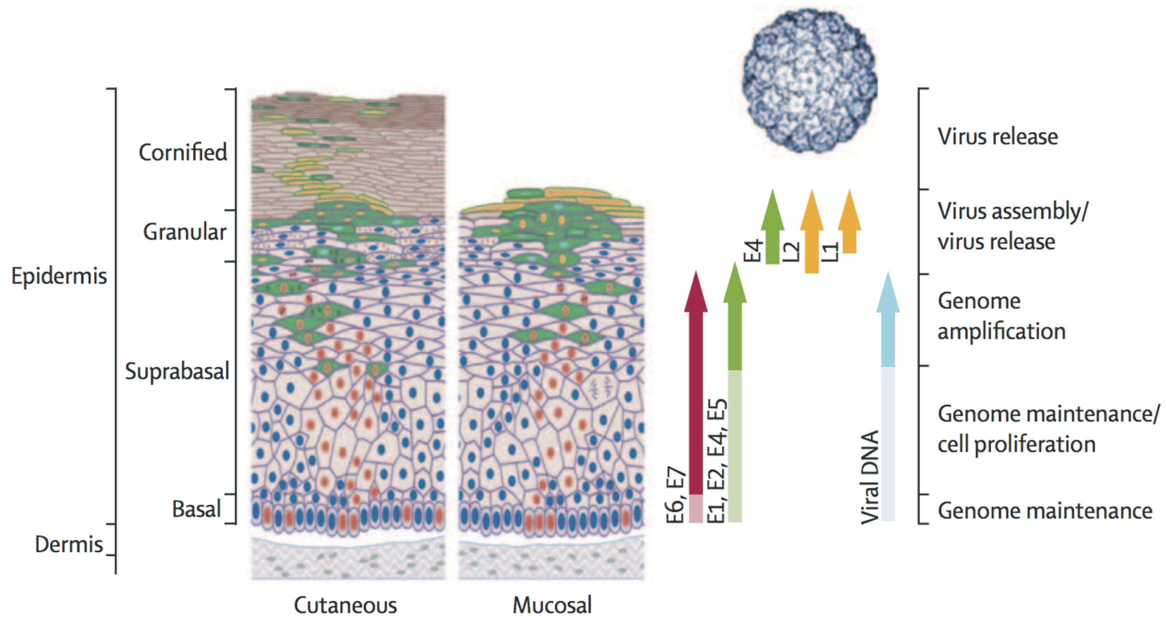


Figure 2.7: Tropism of HPV. In cutaneous or mucosal epithelia E and L genes designate early and late differentiation stage in the epithelium. E1, E2, E5, E6 and E7 are expressed early in the differentiation, meanwhile E4 is expressed at all times. L1 and L2 are expressed during the final stages of differentiation. The viral genome is maintained at the basal layer of the epithelium, where the HPV infection is established. As the basal epithelial cells differentiate, genome amplification, virus assembly and virus release take place. L1 and L2 assemble the viral capsid in which L1 is the major capsid protein and L2 serves as the link to the plasmid DNA (Schiffman et al., 2007).

Strong evidence that support the relationship between the HPVs and cervical cancer; i.e. primarily HPV16 and HPV18 have been found in about 95% of all biopsies gained from cancerous cervixes worldwide. Moreover the HPV16 is the king of all the types, it can be found in more than half of the infections, whereas HPV18 is found between 10 and 20%. Thus, HPV16 and HPV18 are the major threat for development of high grade Cervical Intraepithelial Neoplasia (CIN). There are, however, approx. 5% of cervical cancers without detectable viral DNA, this could mean two things, there are another factors that could induce cervical cancer and there are unidentified types of HPVs that are inducing those "special" cervical cancers.

HPV infections of the oral cavity were first detected in 1985, in those investigations the types that the researchers found were 16, 18, 6, 11 and 2, with a predominance of HPV16. Other regions of the human body were examined later-on finding that patients with laryngeal

papillomatosis with laryngeal or lung cancer contained HPV6 or HPV11 and just one case of HPV16. Moreover, in carcinomas of the nasal sinuses the types that were found were HPV16 and HPV18, whereas HPV57 was more related to an uncommon sinonasal tumour originated by inverted papillomas (Zur Hausen, 1996).

2.6 CONSEQUENCES OF HPV (CANCER)

The association between HPVs and cervical cancer began during the 1990s due to strong evidence of some HPVs infections led to the development of cervical cancer. The scientific community reviewed epidemiological and molecular studies and concluded that due to the lack of another documented hypotheses for the etiology of cervical cancer, the HPVs must be the principal factor. In other words, women must be infected by the viral HPV-DNA in order to develop cervical cancer (Bosch, Lorincz, Munoz, Meijer, & Shah, 2002).

In order to develop cancer there are 4 main steps, “the transmission, viral persistence, progression of a clone of persistently infected cells to precancer, and invasion” (Schiffman et al., 2007).

According to the World Health Organization, now the cervical cancer is on the third most common cancer in women with approx. 529,000 new cases in 2014 (WHO, 2014). The most affected regions by the cervical cancer are Eastern Africa, Melanesia and Southern Africa, whereas Central America is in the sixth place, followed by the Caribbean and South America, meanwhile the less affected regions are Northern America, Australia/New Zealand and Western Asia (WHO, 2012).

2.7 EPIDEMIOLOGY

From country to country around the world there are divergent statistics about the rate of new cases of cervical cancer caused by HPV. In countries like Egypt, China (Shanghai) and many European countries the incidence is 5 new cases per 100,000 women. On the other hand, in the Sub-Saharan Africa more than 45 new cases per 100,000 women are reported. Some interesting findings show that Hispanics living in the USA have higher rates of incidence and mortality

than whites. The incidence and mortality of cervical cancer around the world varies, however the most affected regions are Western and Southeastern Africa and Melanesia. Particularly this type of cancer is the most common in women in Eastern and Middle Africa, see Figures 2.8 and 2.9.

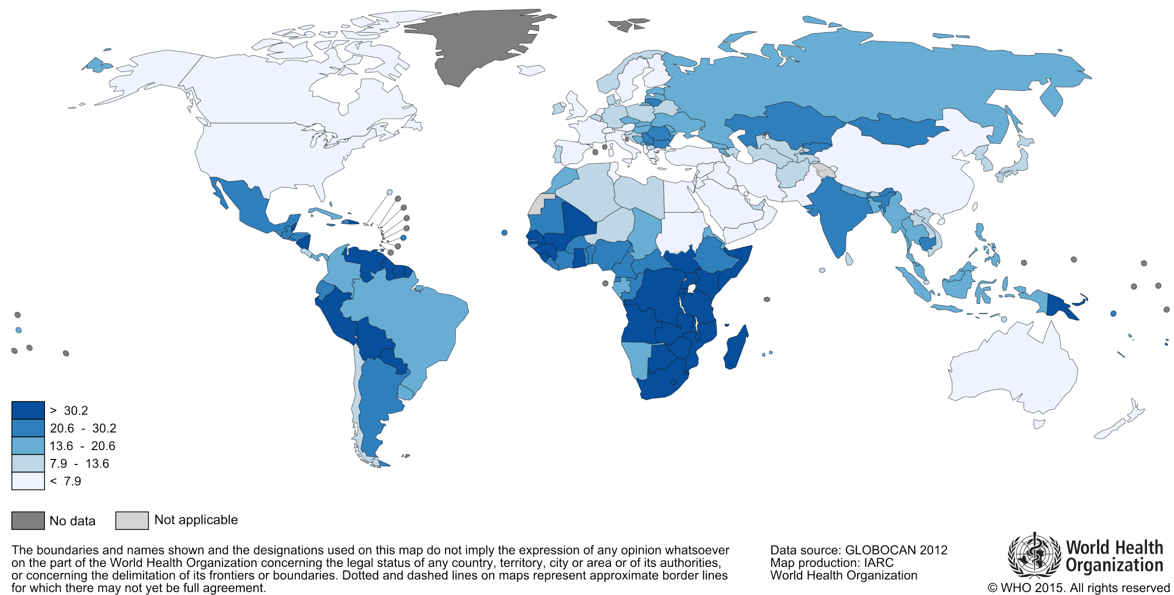


Figure 2.8: Estimated cervical cancer incidence worldwide in 2012 (WHO, 2012).

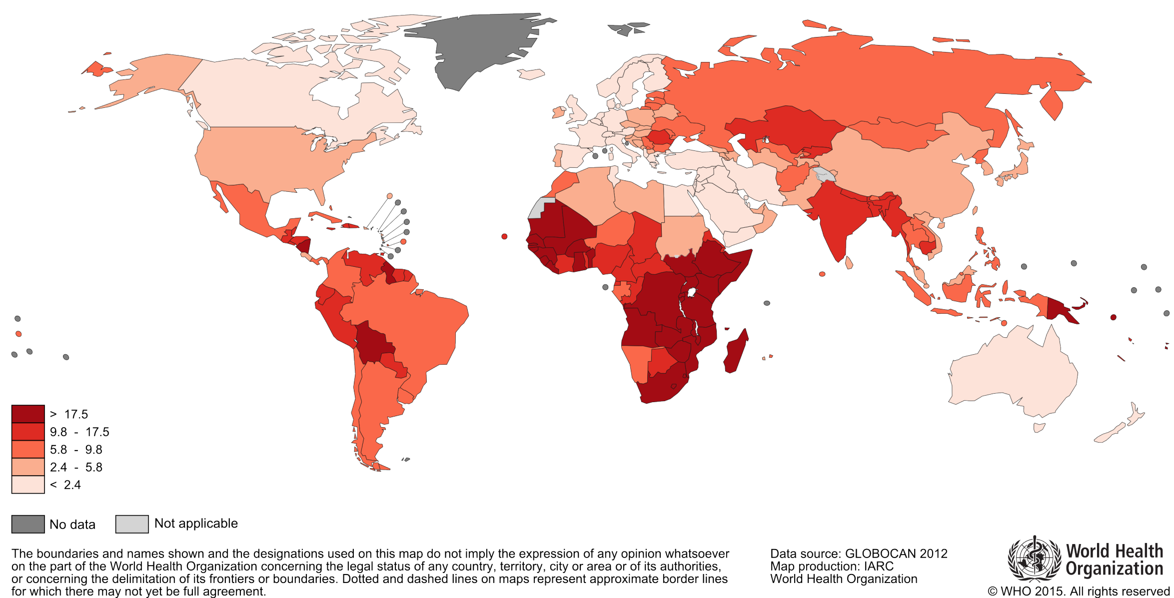


Figure 2.9: Estimated cervical cancer mortality worldwide in 2012 (WHO, 2012).

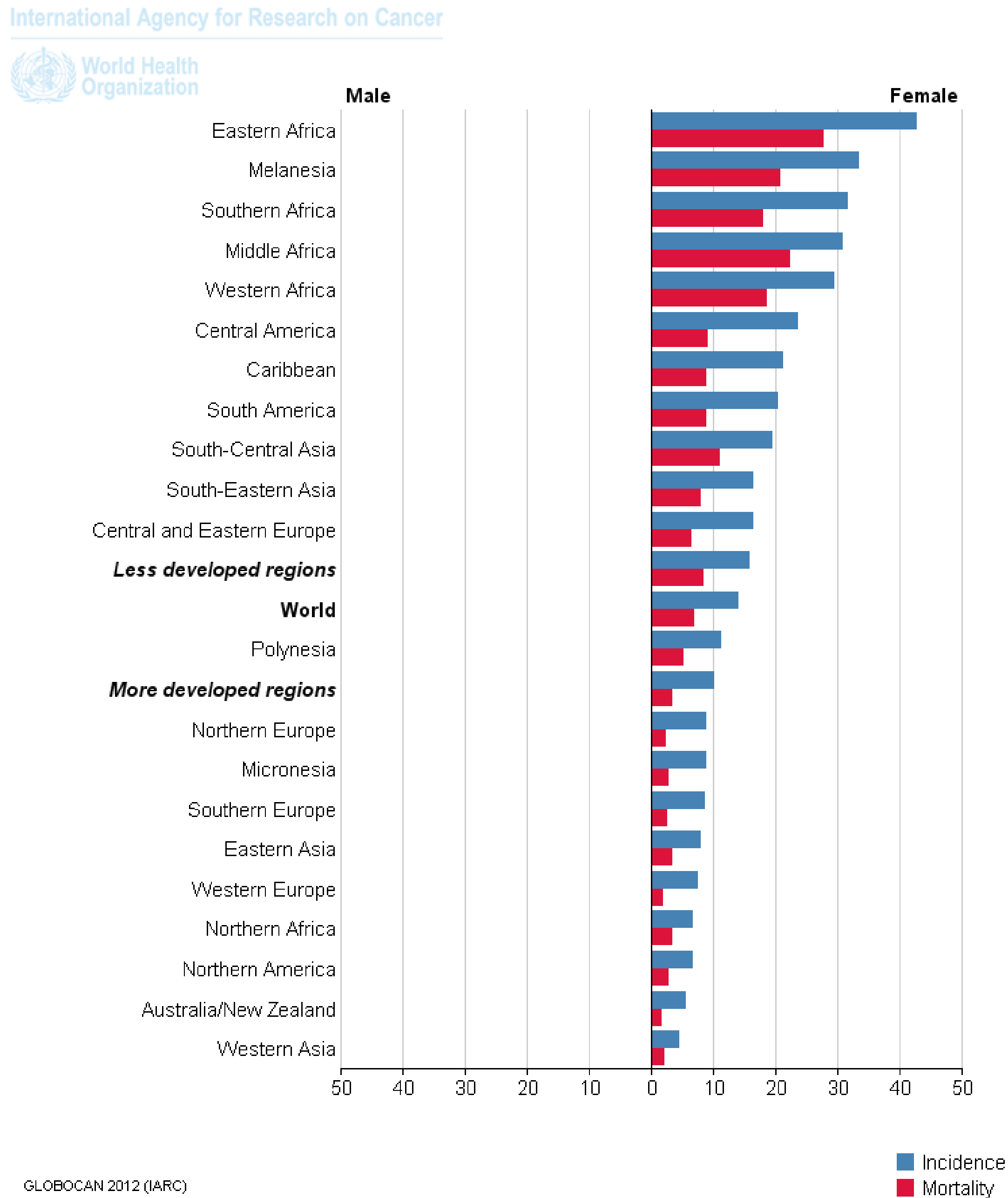


Figure 2.10: Incidence and mortality of cervical cancer by country. Estimated age-standardised rates (World) per 100,000 (WHO, 2012).

In Mexico there has been in the last 10 years a remarkable increase of the number of patients with cancer. However, the incidence of the cervical cancer remains on second place just below breast cancer, with a slope of 7.2 and 9.2 per 100,000 women, respectively (Secretaría de Salud, 2012). The Global pattern has shown that in developed countries, cancer rates are dropping,

whereas in developing countries, cancer rates are increasing (Figure 2.10). The trend of the increasing cancer rate in developing countries is probably related to the adoption of western life style rather than genetic differences between the populations. Nonetheless, there is no certainty of the main difference among the developed and developing countries, although access to better public health systems may be the most important one.

2.7.1 HPV-16 AA

It is believed that not only the life style has an influence upon the women's health when they are infected with HPVs. Although it is widely know that the high risk HPVs such as HPV16 and HPV18 are the main cause of cervical cancer, there may be other variants that affect the Mexican population. Variants differ by only 1-5% of their nucleotide sequence when compared to each other. In order to develop a change in the HPV genome, it takes from 10,000 to 100,000 years to yield a 1% sequence diversity. In other words the HPVs and the human beings have been together for a long time. As a result, certain variants predominate in certain ethnic groups.

In Mexico the prevalence of the HPV16 AA variant among HPV16 infections was of 88% in a study of women living in Monterrey, Nuevo León, Mexico and is the highest ever detected in any part of the world. The prevalence is unexpectedly high and specially frequent in people with an American Indian component but also in Spaniards, perhaps due to the reverse migration from Latin America. It is worth to mention that solely the HPV16 has been confirmed to have an increased carcinogenicity in its variants, if this is true the Mexican women are more prone to develop cervical cancer than previously thought (Calleja-Macias et al., 2004).

Worldwide cervical cancer is a major problem for the women, thus occupies the third place of the incidence of cancer and the fourth reason of death by cancer with approx. 529,000 new cases and 275,000 deaths.

Therefore the World Health Organization has provided a general guidance in which the newest and main elements are: vaccinate 9 to 13-year-old girls with two doses of HPV vaccine, use HPV tests to screen women for cervical cancer prevention and communicate the preventive information more widely to reach a wider audience (WHO, 2014).

2.8 DETECTION OF CERVICAL CANCER

The detection of women in risk to develop cervical cancer has been initially through the Papanicolaou test, which detects aberrant cells in cervical smears, low-grade squamous intraepithelial lesions (LSIL) and high-grade squamous intraepithelial lesions (HSIL). However, due to its inherent limitations in many cases a single Papanicolaou test does not assure the detection of cervical cancer. The false-negative rate (FNR) of a Papanicolaou smear is traditionally calculated as 6.1%. Yet the true FNR has even been calculated as 7.8% by rescreening abnormal smears (Renshaw, DiNisco, Minter, & Cibas, 1997).

Although the HPV is associated with cervical cancer in women, cancer does not develop with the sole presence of the virus. In order to develop cancer a persistent infection must be present throughout the pathology's development. In young women, the presence of HPV is considered a sign of sexual activity, while in women elder than 30 years who have normally more stable sexual relationships, the detection of HPV is interpreted as a persistent infection and thus at risk of cancer depending on the viral type. Of special importance is the continuing presence of just the two viral oncogenes E6 and E7, therefore, HPV-DNA methods have been designed, targeting the more conserved L1 gene as well as the E6 and E7 genes.

Until now, HPVs can not be cultured in the laboratory nor immunological assays are suitable to detect HPV infections. The main tools to detect HPV associated lesions have been cytology and histology. However, in contrast to the past, other tools such as molecular methods based in HPV-DNA have been incorporated in the diagnostics (Burd, 2003). The techniques can be either by target amplification or signal amplification.

2.8.1 Papanicolaou

During the last three decades the Papanicolaou ("Pap") smear or cytology has been the main tool for the detection of high-risk HPVs. It was initially proposed by George Papanicolaou in 1941 and was the first systematic method to detect early cancer in women. In developed countries this method has reduced the mortality and morbidity from cervical cancer, whereas in developing countries the pap smear is not as successful as in the developed countries, this may be due to

the lack of health infrastructure and financial investments to support the laboratory and human resources (WHO, 2014).

The traditional pap smear has three steps: the recollection of the cervical cells using a spatula or brush, the fixation of the cells on a slide with a chemical fixative and the identification of abnormal cells in the transformation zone and the junction of the ecto and endocervix, where cervical dysplasia and cancer originate (Cobo, 2012), see Figures 2.11 and 2.12.

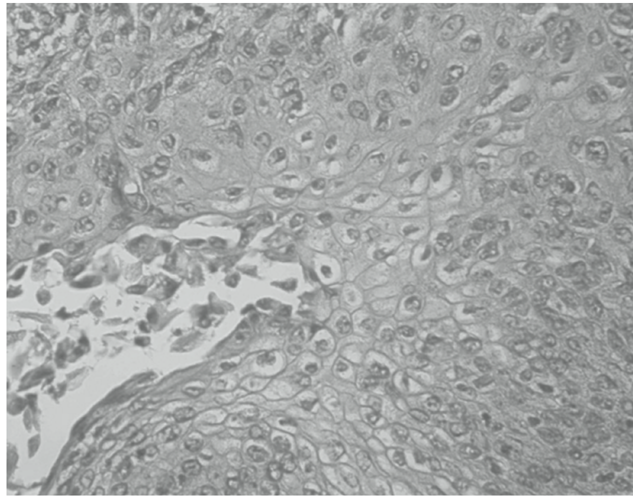


Figure 2.11: Pap smear with koilocytic atypia. Hyperchromatic and crenated nuclei surrounded by a clear halo of cytoplasma. It is widely believed that koilocytosis is correlated with vegetative HPV infections, however does not differentiate high-risk HPV types from low-risk HPV types. (Cobo, 2012).

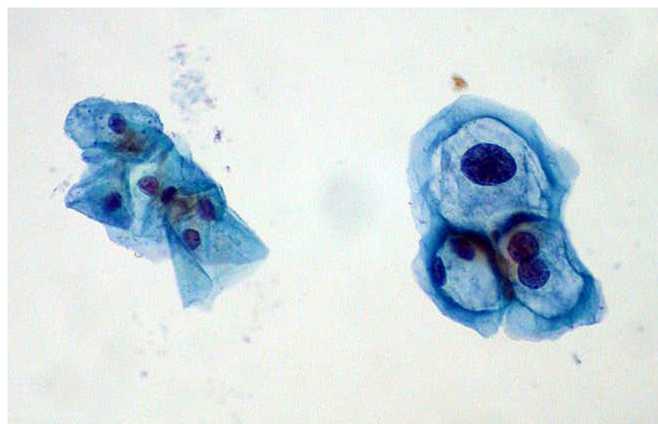


Figure 2.12: ThinPrep pap smear. A ThinPrep liquid-based pap with normal squamous cells on left and HPV-infected cells with a LSIL on right (Uthman, 2006).

2.8.2 End-Point PCR

The Polymerase Chain Reaction (PCR) technique exists since 1983 and was understood by few isolated people at that time, however the credit goes to Kary Mullis, who was awarded in 1993 by the Chemistry Nobel Price (Lewin, Krebs, Kilpatrick, & Goldstein, 2011). This technique broadly speaking has three steps: denaturation, annealing and elongation of the target DNA molecule. The temperatures, cycles and the efficiency of the reaction are variants that relay on the amount of target DNA present at any time (Freeman, Walker, & Vrana, 1999). This technique started by a single idea, "find" a sequence, likewise one would do it in any word processing program or in the Internet in any search engine. Kary Mullis developed this idea in 4 months in 1983 at the Cetus Corporation and his first course of action was to create a short piece of DNA that would find a sequence flanking gene and then it would reproduce itself over and over again. Hence, the chain reaction concept that it is in the name of the technique (PCR). This idea was based in the already known process of DNA auto replication when cells divide. In other words what Mullis basically did; he created the searchers or primers and put them into a tube with DNA, he uncoiled the DNA by heating (using a thermocycler) and added deoxynucleoside triphosphates (dNTP) that would be attached to the freshly uncoiled DNA at the location where primers bonded by a polymerase or the biological glue and finally by decreasing the temperature a new copy of DNA would be created. This technology was based upon a polymerase that could handle high temperatures that was found in the Yellowstone thermal water reservoirs, the bacteria *Thermus aquaticus*, which is stable in temperatures above that which DNA denatures, so it is the perfect fit to "glue" the dNTPs.

To view the product of a PCR, one must run an electrophoresis on an agarose or polyacrylamide gel, staining the gel with ethidium bromide and observing through on ultraviolet (UV) light source (Maurer, 2006).

2.8.3 Hybrid Capture II

The Hybrid Capture II method, by Digene Corporation, is an FDA approved liquid/solid phase signal-amplification technique for the detection of mayor high-risk HPV-DNA in cervical specimens. This method is based on hybridization in solution of long synthetic RNA probes that

match the genomic sequences of 13 high-risk (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68) and optionally 5 low-risk (6, 11, 42, 43, 44) HPV types. This method has five steps: denaturation of the DNA sample, mixing the DNA with the RNA probe, hybridization in the solution, capture of the hybrids by immobilized antibodies bounded to wells of a microtiter plate and the luminescent signal generated by chemical reactions (Cobo, 2012).

2.8.4 MY09/11

Among all the HPVs genomes one can find regions of DNA that share homology. Manos et al. identified in 1990 such regions in 5 HPV types (6,11,16,18 and 33) with 20 to 25 bp in length, this regions were degenerated to amplify distinct regions from over 25 of genital HPVs using an L1 consensus primer pair. Through PCR the set of consensus primers were used to amplify HPV-DNA, the positive strand primer is called MY11 (GCMCAGGGWCATAAYAATGG) whereas the negative strand primer is MY09 (CGTCCMARRGGAWACTGATC); The MY09/11 primer system. This primers are degenerated in different positions to complement to each of the 5 HPVs previously mentioned and produce an approx. 450 bp product. For the degenerated primers M stands for A or C, R for A or G, W for A or T, and Y for C or T. To identify and analyze the PCR products, acrylamide or agarose gel electrophoresis can be used, staining with ethidium bromide will show products that contain at least 200 copies of an HPV (Ting & Manos, 1990). There is only scarce information about type specific sensitivities of the MY09/11 system. Patti E. Gravitt argues that some types such as the HPV 16 and 18 have a sensitivity of around 10 genomes/PCR (Rohan & Shah, 2006). However, in another study that used PCR and RFLP; 25 copies of HPV18 and three copies of HPV16 were detected (Chen, Watanabe, Haruyama, & Kobayashi, 2013).

The sensitivities of the method to detect the HPV through PCR assays have been reported for some HPV types, which were analyzed with plasmids containing their entire genomes (Depuydt et al., 2007). The sensitivities are shown in the Table 2.1.

HPV	Sensitivity (HPV copies)
HPV 16	10^2
HPV 18	10^2
HPV 31	10^1
HPV 33	10^2
HPV 35	10^4
HPV 39	10^2
HPV 45	10^2
HPV 51	10^5
HPV 52	10^4
HPV 53	10^4
HPV 56	10^4
HPV 58	10^2
HPV 59	10^2
HPV 66	10^2
HPV 67	10^3
HPV 68	10^3

Table 2.1: Sensitivity of MY09/11 in PCR assays.

2.8.5 PGMY09/11

As already stated, so far the HPVs cannot be cultured, therefore the detection methods relay upon the detection of viral nucleic acids (Brink, Snijders, & Meijer, 2007). It is worth to mention some others detections systems such as the PGMY09/11 primer system, which is a detection method based on the binding regions used for MY09/11. Instead of using degenerated bases, heterogeneity is created through mixture of various primers sequences (Roulston & Bartlett, 2004), the arrangement of the primers are a set of 5 upstream oligonucleotides (PGMY11) and a set of 13 downstream oligonucleotides (PGMY09) (Gravitt et al., 2000). This consensus primers were designed to improve the reproducibility, sensitivity and specificity of L1 consensus PCR (Coutlée et al., 2002).

2.8.6 GP5+/GP6+

The GP5+/GP6+ detection system is conformed by a mixture of primers with a fixed nucleotide sequence (Qu et al., 1997), amplifying the L1 consensus region. The length of the amplicon is

150 bp and amplify a wide range of HPV types in a PCR, this is accomplished by lowering the annealing temperature during the PCR. However, the products have sufficient internal heterogeneity to produce a new set of specific primers or probes to characterize the HPVs (Roulston & Bartlett, 2004).

2.9 VACCINE

The HPV vaccine is made from empty protein shells or virus-like particles (VLP), this technology is available through biotechnology using recombinant technology, in which empty VLP are synthesized containing neither DNA nor live biological products and hence this particles are non-infectious. Currently there are three HPV prophylactic vaccines: Gardasil® from Sanofi Pasteur MSD, a tetravalent vaccine (HPVs 6, 11, 16 and 18); Gardasil®9 from Merck Sharp and Dome, a nonavalent vaccine (HPVs 6, 11, 16, 18, 31, 33, 45, 52 and 58); and Cervarix® from GlaxoSmithKline Biologicals, a bivalent vaccine (HPVs 16 and 18).

Manufacturer and trade name	Quadrivalent vaccine	Bivalent vaccine
	Merck [Gardasil]	GlaxoSmithKline [Cervarix]
Virus-like particles [VLPs] of genotypes	6, 11, 16, 18	16, 18
Substrate	Yeast [*S. cerevisiae*]	Baculovirus expression system
Adjuvant	Proprietary aluminium hydroxyphosphate sulfate (225µg) (Merck aluminium adjuvant)	Proprietary aluminium hydroxide (500 µg) plus 50 µg 3-deacylated monophosphoryl lipid A (GSK AS04 adjuvant)

Table 2.2: Characteristics of two HPV vaccines (Cutts et al., 2007).

However, the vaccine by itself is only half the answer to address the HPV infections, as in countries with limited or poor health care systems the vaccine is not widely available. The high cost of over $100 dollars per dose (tetravalent vaccine) is a major obstacle. In order to be feasible, immunization programs in developing countries should cost less than $10 to $25 dollars. Thus avoiding, the death of over four million women worldwide over the next decade (Koutsky, 2009). Studies have shown that the HPV vaccines are very effective with protection that lasts for at least 5 years. Nonetheless, continuous screening in vaccinated women will remain important, because although the HPV vaccines reduce the risk of cervical cancer, they do not eliminate it (Cutts et al., 2007).

2.10 BIOINFORMATICS

During the last 15 years, many techniques have been developed in a computer programs as a demonstrative example of how to apply computational concepts to biology, however it seems that now is a real prerequisite for model building and different technologies such as, sequencing technologies, microarrays, spectral analysis, data mining, etc. (Saeys, Inza, & Larrañaga, 2007).

Bioinformatics is a relative new and rapidly evolving discipline with the aim to use computational tools and techniques to manage and analyse biological data sets.

Nowadays with high speed Internet connections around the world and new technologies such as, Next-Generation Sequencing (NGS), biological data is increasing rapidly. Public databases such as GenBank and the Protein Data Bank have increased their catalog exponentially in recent years. Therefore, the information can be accessed rapidly, easily and virtually at no cost from anywhere around the world (Tisdall, 2001).

2.10.1 C

Since the 1940's many programming languages have been created to assist engineers. The original idea of a programming language was to be capable of solving any numerical or logical problem (Rojas et al., 2000). As the years passed by many other programming languages were developed. Yet, it was not until 1972 that the C programming language was developed, which is a general-purpose programming language, among its features one can find economy of expression, modern control flow, data structures and a rich set of operators. The language is not very high-level and it is not specialized to any particular area of application and hence it is more convenient and effective than other programming languages (Kernighan & Ritchie, 1988).

2.10.2 Python

On the other hand there are newer and easier to learn programming languages one of them is Python a very high-level programming language with growing use in commercial and academic applications. Some of its characteristics are: easy syntax to learn, object-oriented programming capabilities, a wide array of libraries, it can interface with optimized code written in C,

C++ or even FORTRAN and in company with `numpy` (Numerical Python), Python is a great choice for scientific programming. In addition, since 1999 Chapman and Chang founded the Biopython project, which is "a mature open source international collaboration of volunteer developers, providing Python libraries for a wide range of bioinformatics problems. Biopython includes modules for reading and writing different sequence file formats and multiple sequence alignments, dealing with 3D macro-molecular structures, interacting with common tools such as BLAST, ClustalW and EMBOSS, accessing key online databases, as well as providing numerical methods for statistical learning" (Cock et al., 2009).

2.11 R PROGRAMMING

R is a computer programming language for statistical data manipulation and analysis, which by itself is an implementation of the S language developed by AT&T, S stands for *statistics*. In addition S was later sold to a small company, which added a graphical user interface, the company renamed the programming language as S-PLUS. However, R has become more popular than S or S-PLUS, because it is free and there is a big online community contributing to its development (Matloff, 2011).

The benefits of using the R programming language for scientific computation are "the existence of a substantial collection of good statical algorithms, access to high-quality numerical routines and integrated data visualization tools" (Gentleman, 2008). One of the main advantages of R is the reproducibility of any statistical analysis, through Sweave or Knit package. The ability to integrate text and code into a single document facilitates the writing of scientific papers and ensures that all figures, graphics and facts are based on the same data and are reproducible by anyone.

CHAPTER 3

MATERIALS AND METHODS

3.1 SEQUENCES

The sequences that were used to design the type-specific HPV primers were obtained from the NCBI GenBank database, and limited to those types that had been previously detected in Clínica Ruiz. The retrieved sequences were studied around the L1; i.e. L2 and the URR. See Table A.3.

3.2 PRIMER DESIGN

The primer design to the specific HPVs was done through an algorithm based on Biopython to select the L1 region and its surroundings (L2 and URR). Primer-BLAST was used to generate each HPV set of primers (Figure 3.1). This software uses Primer3 and BLAST to find the target-specific primers, hence avoids primer pairs that may cause non-specific amplifications (Ye et al., 2012). The used criteria to select the best primers sets are documented in Table 3.1.

Criteria	Parameter
Amplicon size	$<$ 1000 bp
Difference in T_m	min
Self 3' complementary	min

Table 3.1: Criteria for the HPV type-specific primers.

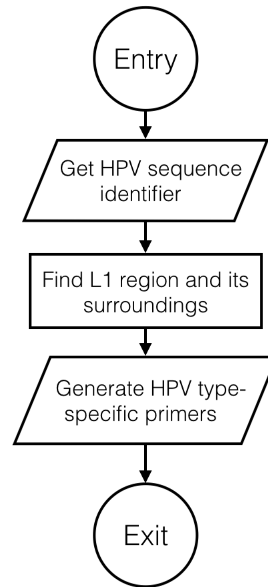


Figure 3.1: Algorithm to find the HPV type-specific primers. The Biopython script searches the HPV sequence identifier in the NCBI database and retrieves the L1 region and its surroundings (L2 and URR). Using Primer-BLAST and the retrieved regions different type-specific primers are generated.

3.3 SAMPLES

The samples were obtained from stored DNA extractions of cervical scrapes previously processed, typed and banked by the laboratory between 2012 to 2015.

3.4 PCR AMPLIFICATION

The PCR reaction for the HPV type-specific amplification used the primers that were designed and synthesized by IDT®. The reaction concentrations were: 2.5 U/μl HotStarTaq DNA Polymerase, 1x enzyme buffer (Tris·Cl, KCl, $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$; pH 8.7 (20°C); the concentrations, except for $MgCl_2$ have not been specified by QIAGEN), 0.2 μM dNTPs, 0.5 μM type-specific primer, 1 μl HPV-DNA and filled up with H_2O grade HPLC in a volume of 25 μl per reaction.

The PCRs were performed in a Veriti® Thermal Cycler. The reactions were submitted to the following amplification program: denaturation of 15 min at 95°C, followed by 40 cycles of

30 sec at 94°C, 30 sec at the optimized hybridization temperature, 1 min at 72°C and 30 min at 72°C at the end of the cycles for extension.

3.5 GEL ELECTROPHORESIS

A polyacrylamide gel at 4.5% was used to reveal the PCR products, TBE was used as the buffer, the gel ran at 200 volts and was stained with EtBr.

3.6 FRESH COMPETENT BACTERIA

Competent bacteria from the One Shot® Mach1™ kit showed low transformation efficiency, therefore the bacterial cells were grown in a plate with solid LB medium and incubated for 24 hours at 37°C. The cells were treated to make them competent using a standard protocol with $CaCl_2$ (Sambrook, Fritsch, & Maniatis, 1989).

3.7 DNA CLONING

The transformation of the cells were done using the TOPO® TA Cloning® Kit and the pCR™ 2.1-TOPO® vector in a final volume of 6 μl and 1 μl of fresh HPV PCR product.

3.8 PLASMID EXTRACTION

The plasmid with the HPV insert was extracted through a standard alkaline lysis protocol with SDS (Sambrook et al., 1989).

3.9 DNA PURIFICATION

The plasmid DNA was purified through a phenol/chloroform extraction. Phenol pH: 8.0, chloroform:isoamyl alcohol (24:1), sodium acetate pH: 5.2, EtOH 100%, EtOH 70% and resuspended with 200 μl of TE buffer.

3.10 DNA QUANTIFICATION

The double stranded DNA extracted and purified was quantified using a LightCycler® 2.0 from Roche and the Quant-iT™ PicoGreen® dsDNA Assay Kit.

3.11 DNA COPY NUMBER

The quantification of the DNA copy number (cp/μl) was obtained by the concentration obtained with the Quant-iT™ PicoGreen® dsDNA Assay Kit, the average weight of a DNA base pair and the Avogadro's number.

3.12 SENSITIVITY

To determine the sensitivity of the MY09/11 system, the plasmids were linearized by restriction digestions, re-purified and diluted. The dilutions were made from 10^{10} to 10^{0} copies per μl with human DNA at 0.025 ng/μl. Each dilution from 10^{5} to 10^{0} was first amplified in duplicate to find the approximate detection limit and then intermediate dilutions were made to determine the exact sensitivity limit. The PCR conditions were: 2 U/μl FastStart™ Taq DNA Polymerase, the buffer enzyme, 0.2 mM dNTPs, 0.4 μM MY09/11 primer mix, 1.7 mM $MgCl_2$, 2 μl of diluted HPV-DNA and filled up with H_2O grade HPLC in a volume of 25 μl per reaction. The PCRs were done in a Veriti® Thermal Cycler following the amplification program: denaturation of 4 min at 95°C, followed by 40 cycles of 15 sec at 94°C, 50 sec at 55°C, 30 sec at 72°C and 2 min at 72°C at the end of the cycles for extension.

3.13 LOGISTIC REGRESSION

A logistic regression model (Kleinbaum & Klein, 2010) was built using the data points obtained from the PCR positive amplifications of the HPV-DNA dilutions. The model was constructed in the R programming language and the sensitivity limit was calculated by interpolation at 95% (R Development Core Team, 2011).

CHAPTER 4

RESULTS

4.1 PRIMER DESIGN

Type-specific primers were design to each of the HPVs that were available at the LCP with the objective of isolate the HPVs to later clone them, 67 different types in total. The sequences were obtained from the NCBI website and are listed in the Table A.3, the sequences complete genomes were analyzed to find the L1 region and its surroundings, both the HPV type and the extracted sequences were compiled into a new database, which was used to generate possible type-specific primers around the region MY09/11, through Primer-BLAST, see Figure 4.1.

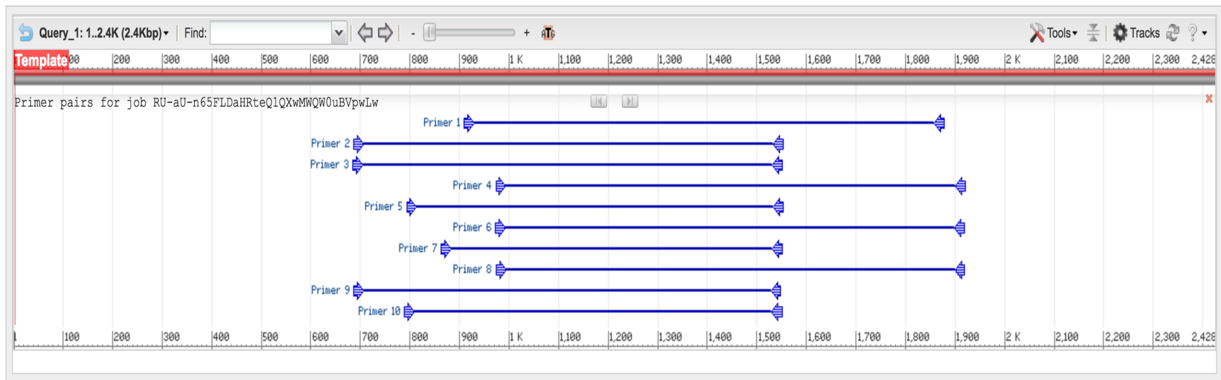


Figure 4.1: Primer-BLAST design. The design of the type-specific primers considered the outside boundaries of the MY09/11 region to later test the specificity of the MY09/11 system within the amplified region.

Ten different primers were generated flanking the MY09/11 region of each HPV (Figure 4.2), the criteria to select the best pair was based on the amplicon size with less than 1000 bp, a minimum difference in the T_m of the forward and reverse primers and a low 3' self-complementarity in both forward and reverse primers. Subsequently the primers were synthesised by IDT®, a complete list is shown in Table A.4.

Primer pair 1

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	AAAGGCTCTGGGTCTACTGC	Plus	20	910	929	59.38	55.00	3.00	2.00
Reverse primer	GGGCCTGGCGCTACAAAATA	Minus	20	1877	1858	60.75	55.00	5.00	2.00
Product length	968								

Primer pair 2

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	TGGCTTTGGTGCTATGGACT	Plus	20	687	706	59.30	50.00	2.00	1.00
Reverse primer	AGGTGGTGGGTGTAGCTTTT	Minus	20	1552	1533	59.15	50.00	4.00	0.00
Product length	866								

Primer pair 3

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	ACTGGCTTTGGTGCTATGGA	Plus	20	685	704	59.30	50.00	2.00	0.00
Reverse primer	GTGGTGGGTGTAGCTTTTCG	Minus	20	1550	1531	59.13	55.00	4.00	2.00
Product length	866								

Primer pair 4

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	TGGTTACCTCTGATGCCCAA	Plus	20	974	993	58.63	50.00	5.00	0.00
Reverse primer	GTGCCAAAAAGCATGCAACC	Minus	20	1920	1901	59.41	50.00	6.00	2.00
Product length	947								

Primer pair 5

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CAGAACCATATGGCGACAGC	Plus	20	794	813	59.06	55.00	8.00	2.00
Reverse primer	GAGGTGGTGGGTGTAGCTTTT	Minus	21	1553	1533	60.20	52.38	4.00	0.00
Product length	760								

Primer pair 6

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	ATGGTTACCTCTGATGCCCA	Plus	20	973	992	58.40	50.00	5.00	2.00
Reverse primer	GCCAAAAAGCATGCAACCGA	Minus	20	1918	1899	60.60	50.00	6.00	0.00
Product length	946								

Primer pair 7

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	TAGGGCTGGTACTGTTGGTG	Plus	20	864	883	59.02	55.00	4.00	0.00
Reverse primer	GGTGGTGGGTGTAGCTTTTC	Minus	20	1551	1532	58.76	55.00	4.00	0.00
Product length	688								

Primer pair 8

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	TGGTTACCTCTGATGCCCAAA	Plus	21	974	994	59.29	47.62	5.00	0.00
Reverse primer	TGCCAAAAAGCATGCAACCG	Minus	20	1919	1900	60.88	50.00	6.00	2.00
Product length	946								

Primer pair 9

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	TGGCTTTGGTGCTATGGACTT	Plus	21	687	707	59.92	47.62	2.00	0.00
Reverse primer	GGTGGGTGTAGCTTTTCGTT	Minus	20	1548	1529	58.40	50.00	4.00	0.00
Product length	862								

Primer pair 10

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	GGTGTCAGAACCATATGGCGA	Plus	21	789	809	60.13	52.38	8.00	2.00
Reverse primer	GTGGTGGGTGTAGCTTTTCGT	Minus	21	1550	1530	60.81	52.38	4.00	0.00
Product length	762								

Figure 4.2: Example of generated primers through Primer-BLAST. HPV16 type-specific primers generated by Primer-BLAST around and including the MY09/11 region.

4.2 TYPE-SPECIFIC AMPLIFICATIONS

To amplify HPVs with the designed type-specific primers, samples from the LCP where selected, which had been previously typed by RFLP. Type-specific primers were matched with the expected HPV in the stored sample. As optimal amplification conditions were not known for these primers, PCR with a temperature gradient across a thermal block of 55-62°C (hybridization temperature) was performed. Amplification conditions that do not yield non-specific products were chosen for subsequent amplifications. No further amplification parameters were optimized. In Figure 4.3 the electrophoresis gels of some products of the type-specific amplifications are shown and based on the temperature the amplification products that were selected were the ones without non-specific amplifications.

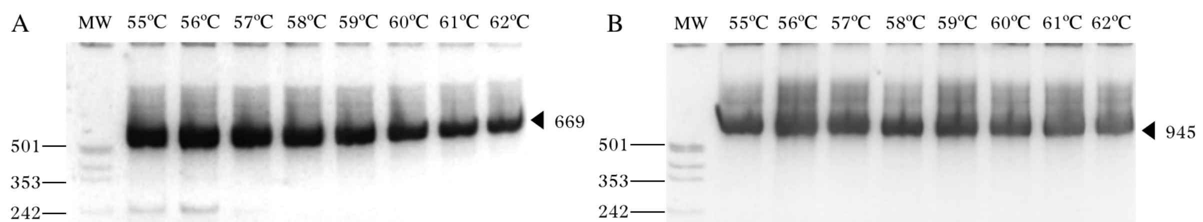


Figure 4.3: Type-specific PCR amplification. Electrophoresis gel of the products of the PCR for type-specific amplifications established on the PCR amplification method with a temperature gradient across the thermal block. MW: molecular weight (pUC18 / HpaII). (A) HPV16 has a better amplification profile above 58°C (hybridization temperature), while (B) HPV18 above 56°C.

4.3 COMPETENT BACTERIA AND DNA CLONING

Competent bacteria from the TOPO® TA Cloning® Kit showed low transformation efficiency. Therefore, fresh competent bacteria were prepared from the supplied Mach1™ *E. coli* strain by standard $CaCl_2$ treatment. The low efficiency bacteria were cultivated on a Petri plate with solid LB medium for 24 hours at 37°C, then one colony was selected to be cultivated in 2 ml of liquid LB medium for 24 hours at 37°C with agitation. The very next day, the OD_{600} was measured and then the bacteria with the medium were diluted with liquid LB medium to OD_{600} = 0.01 and incubated at 37°C, its OD was measured every 30 min until it reached a OD_{600} = 0.2. After that the fresh competent bacteria and the DNA cloning methods were followed obtaining

a satisfactory transformation efficiency of 1.92×10^6 transformants/μg DNA.

4.4 CULTIVATION AND PLASMID EXTRACTION

To cultivate the transformed bacteria, solid LB medium with ampicillin/kanamycin was used. Five isolated transformed bacteria were selected to be grown in liquid LB medium of each HPV type in a volume of 2 ml for 24 hours at 37°C with agitation. The incubated bacteria were later cultivated on solid LB medium with antibiotic and stored at 4°C. One milliliter of the transformed bacteria that were grown in the liquid LB medium was later used it in a minilysates protocol (Sambrook et al., 1989) to extract the vector with the insert of the HPV. After the vector extraction a digestion cut out the insert to visualize it in an agarose gel, see the Figure 4.4 below.

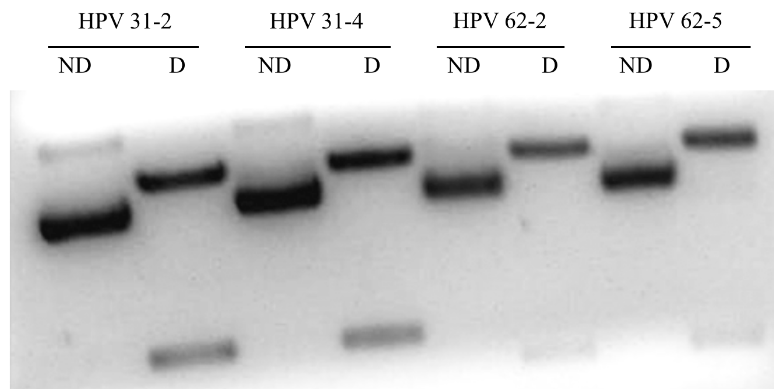


Figure 4.4: HPV Minilysates. Agarose gel with the digested amplicons (EcoRI) of the HPVs 31 (colony 2 and 4) and 62 (colony 2 and 5) that were cloned through the pCR™ 2.1-TOPO® in competent Mach1™ cells. ND = Not Digested, D = Digested.

Once the successful cloning had been confirmed, the transformed strain was grown in 100 ml of liquid LB medium + antibiotic for 24 hours at 37°C with vigorous agitation. The next day, a standard alkaline lysis protocol with SDS (Sambrook et al., 1989) was used to extract the plasmid and resuspend it in 250 μl with TE+RNase.

4.5 DNA PURIFICATION

A good DNA purification step is essential for a correct DNA quantification, therefore the extracted plasmid was purified through a phenol/chloroform extraction in which firstly the DNA extraction was raised to 400 μl with TE, an equal volume of phenol was added to the tube and then a three step process was followed: invert for 10 min, centrifuge 3 min at max speed and transfer the supernatant to a new tube. In addition, an equal volume of chloroform:isoamyl alcohol (24:1) was added to the new tube followed by the three step process. Additionally, the chloroform:isoamyl alcohol step was repeated twice to eliminate impurities. On a new tube with the supernatant one tenth of NaOAc and 2.5 times of EtOH 100% of the supernatant volume were added and then incubated on ice for 10 min. The tube was then centrifuged at max speed for 5 min, the supernatant was thrown away and approx. 2 ml of cold EtOH 70% was added and incubated for 1 min. The tube was centrifuged at max speed for 5 min and the supernatant was thrown away, the DNA pellet was resuspended with 200 μl TE.

After the DNA purification procedure a digestion was performed to ensure that the vector and insert were not altered during its growth in the big volume of 100 ml of liquid LB medium + antibiotic. Albeit this step had already been done in the small scale culture, it was required to assure the quality of the reference material after large scale culture. The digestions are shown in the Figure 4.6.

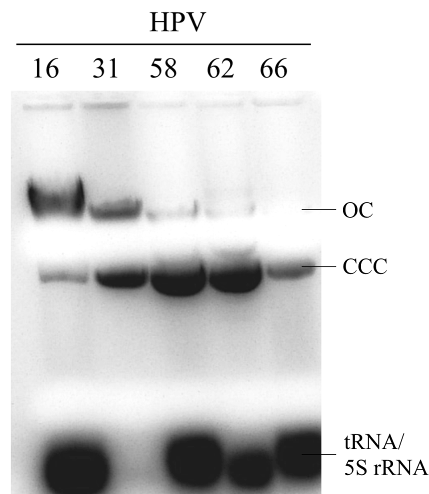


Figure 4.5: DNA's large scale culture gel. HPV 16, 31, 58, 62 and 66 in agarose gel without digestion before DNA quantification. OC = Open Circular, CCC = Covalently Closed Circular.

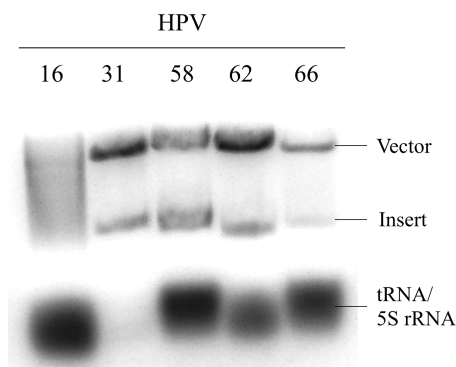


Figure 4.6: Digestions of purified plasmid DNAs. Vector and inserts of HPV 16 (HindIII + EcoRV), 31, 58, 62 and 66 digested with EcoRI that show two fragments indicating successful transformation in a large scale culture before DNA quantification.

4.6 DNA QUANTIFICATION

The DNA quantification of the purified DNA was done with the Quant-iT™ PicoGreen® ds-DNA Assay Kit, which uses a fluorophore that becomes fluorescent upon binding to DNA, the fluorescence intensity depends on the amount of formed complex. This technology gives a more accurate results than the ones obtained by UV absorbance readings, as difficult to remove small RNAs (tRNA, 5S rRNA, see intense low-molecular weight band in Figure 4.5) present in the DNA preparation tend to largely overestimate its DNA concentration (Thermo Fisher Scientific, 2006). A standard curve from 0 ng/μl to 10 ng/μl was used to determine the sample concentration (Figure 4.7). The sample was diluted based on its optical density at 260 nm and the dilution factor was later used to determine the original DNA concentration of the sample. For the experiment a LightCycler® 2.0 from Roche was used. The results are shown in Table 4.2.

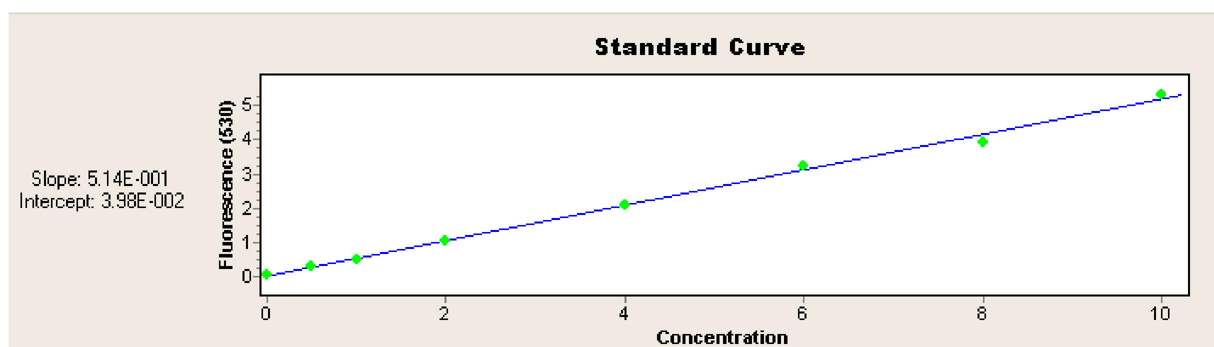


Figure 4.7: Calibration curve of DNA quantification. Concentrations of the different standards for the calibration curve from 0 to 10 ng/μl. See data on Table 4.1.

Concentration (ng/μl)	Standard (ng/μl)
[0.0562]	0
0.537	0.5
0.879	1.00
2.00	2.00
4.03	4.00
6.23	6.00
7.55	8.00
[10.2]	10.0

Table 4.1: Calibration data from the standards. DNA quantification curve to determine the concentration of the HPVs after large scale growth and purification.

HPV	Concentration (ng/μl)
16	232
31	113
58	336
62	309
66	154

Table 4.2: HPV-DNA concentration. Concentrations of the vector + insert after large scale growth and DNA purification. The dilution factor was used to calculate the original DNA concentrations on the samples.

4.7 DNA COPY NUMBER AND SENSITIVITY

To calculate the HPV-DNA copy number in the sample different data was gathered such as, its DNA concentration and the length of the vector + insert. The molecular weight was calculated multiplying the total length times the average weight of a base pair of 650 Da and the Avogadro's number was expressed as $6 \times 10^{11} cp/pmol$. The formula is shown in Equation 4.1 and the results in Table 4.3.

$$\frac{\text{Concentration}\ \ pg/\mu l}{\text{Molecular weight}\ \ pg/pmol} \cdot 6 \times 10^{11}\ \ cp/pmol = \boxed{\text{cp/}\mu\text{l}} \tag{4.1}$$

HPV	Copy Number (cp/μl)
16	4.64×10^{10}
31	2.12×10^{10}
58	6.36×10^{10}
62	6.09×10^{10}
66	2.94×10^{10}

Table 4.3: HPV copy number. Calculus to determine the copy number per μl of the different HPV using the eq. 4.1 after large scale growth, DNA purification and DNA quantification.

Once the concentration in copies per μl had been determined, the sensitivity of the MY09/11 method was elucidated with the help of a dilution series (Figure 4.8). Each dilution was amplified ten times and the number of successful amplifications was recorded (Table 4.4). A logistic model was constructed and the sensitivity limit of the MY09/11 method was determined at 95% of successful amplifications (Figure 4.9). The results are summarized in Table 4.5.

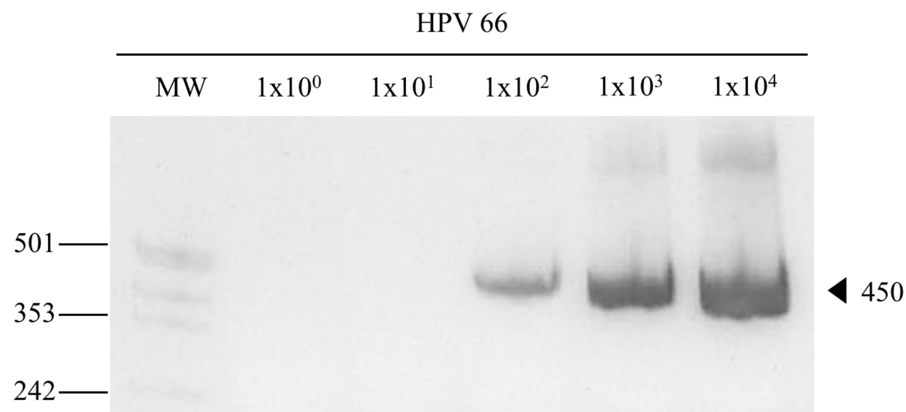


Figure 4.8: Initial dilution of HPV66. Aliquots of 1×10^4 down to 1×10^0 cp/μl of HindIII digested plasmids were subjected to amplification with the MY09/11 system. The approximate sensitivity level lies between 1×10^2 and 1×10^1 cp/μl.

Quantity (cp)	Positive	Negative
1.11	2	8
3.33	5	5
10	7	7
33.33	9	1

Table 4.4: Intermediate dilutions of HPV66 with concentrations between the limits as defined previously were subjected to ten amplifications. The number of successful amplifications has been recorded.

HPV	Sensitivity (cp)
16	–
31	162
58	117
62	423
66	123

Table 4.5: HPV sensitivity limit. A logistic regression model was used to determine the sensitivity limits of the MY09/11 method at 95%.

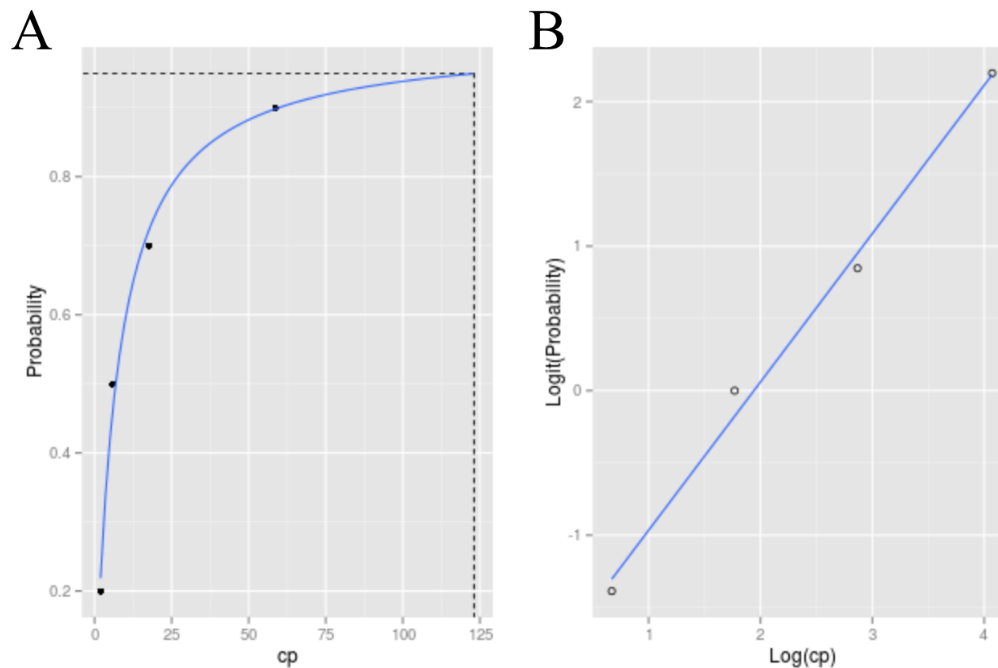


Figure 4.9: Regression curves of the HPV66. (A) Logistic curve of the HPV66 that shows the concentration as cp/μl vs the probability of amplification, using the MY09/11 primers. The sensitivity limit was calculated at the 95% level as 123 cp. (B) The linear regression model for the HPV66 had a R^2 of 0.9928 and a P-value of 0.003605.

CHAPTER 5

DISCUSSION

5.1 HPV SAMPLES

One of the main problems with the samples in the DNA-bank was that many of them were apparently degraded, most probably due to repeated freeze-thaw cycles. Indeed when the samples were re-amplified with the MY09/11 system many of them were reported as negative and new samples were gathered from new patients. Nevertheless, the amplification could have also been hampered by the fact that the amplification conditions for the newly designed type-specific primers had not been optimized completely ($MgCl_2$ concentrations, temperature profile and primer concentrations). Furthermore, variations in the primer binding sites in these less conserved regions cannot be ruled out. However, as the MY09/11 re-amplification had already failed in several of these samples, this is mere speculation.

5.2 TYPE-SPECIFIC PRIMER DESIGN

One of the main goals of this thesis was to design type-specific primers in bulk or large scale. In order to automate this step, a Python script was developed, which through the package `Bio` from Biopython the packages `Entrez` and `SeqIO` allowed sequence files to be imported. `Entrez` is a package that allows the script to connect to NCBI over the World Wide Web, while `SeqIO` is the standard Sequence input/output interface for Biopython that handles different sequence file formats.

A Python script or any other algorithm facilitates greatly the process of retrieving and editing DNA sequences without the struggle to do it by hand and improves the accuracy of the design

because a computer script follows an algorithm and a human being is more prone to make a mistake when being committed to a repetitive task such as retrieve and edit 67 different HPV genomes.

5.3 SENSITIVITY AND DETECTION LIMIT

The MY09/11 sensitivity for the different HPVs was established by PCR with the MY09/11 primers on different dilutions of a type-specific HPV amplicon, which was cloned and isolated by type-specific primers. There are very few reports of the type specific sensitivities of the MY09/11 primers (Table 2.1). Therefore, the main objective of this thesis was to develop a protocol to elucidate the MY09/11 sensitivities and compare them to the scarce information available in the literature (Ting & Manos, 1990; Rohan & Shah, 2006; Chen et al., 2013; Depuydt et al., 2007).

HPV	Results	Literature
31	162 cp	10 cp*
58	117 cp	100 cp*
62	423 cp	NA
66	123 cp	100 cp*

Table 5.1: Comparison of HPV sensitivities. * (Depuydt et al., 2007).

The sensitivity of the MY09/11 system is essential for the validation of the method; to determine it correctly so-called reference materials are required, i.e. a rigorously quantified virus derived DNA. It was obtained by cloning of a relevant part of the viral genomic DNA present in patient samples using type-specific primers. Subsequently, the copy number of the HPV-DNA obtained from a large scale culture was determined by fluorescence spectroscopy with DNA specific dyes.

Sensitivity and amplification efficiency when using the MY09/11 system will be compromised when the degenerated primers do not hybridize correctly on the viral sequences. Especially if the hybridization of the 3' end is affected, the amplification efficiency will drop. However, a "minimal mismatch" may exist within the last 5 to 6 nucleotides at the 3' end of the primer without deleterious effects (Dieffenbach, Lowe, & Dveksler, 1993).

As can be seen in Table 2.1 there is a rather good correlation between the sensitivities reported herein and those documented in literature. It should be noted that the exact procedure applied by Depuydt (2007) has not been well documented and seems to be based on the rather coarse method of endpoint dilution. According to the Poisson distribution this approach may lead to large errors in estimating the sensitivity limit.

The data published by Depuydt has been further analyzed. For example, it should be expected to observe a correlation, although not perfect, between the number of mismatches and the sensitivity limits. Three different plots were constructed; relating the sensitivity and mismatches with the primers MY09, MY11 and both of them as shown in Figures 5.1, 5.2 and 5.3.

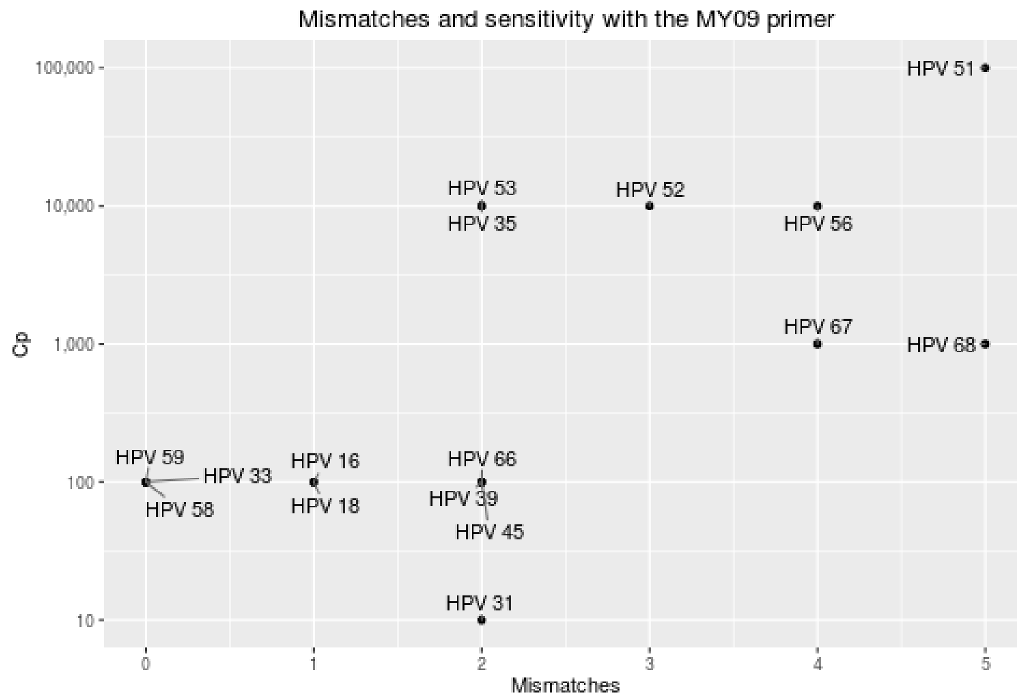


Figure 5.1: Mismatches and sensitivity with the MY09 primer

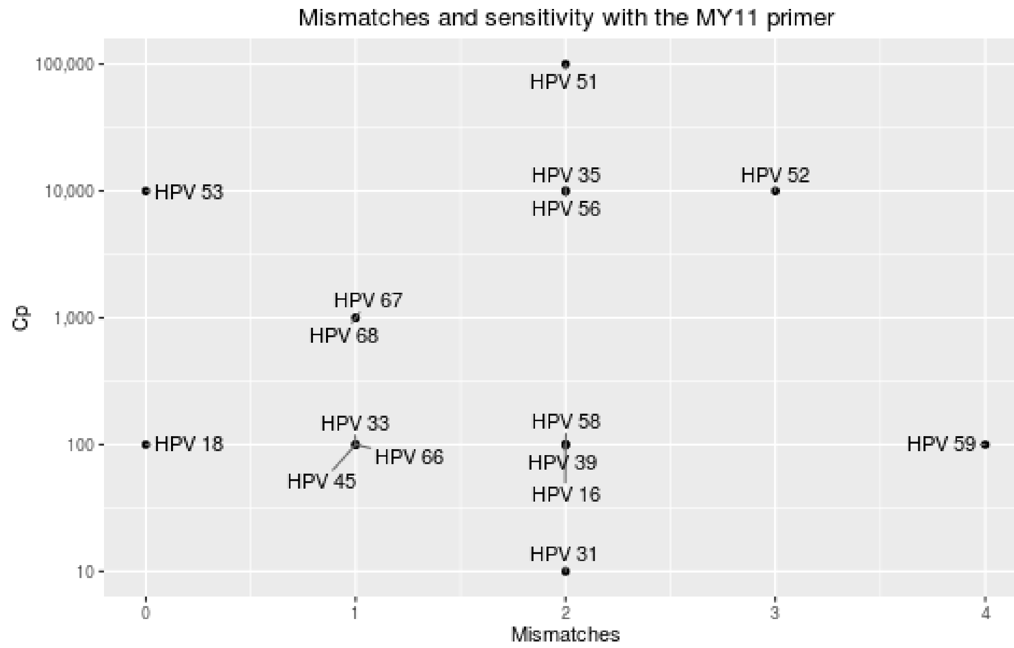


Figure 5.2: Mismatches and sensitivity with the MY11 primer

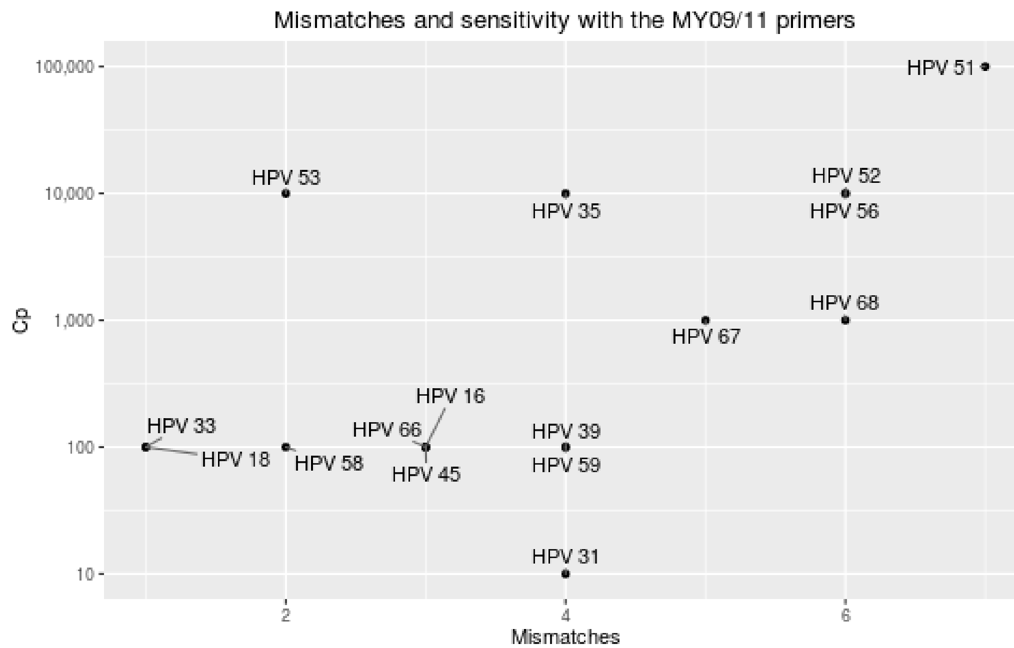


Figure 5.3: Mismatches and sensitivity with the MY09/11 primers

It is apparent that the sensitivity, measured in copies (cp) does not correlate well with the number of mismatches. For example: while HPVs 31, 59, 39 and 35 all have the same total number of mismatches (4) there is a spread of the sensitivities over 3 log units. In addition,

there is no good linear correlation among the mismatches and sensitivity, i.e. the HPV31 with four mismatches has a sensitivity of 10 cp, whereas the HPV53 with just two mismatches has a sensitivity of 10,000 cp.

This simplistic linear model (Fig. 5.4) shows as discussed a bad correlation of the Depuydt's data between the mismatches and the expected sensitivity. However, as stated, it is a rather simplistic model giving each mismatch the same weight, while it is clear that mismatches at the 3' end have a higher impact on the T_m and thus on the amplification efficiency (Kwok et al., 1990).

To continue with the analysis of discrepancies, they are especially evident on different HPV pairs and groups. For example: between the HPV 31 and 35, where the same number of mismatches are present at the 5' end and the central part, the reported sensitivities differ with 3 orders of magnitude (Table 5.2). HPV53 has an apparent sensitivity of 10,000 cp as do the HPVs 35, 52 and 56, but it has only two mismatches located at the 5' end as compared to the 4 mismatches of HPV56 (Table 5.3). And finally the HPV 51 and 68 share the same mismatches on the 3' end and middle with the exception of an extra mismatch on the 5' end of the MY11, however, the sensitives differ surprisingly in two orders of magnitude (Table 5.4).

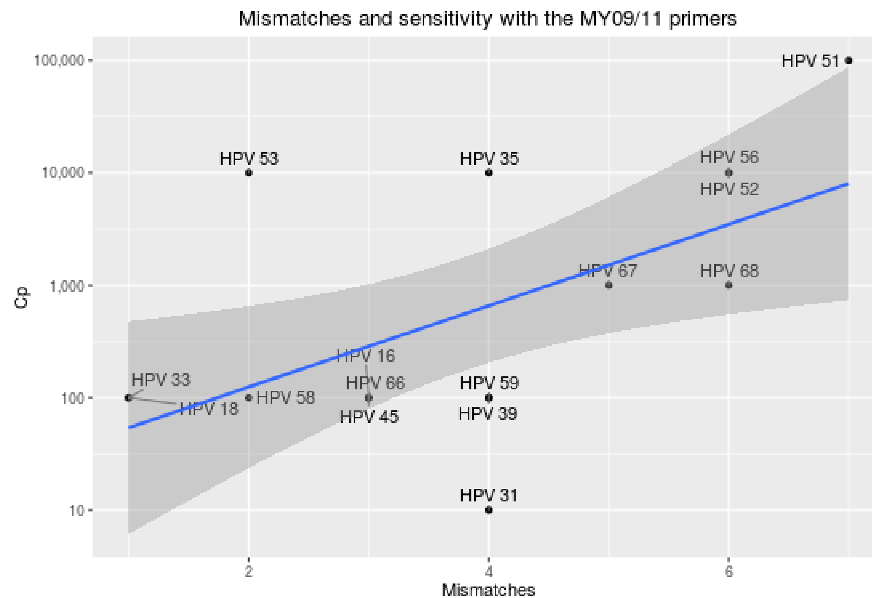


Figure 5.4: Linear model of mismatches and sensitivity with the MY09/11 primers. The linear regression model for the Depuydt's results had a R^2 of 0.2641 and a P-value of 0.04172. * The gray shadow represents the confidence interval at 95%.

Table 5.2: Mismatches on HPV 31 and 35

MY09*	G	A	T	C	A	G	T	W	T	C	C	Y	Y	T	K	G	G	A	C	G
HPV-031	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	T	-	-
HPV-035	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	C	-	-

MY11	G	C	M	C	A	G	G	G	W	C	A	T	A	A	Y	A	A	T	G	G
HPV-031	-	-	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-
HPV-035	-	-	-	-	-	A	-	-	C	-	-	-	-	-	-	-	-	-	-	-

* Complement Sequence (5’ → 3’)

Table 5.3: Mismatches on HPV 35, 52, 53 and 56

MY09*	G	A	T	C	A	G	T	W	T	C	C	Y	Y	T	K	G	G	A	C	G
HPV-035	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	C	-	-
HPV-052	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	T	A	-
HPV-053	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	A	-
HPV-056	-	-	-	-	-	A	-	-	-	-	-	A	-	-	-	-	-	T	A	-

MY11	G	C	M	C	A	G	G	G	W	C	A	T	A	A	Y	A	A	T	G	G
HPV-035	-	-	-	-	-	A	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-052	-	-	G	-	-	-	-	-	C	-	-	C	-	-	-	-	-	-	-	-
HPV-053	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-056	-	-	-	-	-	A	-	-	C	-	-	-	-	-	-	-	-	-	-	-

* Complement Sequence (5’ → 3’)

Table 5.4: Mismatches on HPV 51 and 68

MY09*	G	A	T	C	A	G	T	W	T	C	C	Y	Y	T	K	G	G	A	C	G
HPV-051	-	-	C	-	-	A	-	-	-	G	-	A	-	-	-	-	-	T	-	-
HPV-068	-	-	C	-	-	A	-	-	C	-	-	A	-	-	A	-	-	-	-	-

MY11	G	C	M	C	A	G	G	G	W	C	A	T	A	A	Y	A	A	T	G	G
HPV-051	-	-	G	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-
HPV-068	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-

* Complement Sequence (5’ → 3’)

The effects of the mismatch positions on amplification efficiency have been analyzed in various articles (Dieffenbach et al., 1993; Kwok et al., 1990; Chou, Russell, Birch, Raymond, &

Bloch, 1992; Ledeker, 2012; Boyle, Dallaire, & MacKay, 2009). Mismatches located in the 3' end and the central parts of the primers affect significantly primer binding and amplification efficiency. However, recently it has been documented that mismatches on the 5' end and combinations of mismatches between forward and reverse primers may affect the efficiency of PCR amplifications in higher proportions as previously thought (Ledeker, 2012).

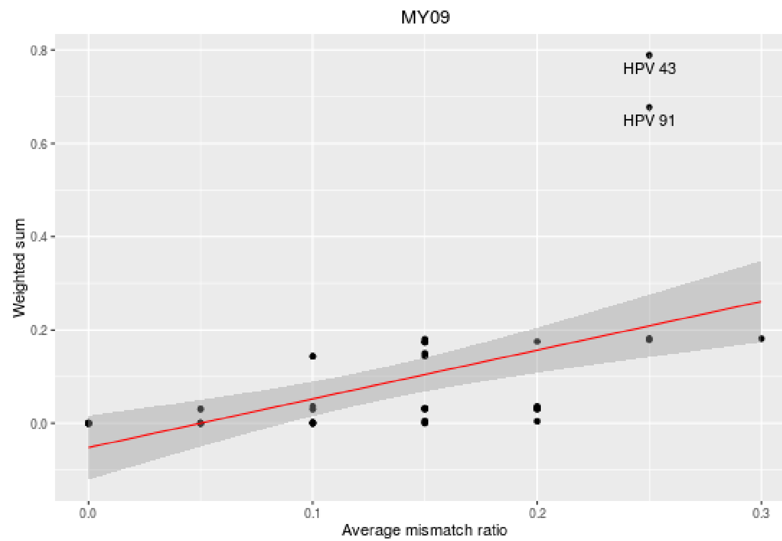


Figure 5.5: Weighted logarithmic model of MY09. The linear regression model had a R^2 of 0.2845 and a P-value of 3.913×10^{-5}. * The gray shadow represents the confidence interval at 95%.

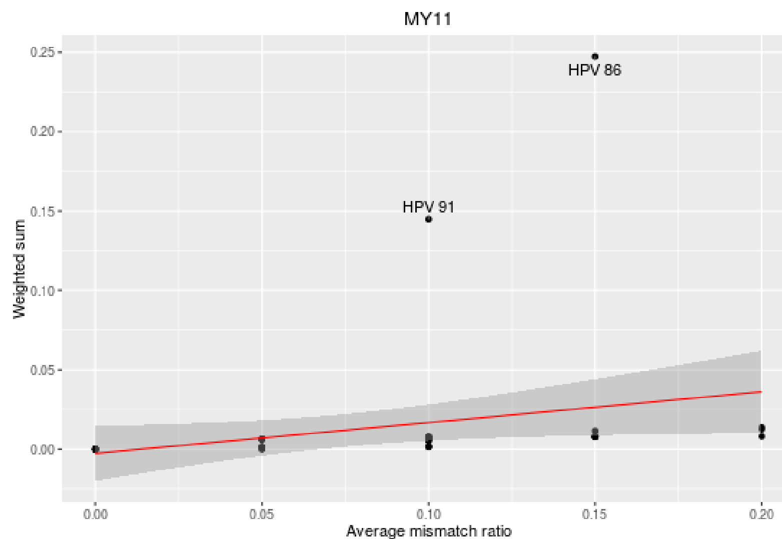


Figure 5.6: Weighted logarithmic model of MY11. The linear regression model had a R^2 of 0.07741 and a P-value of 0.04367. * The gray shadow represents the confidence interval at 95%.

The 3'-terminal position controls the mispriming with the DNA template (Kwok et al., 1990) and within primer pairs. The last one is an undesired phenomenon that results in amplification of the so called primer-dimers thereby lowering the efficiency of the PCR by competing for the DNA polymerase (Chou et al., 1992). In order to improve the relationship of mismatches at different positions and the amplification efficiency we introduced differential weights incrementing exponentially towards the 3' ends of the MY09/11 system. See Figures 5.5 and 5.6. However, again there was no good correlation observable.

The MY11 primer is better designed on the 3' end with a lower weighted sum score (Fig. 5.6), which represents fewer mismatches on the 3' end, whereas the MY09 primer has significantly higher values (Fig. 5.5). For easy interpretation, the higher the score for the weighted sum and the average mismatch ratio the less likely is the amplification of an HPV through the MY09/11 system, but more importantly, HPVs with significant mismatches on 3' end of the MY09 primer affect more the sensitivity and may be correlated with the expected sensitivity for the number of mismatches.

The sensitivity of MY09/11 for the HPV31 was significantly worse than the one reported in the literature (Depuydt et al., 2007).

Interestingly, there are strains of HPV31 which show various numbers of mismatches with regards to the MY primer sequences. The discordant results could thus be explained if by chance the reference materials used in both studies were different. However, further experiments are needed to analyze this hypothesis.

On the other hand, the HPVs 58 and 66 sensitivities correspond or are similar to the ones that are found in the literature, whereas for the HPV62 there is no information available. The match in sensitivities may be due to less mismatches between the MY09/11 system and our reference material of the HPVs 58, 62 and 66 with 1, 4, and 3 mismatches, respectively. Wherein, the mismatches are present on the 5' end with the exception of of the HPV62 with one mismatch on the 3' end (Table A.5 and A.6).

As previously discussed, the data consistency of Depuydt can be challenged and seems to be rather limited with respect to precision. In addition, as in our study, it relied mainly on patient derived reference materials. As discussed above for HPV31, there exist variants of each virus

type infecting the global human population. The reference sequences as defined by GenBank are based mainly on the first published sequence of a given type (Table A.3) irrespective of its worldwide frequency. In order to document the variability within a viral type was calculated as the fraction of unique sequences among different isolates of the same HPV type (Table 5.5). The unique sequences are listed in the Tables A.7 and A.8 with its respective count of appearance. The complete list of the HPV variants within its type is shown in Table A.9.

Table 5.5: Unique HPV sequences among different isolates

HPV	MY09		MY11	
HPV016	1.34%	(4/298)	1.34%	(4/298)
HPV018	5.26%	(4/76)	1.32%	(1/76)
HPV031	4.65%	(2/43)	4.65%	(2/43)
HPV033	2.70%	(1/37)	5.41%	(2/37)
HPV035	7.69%	(4/52)	1.92%	(1/52)
HPV039	13.33%	(4/30)	10.00%	(3/30)
HPV045	2.70%	(1/37)	2.70%	(1/37)
HPV051	3.70%	(2/54)	3.70%	(2/54)
HPV052	1.06%	(2/188)	0.53%	(1/188)
HPV053	11.54%	(3/26)	7.69%	(2/26)
HPV056	36.36%	(4/11)	27.27%	(3/11)
HPV058	0.52%	(1/193)	1.04%	(2/193)
HPV059	22.22%	(4/18)	5.56%	(1/18)
HPV066	26.67%	(4/15)	6.67%	(1/15)
HPV067	11.11%	(1/9)	11.11%	(1/9)
HPV068	18.52%	(5/27)	18.52%	(5/27)

Finally, as a secondary result of the performed sequence analyses I propose to change some degenerated bases to increase the sensitivity of the MY09/11 system. MY09 should be changed to (GATCA**R**TWTCCYYT**D**GG**WM**G) and MY11 to (GCMCA**R**GG**H**CA**Y**AAYAATGG).

These changes would result in 52.63% and 66.67% less mismatches, respectively among different variants of the same HPV type (Table A.9). Of course these proposed modifications have to be carefully analyzed whether they do not increase self- and cross- complementarity within the primer pairs, which are the trigger for the formation of primer-dimers during the amplification. Excessive primer-dimer formation reduces significantly the sensitivity. Anyway, whether the MY09/11 system is modified or not, it is necessary to determine the sensitivities

for all the HPVs. However, the sensitivities are bound to the analyzed reference sequences. Considering the variability within each HPV and the particular pattern of differences it is quite obvious that the sensitivities may vary significantly from one isolate to another within a given HPV type. Sensitivities will thus be just a small window into the real distribution of sensitivities.

CHAPTER 6

CONCLUSION

This study was set to determine the sensitivity of the MY09/11 detection system through the design of a protocol that is able to elucidate the sensitivities of different HPVs. The literature on this subject is scarce and inconclusive due to the limited quality of the information for the discussed reasons. The main question to answer in this thesis was:

1. May the sensitivity of the MY09/11 detection system be determined for each HPV type?

Different methods or systems are available to detect HPV based mainly on degenerate or type specific primers. The MY09/11 amplify a broad spectrum of HPV and they are widely used, hence the sensitivity is an important piece of information. The sensitivities of the HPV 31, 58, 62 and 66 by the MY09/11 detection system were determined as a proof-of-concept, however due to limitations in time it has not been possible to determine the sensitivities of other HPVs that had been successfully amplified by the MY09/11 system. Nonetheless this thesis contains the protocol and information to continue with the search of the sensitivities of all the others HPVs, the type-specific primers are shown in Table A.4.

The sensitivity of the MY09/11 system may help as an auxiliary parameter to interpret a negative result. Not surprisingly the laboratories are required to know the sensitivities of their methods for the validation and certification of their tests, therefore the sensitivities for each or at least the most important HPVs is a requirement that should be addressed in every clinical laboratory to ensure the quality of their results and to generate values that may serve doctors to interpret their results.

It cannot be stressed enough that the sensitivities as determined by this protocol may be limited, since within the types there is some variability and if this variability affects the binding sites of the primers then the sensitivities will be affected. It would therefore be useful not to only use a reference sequence but a small sample of sequences and determine in each of them the sensitivity (intra-type), to determine how much varies the sensitivity of an isolate to another and verify if they are equal, different or the difference between each other, since the variations may be large.

It is possible to improve the results presented in this thesis by considering different sequences of the same HPV and analyze the differences in the sensitivities within the HPV group of the same type. Moreover, using more than one reference material as sample to ensure the diversity of the different HPVs that are present within a certain population, in our case a fraction of the patients from the state of Puebla, Mexico. Therefore, by combining the two strategies the interpretation of negative results will improve significantly.

REFERENCES

Antonsson, A., Erfurt, C., Hazard, K., Holmgren, V., Simon, M., Kataoka, A., ... Hansson, B. G. (2003). Prevalence and type spectrum of human papillomaviruses in healthy skin samples collected in three continents. *Journal of General Virology*, *84*(7), 1881–1886.

Applied Biosystems. (2003). *Real-Time PCR Vs. Traditional PCR*. Retrieved from http://www6.appliedbiosystems.com/support/tutorials/pdf/rtpcr_vs_tradpcr.pdf

Badillo-Almaraz, I., Zapata-Benavides, P., Saavedra-Alonso, S., Zamora-Avila, D., Reséndez-Pérez, D., Tamez-Guerra, R., ... Rodríguez-Padilla, C. (2012). Human papillomavirus 16/18 infections in lung cancer patients in Mexico. *Intervirology*, *56*(5), 310–315.

Baldur-Felskov, B., Dehlendorff, C., Junge, J., Munk, C., & Kjaer, S. K. (2014). Incidence of cervical lesions in danish women before and after implementation of a national HPV vaccination program. *Cancer Causes & Control*, *25*(7), 915–922.

Bernard, H.-U. (2005). The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses. *Journal of clinical virology*, *32*, 1–6.

Bernard, H.-U., Burk, R. D., Chen, Z., van Doorslaer, K., zur Hausen, H., & de Villiers, E.-M. (2010). Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology*, *401*(1), 70–79.

Bosch, F., Lorincz, A., Munoz, N., Meijer, C., & Shah, K. (2002). The causal relation between human papillomavirus and cervical cancer. *Journal of clinical pathology*, *55*(4), 244–265.

Boyle, B., Dallaire, N., & MacKay, J. (2009). Evaluation of the impact of single nucleotide polymorphisms and primer mismatches on quantitative PCR. *BMC Biotechnology*, *9*(1), 75.

Brink, A. A., Snijders, P. J., & Meijer, C. J. (2007). HPV detection methods. *Disease markers*, *23*(4), 273–281.

Bruni, L., Barrionuevo-Rosas, L., Albero, G., Aldea, M., Serrano, B., Valencia, S., ... Castellsagué, X. (2015). *Human Papillomavirus and Related Diseases Report*. ICO Information Centre on HPV and Cancer (HPV Information Centre).

Burchell, A. N., Winer, R. L., de Sanjosé, S., & Franco, E. L. (2006). Epidemiology and transmission dynamics of genital HPV infection. *Vaccine*, *24*, S52–S61.

Burd, E. M. (2003). Human Papillomavirus and Cervical Cancer. *Clinical Microbiology Re-*

views, *16*(1), 1–17.

Calleja-Macias, I. E., Kalantari, M., Huh, J., Ortiz-Lopez, R., Rojas-Martinez, A., Gonzalez-Guerrero, J. F., ... others (2004). Genomic diversity of human papillomavirus-16, 18, 31, and 35 isolates in a mexican population and relationship to european, african, and native american variants. *Virology*, *319*(2), 315–323.

Chen, L., Watanabe, K., Haruyama, T., & Kobayashi, N. (2013). Simple and rapid human papillomavirus genotyping method by restriction fragment length polymorphism analysis with two restriction enzymes. *Journal of Medical Virology*, *85*(7), 1229–1234.

Chou, Q., Russell, M., Birch, D. E., Raymond, J., & Bloch, W. (1992). Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nucleic acids research*, *20*(7), 1717–1723.

Cobo, F. (2012). *Diagnosis of HPV infection.* Elsevier.

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., ... de Hoon, M. J. L. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics (Oxford, England)*, *25*(11), 1422–1423.

Coutlée, F., Gravitt, P., Kornegay, J., Hankins, C., Richardson, H., Lapointe, N., ... others (2002). Use of pgmy primers in l1 consensus pcr improves detection of human papillomavirus dna in genital samples. *Journal of clinical microbiology*, *40*(3), 902–907.

Currin, L. G., Jack, R. H., Linklater, K. M., Mak, V., Møller, H., & Davies, E. A. (2009). Inequalities in the incidence of cervical cancer in South East England 2001–2005: an investigation of population risk factors. *BMC Public Health*, *9*(1), 62.

Cutts, F., Franceschi, S., Goldie, S., Castellsague, X., De Sanjose, S., Garnett, G., ... others (2007). Human papillomavirus and hpv vaccines: a review. *Bulletin of the World Health Organization*, *85*(9), 719–726.

Danos, O., Katinka, M., & Yaniv, M. (1982). Human papillomavirus 1a complete DNA sequence: a novel type of genome organization among papovaviridae. *The EMBO journal*, *1*(2), 231–236.

Depuydt, C. E., Boulet, G. A. V., Horvath, C. A. J., Benoy, I. H., Vereecken, A. J., & Bogers, J. J. (2007). Comparison of MY09/11 consensus PCR and type-specific PCRs in the detection of oncogenic HPV types. *Journal of Cellular and Molecular Medicine*, *11*(4), 881–891.

De Villiers, E.-M., Fauquet, C., Broker, T. R., Bernard, H.-U., & zur Hausen, H. (2004). Classification of papillomaviruses. *Virology*, *324*(1), 17–27.

Dieffenbach, C. W., Lowe, T. M., & Dveksler, G. S. (1993). General concepts for PCR primer design. *Genome Research*, *3*(3), S30–S37.

Doorbar, J. (2005). The papillomavirus life cycle. *Journal of clinical virology*, *32*, 7–15.

Doorbar, J. (2006). Molecular biology of human papillomavirus infection and cervical cancer. *Clinical science*, *110*, 525–541.

Eckert, R. L., & Rorke, E. A. (1989). Molecular biology of keratinocyte differentiation. *Environmental health perspectives*, *80*, 109–116.

Eide, M. L., & Debaque, H. (2012). HPV detection methods and genotyping techniques in screening for cervical cancer. *Annales de pathologie*, *32*(6), e15–23– 401–9.

Fareed, G. C., & Davoli, D. (1977). Molecular biology of papovaviruses. *Annual review of biochemistry*, *46*(1), 471–522.

Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., & Ball, L. A. (2005). *Virus taxonomy: VIIIth report of the international committee on taxonomy of viruses*. Academic Press.

Flores-Munguia, R., Siegel, E., Klimecki, W. T., & Giuliano, A. R. (2004). Performance assessment of eight high-throughput PCR assays for viral load quantitation of oncogenic HPV types. *The Journal of Molecular Diagnostics*, *6*(2), 115–124.

Freeman, W. M., Walker, S. J., & Vrana, K. E. (1999). Quantitative RT-PCR: pitfalls and potential. *Biotechniques*, *26*, 112–125.

Fuchsmann, C., Ayari-Khalfallah, S., Coulombeau, B., & Froehlich, P. (2011). Papilomatosis laríngea. *EMC-Otorrinolaringología*, *40*(4), 1–8.

Gál, Á. B., Carnwath, J. W., Dinnyes, A., Herrmann, D., Niemann, H., & Wrenzycki, C. (2006). Comparison of real-time polymerase chain reaction and end-point polymerase chain reaction for the analysis of gene expression in preimplantation embryos. *Reproduction, Fertility and Development*, *18*(3), 365–371.

Gentleman, R. (2008). *R Programming for Bioinformatics*. CRC Press.

Gravitt, P., Peyton, C., Alessi, T., Wheeler, C., Coutlee, F., Hildesheim, A., … Apple, R. (2000). Improved amplification of genital human papillomaviruses. *Journal of clinical microbiology*, *38*(1), 357–361.

Hazard, K., Karlsson, A., Andersson, K., Ekberg, H., Dillner, J., & Forslund, O. (2007). Cutaneous human papillomaviruses persist on healthy skin. *Journal of Investigative Dermatology*, *127*(1), 116–119.

Hernandez-Suarez, G., Pineros, M., Vargas, J., Orjuela, L., Hernandez, F., Peroza, C., … Perez, G. (2013). Human papillomavirus genotypes in genital warts in Latin America: a cross-sectional study in Bogota, Colombia. *International journal of STD & AIDS*, *24*(7), 567–572.

Hidalgo-Martínez, A. C. (2006). El cáncer cérvico-uterino, su impacto en méxico y el porqué no funciona el programa nacional de detección oportuna. *Rev Biomed*, *17*, 81–84.

Ho, L., Chan, S., Burk, R., Das, B., Fujinaga, K., Icenogle, J., … Mavromara-Nazos, P. (1993). The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *Journal of Virology*, *67*(11), 6413–6423.

Jemal, A., Center, M. M., DeSantis, C., & Ward, E. M. (2010). Global patterns of cancer incidence and mortality rates and trends. *Cancer Epidemiology Biomarkers & Prevention*, *19*(8), 1893–1907.

Juarez-Figueroa, L. A., Wheeler, C. M., Uribe-Salas, F. J., Conde-Glez, C. J., Zamilpa-Mejia, L. G., Garcia-Cisneros, S., & Hernandez-Avila, M. (2001). Human papillomavirus: a highly prevalent sexually transmitted disease agent among female sex workers from Mexico City. *Sexually transmitted diseases*, *28*(3), 125–130.

Kernighan, B. W., & Ritchie, D. M. (1988). *The c programming language* (Vol. 2). Prentice-Hall Englewood Cliffs.

Kleinbaum, D. G., & Klein, M. (2010). *Logistic regression: a self-learning text*. Springer Science & Business Media.

Koutsky, L. (2009, April). The Epidemiology behind the HPV Vaccine Discovery. *Annals of Epidemiology*, *19*(4), 239–244.

Kumate, J. (2010). *Infectología Clínica* (17a. ed.). Méndez Editores.

Kwok, S., Kellogg, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C., & Sninsky, J. J. (1990). Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. *Nucleic acids research*, *18*(4), 999–1005.

Laboratory of Tumor Virus Biology. (1986). *Papilloma Virus (HPV).* Retrieved 2016-02-14, from https://visualsonline.cancer.gov/details.cfm?imageid=2255

Ledeker, B. M. (2012). *Determining the Effect of Primer Mismatches on Quantitative PCR Accuracy and Developing Guidance for Design of Primers Targeting Genes with Sequence Variations* (Master's thesis). Colorado State University.

Lewin, B., Krebs, J., Kilpatrick, S. T., & Goldstein, E. S. (2011). *Lewin's genes X* (Vol. 10). Jones & Bartlett Learning.

Li, J., Cai, H., Xu, Z., Wang, Q., Hang, D., Shen, N., ... Ke, Y. (2012). Nine complete genome sequences of cutaneous human papillomavirus genotypes isolated from healthy skin of individuals living in rural He Nan Province, China. *Journal of virology*, *86*(21), 11936–11936.

Longo, D., Fauci, A., Kasper, D., Hauser, S., Jameson, J., & Loscalzo, J. (2011). *Harrison's principles of internal medicine: Volumes 1 and 2* (18th ed.). McGraw-Hill Professional.

Matloff, N. (2011). *The art of r programming: A tour of statistical software design*. No Starch Press.

Maurer, J. (2006). *PCR methods in foods*. Springer Science & Business Media.

Mistry, N. (2007). *Human papillomavirus tropism: determinants of viral tissue specificity.* Klinisk mikrobiologi.

Mohabatkar, H. (2007). Prediction of epitopes and structural properties of iranian HPV-16 E6 by bioinformatics methods. *Asian Pac. J. Cancer Prev*, *8*(4), 602–606.

Münger, K., & Howley, P. M. (2002). Human papillomavirus immortalization and transformation functions. *Virus research*, *89*(2), 213–228.

Orth, G., Favre, M., & Croissant, O. (1977). Characterization of a new type of human papillomavirus that causes skin warts. *Journal of virology*, *24*(1), 108–120.

Palsson, B. O. (1997). Bioinformatics: What lies beyond bioinformatics? *Nature Biotechnology*, *15*(1), 3–4.

Qu, W., Jiang, G., Cruz, Y., Chang, C. J., Ho, G., Klein, R. S., & Burk, R. D. (1997). Pcr detection of human papillomavirus: comparison between my09/my11 and gp5+/gp6+ primer systems. *Journal of clinical microbiology*, *35*(6), 1304–1310.

R Development Core Team. (2011). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org`

Renshaw, A. A., DiNisco, S. A., Minter, L. J., & Cibas, E. S. (1997). A more accurate measure of the false-negative rate of papanicolaou smear screening is obtained by determining the false-negative rate of the rescreening process. *Cancer Cytopathology*, *81*(5), 272–276.

Rodrigo, A. G., Goracke, P. C., Rowhanian, K., & Mullins, J. I. (1997). Quantitation of target molecules from polymerase chain reaction-based limiting dilution assays. *AIDS research and human retroviruses*, *13*(9), 737–742.

Rohan, T. E., & Shah, K. V. (2006). *Cervical cancer: from etiology to prevention* (Vol. 2). Springer Science & Business Media.

Rojas, R., Göktekin, C., Friedland, G., Krüger, M., Langmack, O., & Kuniß, D. (2000). *Plankalkül: the first high level programming language and its implementation*. Citeseer.

Roulston, J. E., & Bartlett, J. M. S. (Eds.). (2004). *Molecular Diagnosis of Cancer: Methods and Protocols* (2nd edition ed.). Totowa, N.J: Humana Press.

Ruiz, F. E., Pérez, P. S., & Bonev, B. I. (2009). *Information theory in computer vision and pattern recognition*. Springer Science & Business Media.

Rychlik, W. (2009). *Oligo 7 Tutorial & Examples.* Retrieved from `http://www.oligo.net/oligo_7_tutorial.pdf`

Saeys, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics*, *23*(19), 2507–2517.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning* (Vol. 2). Cold spring harbor laboratory press New York.

Schiffman, M., Castle, P. E., Jeronimo, J., Rodriguez, A. C., & Wacholder, S. (2007). Human papillomavirus and cervical cancer. *The Lancet*, *370*(9590), 890–907.

Schiller, J. T., Day, P. M., & Kines, R. C. (2010). Current understanding of the mechanism of hpv infection. *Gynecologic oncology*, *118*(1), S12–S17.

Secretaría de Salud. (2012). *Panorama Epidemiológico y Estadístico de la Mortalidad*

en México 2010. Retrieved from http://www.epidemiologia.salud.gob.mx/dgae/infoepid/publicaciones2012.html

Singh, G. K., Azuine, R. E., & Siahpush, M. (2012). Global inequalities in cervical cancer incidence and mortality are linked to deprivation, low socioeconomic status, and human development. *Int J MCH AIDS*, *1*(1), 17–30.

Sirotković-Skerlev, M., & Zekan, M. S. J. (2011). *Oncogenic aspects of hpv infections of the female genital tract*. InTech.

Stanley, M. (1994). *Immunology of human papillomaviruses*. Springer US.

Stanley, M. (2006). Immune responses to human papillomavirus. *Vaccine*, *24*, S16–S22.

Stanley, M. (2014). Hpv vaccination in boys and men. *Human vaccines & immunotherapeutics*, *10*(7), 2106–2108.

Syrjänen, K., Gissmann, L., & Koss, L. G. (1987). *Papillomaviruses and human disease*. Springer Berlin Heidelberg.

Taswell, C. (1981). Limiting dilution assays for the determination of immunocompetent cell frequencies. I. Data analysis. *The Journal of Immunology*, *126*(4), 1614–1619.

Thermo Fisher Scientific. (2006). *Quant-iT™ Assays for high-throughput quantitation of DNA, RNA, and oligos.* Retrieved 2016-03-25, from https://www.thermofisher.com/order/catalog/product/P11496

Thompson, B., Vilchis, H., Moran, C., Copeland, W., Holte, S., & Duggan, C. (2014). Increasing cervical cancer screening in the United States-Mexico border region. *The Journal of Rural Health*, *30*(2), 196–205.

Ting, Y., & Manos, M. M. (1990). Detection and typing of genital human papillomaviruses. *PCR protocols: a guide to methods and applications*, 356–367.

Tisdall, J. (2001). *Beginning Perl for Bioinformatics*. O'Reilly Media, Inc.

Uthman, E. (2006). *HPV/LSIL On Pap Smear.* Retrieved 2016-03-11, from http://flickr.com/photos/euthman/194024495/

WHO. (2012). *Fact Sheets by Cancer.* Retrieved from http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx

WHO. (2014). *Comprehensive cervical cancer control: a guide to essential practice.* World Health Organization.

Woodman, C. B. J., Collins, S. I., & Young, L. S. (2007). The natural history of cervical HPV infection: unresolved issues. *Nature reviews. Cancer*, *7*(1), 11–22.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics*, *13*(1), 1–11.

Yim, E.-K., & Park, J.-S. (2005). The role of HPV E6 and E7 oncoproteins in HPV-associated cervical carcinogenesis. *Cancer research and treatment: official journal of Korean Can-*

cer Association, *37*(6), 319.

Zur Hausen, H. (1996). Papillomavirus infections—a major cause of human cancers. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, *1288*(2), F55–F78.

APPENDIX A

TABLES

Table A.1: Primer MY09: comparison between the different HPVs to HPV16.

MY09•	G	A	T	C	A	G	T	W	T	C	C	Y	Y	T	K	G	G	A	C	G
HPV-016*	G	A	T	C	A	G	T	T	T	C	C	T	T	T	A	G	G	A	C	G
HPV-002	-	-	C	-	-	A	-	-	-	-	-	C	-	-	G	-	-	T	-	-
HPV-006	-	-	-	-	-	-	-	A	-	-	-	-	-	-	G	-	-	-	-	-
HPV-011	-	-	-	-	-	-	-	-	-	-	-	C	C	-	T	-	-	-	-	-
HPV-013	-	-	-	-	-	-	-	A	-	-	-	C	C	-	T	-	-	C	A	-
HPV-018	-	-	-	-	-	A	-	A	-	-	-	C	C	-	T	-	-	-	-	-
HPV-026	-	-	-	-	-	A	-	-	-	-	-	A	C	-	-	-	-	G	-	-
HPV-027	-	-	C	-	-	A	-	-	-	-	-	-	-	-	G	-	-	T	-	-
HPV-030	-	-	C	-	-	A	-	-	C	-	-	A	C	-	G	-	-	C	A	-
HPV-031	-	-	-	-	-	-	-	-	-	-	-	A	C	-	G	-	-	T	-	-
HPV-032	-	-	-	-	-	-	-	-	-	-	-	A	-	-	G	-	-	T	A	-
HPV-033	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-
HPV-034	-	-	-	-	-	-	-	-	-	G	-	C	-	-	-	-	-	T	A	-
HPV-035	-	-	-	-	-	A	-	-	-	-	-	G	-	-	G	-	-	C	-	-
HPV-039	-	-	-	-	-	A	-	-	C	-	-	-	-	-	G	-	-	-	-	-
HPV-040	-	-	-	-	-	A	-	-	-	-	-	A	-	-	-	-	-	-	-	-
HPV-042	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	T	A	-
HPV-043	A	C	C	-	-	-	-	-	-	-	-	C	-	-	-	-	-	G	-	-
HPV-044	-	-	-	-	-	A	-	A	-	-	-	C	C	-	T	-	-	T	A	-
HPV-045	-	-	-	-	-	A	-	A	-	-	-	C	C	-	T	-	-	T	-	-
HPV-051	-	-	C	-	-	A	-	-	-	G	-	A	-	-	G	-	-	T	-	-
HPV-052	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	A	-
HPV-053	-	-	-	-	-	-	-	-	-	-	-	-	C	-	T	-	-	C	A	-
HPV-054	-	-	C	-	-	-	-	-	-	-	-	C	-	-	G	-	-	T	-	-

MY09•	G	A	T	C	A	G	T	W	T	C	C	Y	Y	T	K	G	G	A	C	G
HPV-016*	G	A	T	C	A	G	T	T	T	C	C	T	T	T	A	G	G	A	C	G
HPV-056	-	-	-	-	-	A	-	-	-	-	-	A	C	-	G	-	-	T	A	-
HPV-057	-	-	C	-	-	A	-	-	C	-	-	C	C	-	G	-	-	-	-	-
HPV-058	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-
HPV-059	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-
HPV-061	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	G	-	-
HPV-062	-	-	C	-	-	-	-	-	-	-	-	-	-	-	G	-	-	G	-	-
HPV-066	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	T	A	-
HPV-067	-	-	-	-	-	A	-	-	-	-	-	C	-	-	-	-	-	T	A	-
HPV-068	-	-	C	-	-	A	-	-	C	-	-	A	-	-	-	-	-	-	-	-
HPV-069	-	-	-	-	-	A	-	-	-	-	-	A	C	-	G	-	-	C	A	-
HPV-070	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	G	-	-
HPV-071	-	-	-	-	-	-	-	-	-	-	-	A	-	-	G	-	-	G	-	-
HPV-072	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
HPV-073	-	-	C	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	A	-
HPV-074	-	-	-	-	-	A	-	A	C	-	-	C	C	-	T	-	-	T	A	-
HPV-081	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	T	-	-
HPV-082	-	-	-	-	-	-	-	-	-	G	-	A	-	-	G	-	-	T	-	-
HPV-083	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	C	-	-
HPV-084	-	-	-	-	-	-	-	A	-	-	-	C	-	-	G	-	-	T	-	-
HPV-085	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-086	-	-	C	-	-	A	-	-	-	-	-	A	C	-	G	-	-	C	-	-
HPV-087	-	-	C	-	-	A	-	-	-	-	-	-	-	-	G	-	-	C	-	-
HPV-089	-	-	C	-	-	-	-	-	C	-	-	-	-	-	G	-	-	T	-	-
HPV-090	-	-	C	-	-	A	-	-	C	-	-	-	C	-	T	-	-	C	A	-
HPV-091	A	C	-	-	-	A	-	-	-	-	-	A	-	-	-	-	-	-	-	-
HPV-097	-	-	-	-	-	A	-	A	-	-	-	-	C	-	T	-	-	-	-	-
HPV-102	-	-	-	-	-	A	-	-	-	-	-	-	-	-	G	-	-	G	-	-
HPV-106	-	-	C	-	-	-	-	-	-	-	-	C	-	-	G	-	-	T	A	-
HPV-114	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	C	-	-

* Denotes the HPV16, which was used as a reference to compare the other HPVs. “-” represents no difference to the reference and a letter represents the change in that particular site.

• Complement Sequence (5’ $\rightarrow$ 3’)

Table A.2: Primer MY11: comparison between the different HPVs to HPV16.

MY11	G	C	M	C	A	G	G	G	W	C	A	T	A	A	Y	A	A	T	G	G
HPV-016*	G	C	A	C	A	G	G	G	C	C	A	C	A	A	T	A	A	T	G	G
HPV-002	-	-	C	-	-	A	-	-	G	-	-	-	-	-	C	-	-	-	-	-
HPV-006	-	-	C	-	-	-	-	-	A	-	-	T	-	-	C	-	-	-	-	-
HPV-011	-	-	T	-	-	-	-	-	A	-	-	T	-	-	C	-	-	-	-	-
HPV-013	-	-	C	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-
HPV-018	-	-	-	-	-	-	-	-	T	-	-	T	-	-	C	-	-	-	-	-
HPV-026	-	-	-	-	-	-	-	-	T	-	-	T	-	-	-	-	-	-	-	-
HPV-027	-	-	C	-	-	-	-	-	A	-	-	T	-	-	-	-	-	-	-	-
HPV-030	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-
HPV-031	-	-	T	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-
HPV-032	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-033	-	-	-	-	-	A	-	-	T	-	-	T	-	-	-	-	-	-	-	-
HPV-034	-	-	C	-	-	-	-	-	A	-	-	A	-	-	C	-	-	-	-	-
HPV-035	-	-	-	-	-	A	-	-	-	-	-	T	-	-	-	-	-	-	-	-
HPV-039	-	-	C	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-
HPV-040	-	-	C	-	-	-	-	-	-	-	-	T	-	-	C	-	-	-	-	-
HPV-042	-	-	-	-	-	A	-	-	A	-	-	-	-	-	-	-	-	-	-	-
HPV-043	-	-	C	-	-	-	-	-	A	-	-	T	-	-	-	-	-	-	-	-
HPV-044	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-045	-	-	C	-	-	-	-	-	-	-	-	T	-	-	C	-	-	-	-	-
HPV-051	-	-	G	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-
HPV-052	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-053	-	-	C	-	-	-	-	-	A	-	-	T	-	-	-	-	-	-	-	-
HPV-054	-	-	C	-	-	-	-	-	T	-	-	A	-	-	C	-	-	-	-	-
HPV-056	-	-	C	-	-	A	-	-	-	-	-	T	-	-	-	-	-	-	-	-
HPV-057	-	-	C	-	-	-	-	-	A	-	-	T	-	-	C	-	-	-	-	-
HPV-058	-	-	-	-	-	A	-	-	T	-	-	T	-	-	C	-	-	-	-	-
HPV-059	-	-	T	-	-	-	-	-	T	T	T	A	-	-	C	-	-	-	-	-
HPV-061	-	-	C	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-
HPV-062	-	-	G	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-
HPV-066	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-
HPV-067	-	-	-	-	-	A	-	-	T	-	-	T	-	-	C	-	-	-	-	-
HPV-068	-	-	-	-	-	-	-	-	A	-	-	-	-	-	C	-	-	-	-	-
HPV-069	-	-	C	-	-	-	-	-	T	-	-	T	-	-	-	-	-	-	-	-
HPV-070	-	-	C	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-
HPV-071	-	-	-	-	-	-	-	-	-	A	C	A	-	-	C	-	-	-	-	-

MY11	G	C	M	C	A	G	G	G	W	C	A	T	A	A	Y	A	A	T	G	G
HPV-016*	G	C	A	C	A	G	G	G	C	C	A	C	A	A	T	A	A	T	G	G
HPV-072	-	-	C	-	-	-	-	-	T	-	-	-	-	-	C	-	-	-	-	-
HPV-073	-	-	-	-	-	-	-	-	A	-	-	A	-	-	-	-	-	-	-	-
HPV-074	-	-	G	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-081	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-
HPV-082	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-083	-	-	C	-	-	-	-	-	A	-	-	T	-	-	-	-	-	-	-	-
HPV-084	-	-	C	-	-	-	-	-	T	-	-	T	-	-	C	-	-	-	-	-
HPV-085	-	-	C	-	-	-	-	-	A	-	-	A	-	-	-	-	-	-	-	-
HPV-086	-	-	G	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	A	-
HPV-087	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-089	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-
HPV-090	-	-	-	-	-	-	-	-	-	-	C	T	-	-	C	-	-	-	-	-
HPV-091	-	-	-	-	-	-	-	-	G	-	-	T	-	-	-	-	-	C	-	-
HPV-097	-	-	C	-	-	-	-	-	-	-	-	T	-	-	C	-	-	-	-	-
HPV-102	-	-	C	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-
HPV-106	-	-	-	-	-	-	-	-	-	-	G	A	-	-	C	-	-	-	-	-
HPV-114	-	-	T	-	-	-	-	-	A	-	-	-	-	-	C	-	-	-	-	-

* Denotes the HPV16, which was used as a reference to compare the other HPVs. “-” represents no difference to the reference and a letter represents the change in that particular site.

Table A.3: HPV sequences and their identifiers.

HPV	Sequence Identifier
2	NC_001352.1
6	NC_001355.1
7	NC_001595.1
8	M12737.1
10	NC_001576.1
11	M14119.1
13	DQ344807.1
16	NC_001526.2
18	NC_001357.1
23	U31781.1
26	NC_001583.1
27	X74473.1
30	X74474.1
31	J04353.1
32	NC_001586.1
33	M12732.1
34	NC_001587.1
35	M74117.1
39	M62849.1
40	X74478.1
42	GQ472847.1
43	AJ620205.1
44	U31788.1
45	X74479.1
51	M62877.1
52	X74481.1
53	NC_001593.1
54	NC_001676.1
55	U31791.1
56	X74483.1
57	AB361563.1
58	D90400.1
59	X77858.1
61	NC_001694.1
62	AY395706.1
66	U31794.1
67	D21208.1
68	EU918769.1
69	AB027020.1
70	U21941.1
71	AY330623.1
72	X94164.1
73	X94165.1
74	AF436130.1
81	AJ620209.1
82	AB027021.1
83	AF151983.1
84	AF293960.1
85	AF131950.1
86	AF349909.1
87	AJ400628.2
89	AF436128.1
90	NC_004104.1
91	AF419318.1
97	DQ080080.1
98	FM955837.2
100	FM955839.1
101	NC_008189.1
102	DQ080083.1
103	NC_008188.1
104	FM955840.1
105	FM955841.1
106	DQ080082.1
108	NC_012213.1
111	EU410349.1
113	FM955842.1
145	HM999997.1

Table A.4: HPV forward and reverse primers from the samples that were available in LCP and generated by Primer-BLAST

HPV	Forward Primer	Reverse Primer
2	GGTGAGCACTGGTCTAAGGG	ACGCCTTACCCGTTTTCGTT
6	AAATCGCACGTCTGTAGGGAG	TGCTAAGGGCAACCGAAAAC
7	GGCTTTGGTGCCATGGATTT	TTACGTTTGGGGGTGACTGA
8	TCCATGTGCAGAGGACCAAC	CCTGACGGAGCCCCTAGATA
10	AATCGGGCTAGTGCTGTTGG	TGTACCGAAATCGGTCTCCTG
11	CCTGTGCCTGATGACCTGTT	TACGTTTTCGTTTGGGGGCT
13	ACTCCCAGTGGCTCTCTTGT	ACAGCGTGCTGCAACAAATC
16	TAGGGCTGGTACTGTTGGTG	GGTGGTGGGTGTAGCTTTTC
18	TACGGCGTGAGCAGCTTTTT	CTAGGGCGCAACCACATAAC
23	GCACACCGTGCTTAGGTGAA	TTCGAGTTGCAGGCCGTTT
26	GGGAGAGCACTGGGGTATTG	GGGACGTTTGGTGCCTAGTT
27	AGGGCTGGTAAGATGGGTGA	AAACTGCACGCAACCGAAAA
30	ATAGGGCAGGTGCTATTGGT	AAAGGGTGGAGGACATAGGGA
31	AGCGGCTCCATGGTTACTTC	CGGACCGGGTGTACAACTTT
32	TTGGCGAGCATTGGGGTAAG	TGACTGGAAGATGCTGTTCGTT
33	GTAAGCCTCCAACAGGGGAA	TTTTGCAGACGATGTGCGGG
34	GGCAGGTACTGTAGGCGATG	CCTGTAACCTAGGACGGGCA
35	TGGTAACCTCCGATGCACAA	AGCTTACCAACAGCAACCGA
39	ATACGTGCAAACCCCGGTAG	CTTCGGTCGCCACAAAATGG
40	GGCGACATGGTGGATACTGG	GCAAGGGTGTATTCCCCAGA
42	GCCTGTACACCACAGTCCAA	GCTGTAGACGCCTTTCGTTT
43	CCCAGTGGGTCTTTGGTTACTT	GGAAGACGTTCAGACGCCA
44	CCGTTAGGTGAGCATTGGGG	GACGTTTCCTACCCACACGAA
45	CCCTTCTCCCAGTGGCTCTAT	ATCGGTTGCACAGCAAAATGG
51	CTTGTAGGTGTTGGGGAAGACA	GGTAGGGCAATAGGGACGC
52	GGGGATTGTCCTCCCCTACA	CGATGATGCAGGGCGTTTTAG
53	GCAGGCGTTATTGGTGAGGA	AAACAGGGACGCCACCTAAC
54	GAACACTGGGCTAAAGGCAAC	GCACGCTTTACAGGCCGAAG
55	GGGAGCATTGGGGTAAGGG	TCTACCCACACGAACAGAGGA
56	TAGAGAACCCCCTCCGAGTT	GCAACCACGCGTAAAAGCA
57	TTCTTCAATCGGGGTGGGTC	CATGCAGCGGGGACAATTAC
58	TTGGGGTAAAGGTGTTGCCT	GCCGAACGTTTTAGTCTGGG
59	TGCCAACCCAGGCAGTTATT	ACACGTTTTGGTGATGGGGT
61	ACAGGATGGCGATATGGTGG	TTTTGCGTTTTGTAGCAGGGG
62	TTGCAAATGGCTGCAGAACC	TACGGGACACAGCAACAGAC

66	AGGGGATTGTCCACCTCTTG	AGGCAGATACACTAGCCTTGG
67	ATGGCTAGCGAGGCTTATGG	GGCATGGGCAGGGATATAGG
68	TGTATATGCCCCCTCGCCT	GTCACAGGGTGCAACCACAA
69	GTCAAGGTCGTGAACCTCCAA	TTCGGTTCGCAAAAAGGAACA
70	GGTGGGCGACACAATACCTT	ACAGGGTGCAACCACAAATG
71	CACAGGTTTTGGGGCAATGG	TGCAACCCTTTTACGGGACA
72	CCAGCGGCTCTATGGTGTC	GAAAAGGCGGGAGACCGAAA
73	TGGCTGCTGATCCCTATGGT	CTTAGGTGTGGCACTTGTGGT
74	TATGCATGGTAGGCTGTGCT	GCCGAACGCTTTTTACCTACA
81	TGCTCCAGCACAAAACGAATG	TACGGGACACAGCAAAAGACC
82	TTGGTGATGCCATTCCAGACA	ATAGGGACGCCACCTAAACAC
83	GGTACTGTGTGTTCCGGTGT	AGACACAGACACAGAACGGG
84	GGGGAACATTGGGGTAAGGG	AGACGCTGTACGTTTACGGG
85	GGAGATGCAGTACCGGAAACA	AAGGGCACACAGGATACACAG
86	GCAGGCTATGGTGCTATGGA	AGTTTTGCGACGTTTGGAGG
87	AGGATGGCGACATGGTTGAG	AGTAGAGGTAGAGGCGGGAC
89	TATGTATATGCCCCAACGCCC	AACCGGTCGCTGAACTCAAA
90	GAGCAGGTACTATGGGCGAC	TTTGTAGAGGGGGCGGTAGA
91	AACAGGGCGAACATAGCAGG	GCCATGACTACCACCAGACC
97	AGAGCTGGAACAATGGGTGA	TACGACGGACACGTAAACGC
98	TTTCGTTCGTGGGGGTAAGG	CGTGTTCAGACCGACAGTGA
100	CCACCTGCCACATTGGAAAAC	TGATTAACCGGGTGCGGTC
101	GTGACAGTGTTCCAGAGGAGG	CCCAGGTGTGGCCAAAATAC
102	ACCCCCTATTGGGGAACACT	GGCAGACGTGGATGTAGAGG
103	ATGTGCAGATGCTAGGCAGG	GGTCTCTACACGAGACCGCT
104	TGCAGGAGATCAACAGGGAA	CTTCGTTTTGCAGTACCCCG
105	CCAGGATGGGGATATGGCAG	TTCTTTTGCGTTTGGTTCCCC
106	TACTACTACATCCCGCGCCT	GGCCACAAGGACTGGACTAC
108	TGACATTGGCTTTGGTGCCT	AGGACCGTGGGCGTTTTATAG
111	CGTCTATGGGGACGCTTGTT	GCGTCTGCGTTTTGCAGATT
113	GACTCTGAAGCAGGCAACCA	AAGAGCGTTTTCGCGTTTGT
145	GCGAGAGCAATGTTATGCCAG	GGGCGCGGTTACTAGTGTT

Table A.5: Primer MY09 vs HPV sequences: comparison between the MY09 primer to different HPV sequences.

MY09•	G	A	T	C	A	G	T	W	T	C	C	Y	Y	T	K	G	G	A	C	G
HPV-002	-	-	C	-	-	A	-	-	-	-	-	-	-	-	-	-	-	T	-	-
HPV-006	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-011	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-013	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	A	-
HPV-016	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-
HPV-018	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-026	-	-	-	-	-	A	-	-	-	-	-	A	-	-	A	-	-	G	-	-
HPV-027	-	-	C	-	-	A	-	-	-	-	-	-	-	-	-	-	-	T	-	-
HPV-030	-	-	C	-	-	A	-	-	C	-	-	A	-	-	-	-	-	C	A	-
HPV-031	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	T	-	-
HPV-032	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	T	A	-
HPV-033	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-034	-	-	-	-	-	-	-	-	-	G	-	-	-	-	A	-	-	T	A	-
HPV-035	-	-	-	-	-	A	-	-	-	-	-	G	-	-	-	-	-	C	-	-
HPV-039	-	-	-	-	-	A	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-040	-	-	-	-	-	A	-	-	-	-	-	A	-	-	A	-	-	-	-	-
HPV-042	-	-	-	-	-	A	-	-	-	-	-	-	-	-	A	-	-	T	A	-
HPV-043	A	C	C	-	-	-	-	-	-	-	-	-	-	-	A	-	-	G	-	-
HPV-044	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	T	A	-
HPV-045	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	T	-	-
HPV-051	-	-	C	-	-	A	-	-	-	G	-	A	-	-	-	-	-	T	-	-
HPV-052	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	T	A	-
HPV-053	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	A	-
HPV-054	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
HPV-056	-	-	-	-	-	A	-	-	-	-	-	A	-	-	-	-	-	T	A	-
HPV-057	-	-	C	-	-	A	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-058	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-059	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-061	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
HPV-062	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
HPV-066	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	A	-
HPV-067	-	-	-	-	-	A	-	-	-	-	-	-	-	-	A	-	-	T	A	-
HPV-068	-	-	C	-	-	A	-	-	C	-	-	A	-	-	A	-	-	-	-	-
HPV-069	-	-	-	-	-	A	-	-	-	-	-	A	-	-	-	-	-	C	A	-
HPV-070	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-

MY09•	G	A	T	C	A	G	T	W	T	C	C	Y	Y	T	K	G	G	A	C	G
HPV-071	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	G	-	-
HPV-072	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	G	-	-
HPV-073	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-
HPV-074	-	-	-	-	-	A	-	-	C	-	-	-	-	-	-	-	-	T	A	-
HPV-081	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
HPV-082	-	-	-	-	-	-	-	-	-	G	-	A	-	-	-	-	-	T	-	-
HPV-083	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
HPV-084	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
HPV-085	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-
HPV-086	-	-	C	-	-	A	-	-	-	-	-	A	-	-	-	-	-	C	-	-
HPV-087	-	-	C	-	-	A	-	-	-	-	-	-	-	-	-	-	-	C	-	-
HPV-089	-	-	C	-	-	-	-	-	C	-	-	-	-	-	-	-	-	T	-	-
HPV-090	-	-	C	-	-	A	-	-	C	-	-	-	-	-	-	-	-	C	A	-
HPV-091	A	C	-	-	-	A	-	-	-	-	-	A	-	-	A	-	-	-	-	-
HPV-097	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-102	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	G	-	-
HPV-106	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	A	-
HPV-114	-	-	-	-	-	A	-	-	-	-	-	-	-	-	A	-	-	C	-	-

“-” Represents no difference to the reference and a letter represents the change in that particular site.

• Complement Sequence (5’ → 3’)

Table A.6: Primer MY11 vs HPV sequences: comparison between the MY11 primer to different HPV sequences.

MY11	G	C	M	C	A	G	G	G	W	C	A	T	A	A	Y	A	A	T	G	G
HPV-002	-	-	-	-	-	A	-	-	G	-	-	C	-	-	-	-	-	-	-	-
HPV-006	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-011	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-013	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-
HPV-016	-	-	-	-	-	-	-	-	C	-	-	C	-	-	-	-	-	-	-	-
HPV-018	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-026	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-027	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-030	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-
HPV-031	-	-	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-
HPV-032	-	-	-	-	-	A	-	-	C	-	-	C	-	-	-	-	-	-	-	-
HPV-033	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-034	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-
HPV-035	-	-	-	-	-	A	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-039	-	-	-	-	-	-	-	-	C	-	-	C	-	-	-	-	-	-	-	-
HPV-040	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-042	-	-	-	-	-	A	-	-	-	-	-	C	-	-	-	-	-	-	-	-
HPV-043	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-044	-	-	G	-	-	-	-	-	C	-	-	C	-	-	-	-	-	-	-	-
HPV-045	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-051	-	-	G	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-
HPV-052	-	-	G	-	-	-	-	-	C	-	-	C	-	-	-	-	-	-	-	-
HPV-053	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-054	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-
HPV-056	-	-	-	-	-	A	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-057	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-058	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-059	-	-	T	-	-	-	-	-	-	T	T	A	-	-	-	-	-	-	-	-
HPV-061	-	-	-	-	-	-	-	-	C	-	-	C	-	-	-	-	-	-	-	-
HPV-062	-	-	G	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-066	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-067	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-068	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-
HPV-069	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-070	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-

MY11	G	C	M	C	A	G	G	G	W	C	A	T	A	A	Y	A	A	T	G	G
HPV-071	-	-	-	-	-	-	-	-	C	A	C	A	-	-	-	-	-	-	-	-
HPV-072	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-
HPV-073	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-
HPV-074	-	-	G	-	-	A	-	-	C	-	-	C	-	-	-	-	-	-	-	-
HPV-081	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-082	-	-	-	-	-	-	-	-	C	-	-	C	-	-	-	-	-	-	-	-
HPV-083	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-084	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HPV-085	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-
HPV-086	-	-	G	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	A	-
HPV-087	-	-	-	-	-	-	-	-	C	-	-	C	-	-	-	-	-	-	-	-
HPV-089	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-090	-	-	-	-	-	-	-	-	C	-	C	-	-	-	-	-	-	-	-	-
HPV-091	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	C	-	-
HPV-097	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-
HPV-102	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-
HPV-106	-	-	-	-	-	-	-	-	C	-	G	A	-	-	-	-	-	-	-	-
HPV-114	-	-	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-

"-" Represents no difference to the reference and a letter represents the change in that particular site.

Table A.7: Unique MY09 sequences among the same HPV types

Type	Sequence	MY09•	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	N
HPV016	HQ644284	GATCAGTTTCCTTTAGGACG	G	A	T	C	A	G	T	T	T	C	C	T	T	T	A	G	G	A	C	G	294
HPV016	AB889492	GATCAGTTTCCTTTAGGTCG	G	A	T	C	A	G	T	T	T	C	C	T	T	T	A	G	G	T	C	G	1
HPV016	HQ644257	GATCATTTTCCTTTAGGACG	G	A	T	C	A	T	T	T	T	C	C	T	T	T	A	G	G	A	C	G	1
HPV016	HQ644294	TATCAGTTTCCTTTAGGACG	T	A	T	C	A	G	T	T	T	C	C	T	T	T	A	G	G	A	C	G	2
HPV018	NC001357	GATCAATATCCCCTTGGACG	G	A	T	C	A	A	T	A	T	C	C	C	C	T	T	G	G	A	C	G	72
HPV018	DQ486472	GATCAGTATCCCCTTGGACG	G	A	T	C	A	G	T	A	T	C	C	C	C	T	T	G	G	A	C	G	5
HPV018	JQ917454	GATCAATATCCCCTTGGGCG	G	A	T	C	A	A	T	A	T	C	C	C	C	T	T	G	G	G	C	G	1
HPV018	EF202151	AATCAATATCCCCTTGGACG	A	A	T	C	A	A	T	A	T	C	C	C	C	T	T	G	G	A	C	G	1
HPV031	J04353	GATCAGTTTCCACTGGGTCG	G	A	T	C	A	G	T	T	T	C	C	A	C	T	G	G	G	T	C	G	41
HPV031	KJ754578	GATCAGTTTCCCCTGGGTCG	G	A	T	C	A	G	T	T	T	C	C	C	C	T	G	G	G	T	C	G	1
HPV033	A12360	GATCAGTTTCCTTTGGGACG	G	A	T	C	A	G	T	T	T	C	C	T	T	T	G	G	G	A	C	G	241
HPV035	M74117	GATCAGTTTCCGTTGGGCCG	G	A	T	C	A	G	T	T	T	C	C	G	T	T	G	G	G	C	C	G	1
HPV035	HQ537722	GATCAATTTCCGTTGGGCCG	G	A	T	C	A	A	T	T	T	C	C	G	T	T	G	G	G	C	C	G	49
HPV035	JN104062	GGTCTATTTCCGTTGGGCCG	G	G	T	C	T	A	T	T	T	C	C	G	T	T	G	G	G	C	C	G	1
HPV035	GQ479032	GATCAATTTCCGTTAGGCCG	G	A	T	C	A	A	T	T	T	C	C	G	T	T	A	G	G	C	C	G	1
HPV039	JN104068	GATCAATTCCCTTTGGGACG	G	A	T	C	A	A	T	T	C	C	C	T	T	T	G	G	G	A	C	G	21
HPV039	JQ902116	GATCAGTTCCCTTTGGGACG	G	A	T	C	A	G	T	T	C	C	C	T	T	T	G	G	G	A	C	G	1
HPV039	JN104069	GATCAATTCCCCTTGGGACG	G	A	T	C	A	A	T	T	C	C	C	C	T	T	G	G	G	A	C	G	1
HPV039	U45899	GATCAGTATCCCTTGGGACG	G	A	T	C	A	G	T	A	T	C	C	C	T	T	G	G	G	A	C	G	7
HPV045	EF202157	GATCAATATCCCCTTGGTCG	G	A	T	C	A	A	T	A	T	C	C	C	C	T	T	G	G	T	C	G	37
HPV051	KF436866	GACCAATTTGCATTGGGTCG	G	A	C	C	A	A	T	T	T	G	C	A	T	T	G	G	G	T	C	G	53
HPV051	KF707621	TACCAATTTGCATTGGGTCG	T	A	C	C	A	A	T	T	T	G	C	A	T	T	G	G	G	T	C	G	1
HPV052	X74481	GATCAGTTTCCTTTAGGTAG	G	A	T	C	A	G	T	T	T	C	C	T	T	T	A	G	G	T	A	G	174
HPV052	EU077221	GATCAGTTTCCTTTAGGCAG	G	A	T	C	A	G	T	T	T	C	C	T	T	T	A	G	G	C	A	G	14
HPV053	KF436825	GATCAGTTTCCTCTTGGCAG	G	A	T	C	A	G	T	T	T	C	C	T	C	T	T	G	G	C	A	G	23
HPV053	DQ486475	GATCAGTTTCCCCTTGGCAG	G	A	T	C	A	G	T	T	T	C	C	C	C	T	T	G	G	C	A	G	1
HPV053	KF436823	GATCAGTTCCCTCTTGGCAG	G	A	T	C	A	G	T	T	C	C	C	T	C	T	T	G	G	C	A	G	2
HPV056	X74483	GATCAATTTCCACTGGGTAG	G	A	T	C	A	A	T	T	T	C	C	A	C	T	G	G	G	T	A	G	8
HPV056	KC815982	GATCAGTTTCCCTTTGGACG	G	A	T	C	A	G	T	T	T	C	C	C	T	T	T	G	G	A	C	G	3
HPV056	KC815983	GATCAGTTTCCTCTTGGACA	G	A	T	C	A	G	T	T	T	C	C	T	C	T	T	G	G	A	C	A	1
HPV056	KC815981	GATCAGTATCCCTTGGGACA	G	A	T	C	A	G	T	A	T	C	C	C	T	T	G	G	G	A	C	A	1
HPV058	FJ385261	GATCAGTTTCCTTTGGGACC	G	A	T	C	A	G	T	T	T	C	C	T	T	T	G	G	G	A	C	C	1
HPV059	AF374230	GATCAGTTTCCTTTTGGACG	G	A	T	C	A	G	T	T	T	C	C	T	T	T	T	G	G	A	C	G	1
HPV059	JQ902128	GATCAGTTTCCCTTTGGGCG	G	A	T	C	A	G	T	T	T	C	C	C	T	T	T	G	G	G	C	G	1
HPV059	AB437933	GATCAGTATCCCCTGGGACG	G	A	T	C	A	G	T	A	T	C	C	C	C	T	G	G	G	A	C	G	1
HPV059	AB437934	GATCAGTATCCCCGGGGACG	G	A	T	C	A	G	T	A	T	C	C	C	C	G	G	G	G	A	C	G	1
HPV066	EF177189	GATCAGTTTCCTTTGGGTAG	G	A	T	C	A	G	T	T	T	C	C	T	T	T	G	G	G	T	A	G	11
HPV066	AY147908	GATCAGTATCCNTTGGGTAG	G	A	T	C	A	G	T	A	T	C	C	N	T	T	G	G	G	T	A	G	1
HPV066	EF177182	GATCAGCTTCCTTTGGGTAG	G	A	T	C	A	G	C	T	T	C	C	T	T	T	G	G	G	T	A	G	1
HPV066	U12498	GATCAGTTTCCTTTTGGTAG	G	A	T	C	A	G	T	T	T	C	C	T	T	T	T	G	G	T	A	G	2
HPV067	D21208	GATCAATTTCCCTTAGGTAG	G	A	T	C	A	A	T	T	T	C	C	C	T	T	A	G	G	T	A	G	8
HPV068	M73258	GACCAGTTTCCTTTAGGACG	G	A	C	C	A	G	T	T	T	C	C	T	T	T	A	G	G	A	C	G	17
HPV068	Y17206	GATCAGTTTCCCCTTGGACG	G	A	T	C	A	G	T	T	T	C	C	C	C	T	T	G	G	A	C	G	1
HPV068	AJ831567	GATCAGTATCCSTTTGGACG	G	A	T	C	A	G	T	A	T	C	C	S	T	T	T	G	G	A	C	G	1
HPV068	DQ080079	GACCAATTCCCATTAGGACG	G	A	C	C	A	A	T	T	C	C	C	A	T	T	A	G	G	A	C	G	5
HPV068	KC470270	GACCAATTTCCATTAGGACG	G	A	C	C	A	A	T	T	T	C	C	A	T	T	A	G	G	A	C	G	2

• Complement Sequence (5’ → 3’)

Table A.8: Unique MY11 sequences among the same HPV types

Type	Sequence	MY11	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	N
HPV016	HQ644284	GCACAGGGCCACAATAATGG	G	C	A	C	A	G	G	G	C	C	A	C	A	A	T	A	A	T	G	G	296
HPV016	HQ644246	GCACAGGGCCACAATACTGG	G	C	A	C	A	G	G	G	C	C	A	C	A	A	T	A	C	T	G	G	1
HPV016	AF084952	ACACAGGGCCACAATAATGG	A	C	A	C	A	G	G	G	C	C	A	C	A	A	T	A	A	T	G	G	1
HPV016	HM057182	GCACAGGGTCATAACAATGG	G	C	A	C	A	G	G	G	T	C	A	T	A	A	C	A	A	T	G	G	85
HPV018	U45891	GCACAGGGACATAACAATGG	G	C	A	C	A	G	G	G	A	C	A	T	A	A	C	A	A	T	G	G	2
HPV031	J04353	GCTCAGGGACACAATAATGG	G	C	T	C	A	G	G	G	A	C	A	C	A	A	T	A	A	T	G	G	42
HPV031	AJ831566	GCCCAGGGTCATAAGAATGG	G	C	C	C	A	G	G	G	T	C	A	T	A	A	G	A	A	T	G	G	1
HPV033	A12360	GCACAAGGTCATAATAATGG	G	C	A	C	A	A	G	G	T	C	A	T	A	A	T	A	A	T	G	G	32
HPV033	DQ486473	GCACAGGGTCATAATAATGG	G	C	A	C	A	G	G	G	T	C	A	T	A	A	T	A	A	T	G	G	13
HPV035	M74117	GCACAAGGCCATAATAATGG	G	C	A	C	A	A	G	G	C	C	A	T	A	A	T	A	A	T	G	G	52
HPV039	JN104068	GCCCAGGGCCACAACAATGG	G	C	C	C	A	G	G	G	C	C	A	C	A	A	C	A	A	T	G	G	22
HPV039	JQ902116	GCCCAGGGCCACAATAATGG	G	C	C	C	A	G	G	G	C	C	A	C	A	A	T	A	A	T	G	G	1
HPV039	U45899	GCCCAGGGTCATAATAATGG	G	C	C	C	A	G	G	G	T	C	A	T	A	A	T	A	A	T	G	G	7
HPV045	EF202157	GCCCAGGGCCATAACAATGG	G	C	C	C	A	G	G	G	C	C	A	T	A	A	C	A	A	T	G	G	37
HPV051	KF436866	GCGCAGGGTCACAATAATGG	G	C	G	C	A	G	G	G	T	C	A	C	A	A	T	A	A	T	G	G	48
HPV051	KF707617	GCGCAGGGCCACAATAATGG	G	C	G	C	A	G	G	G	C	C	A	C	A	A	T	A	A	T	G	G	192
HPV052	KJ676024	GCTCAGGGCCACAATAATGG	G	C	T	C	A	G	G	G	C	C	A	C	A	A	T	A	A	T	G	G	1
HPV053	KF436825	GCCCAGGGACATAATAATGG	G	C	C	C	A	G	G	G	A	C	A	T	A	A	T	A	A	T	G	G	25
HPV053	DQ486475	GCACAGGGACATAATAATGG	G	C	A	C	A	G	G	G	A	C	A	T	A	A	T	A	A	T	G	G	1
HPV056	X74483	GCCCAAGGCCATAATAATGG	G	C	C	C	A	A	G	G	C	C	A	T	A	A	T	A	A	T	G	G	8
HPV056	KC815982	GCCCAAGGACATAATAATGG	G	C	C	C	A	A	G	G	A	C	A	T	A	A	T	A	A	T	G	G	2
HPV056	KC815983	GCCCAAGGACATAACAATGG	G	C	C	C	A	A	G	G	A	C	A	T	A	A	C	A	A	T	G	G	1
HPV058	D90400	GCACAAGGTCATAACAATGG	G	C	A	C	A	A	G	G	T	C	A	T	A	A	C	A	A	T	G	G	193
HPV058	HQ537773	GCACAAGGTCATAACAATGA	G	C	A	C	A	A	G	G	T	C	A	T	A	A	C	A	A	T	G	A	1
HPV059	KC470263	GCTCAGGGTTTAAACAATGG	G	C	T	C	A	G	G	G	T	T	T	A	A	A	C	A	A	T	G	G	13
HPV066	EF177189	GCACAGGGCCATAATAATGG	G	C	A	C	A	G	G	G	C	C	A	T	A	A	T	A	A	T	G	G	10
HPV067	U12492	GCCCAGGGACATAACAATGG	G	C	C	C	A	G	G	G	A	C	A	T	A	A	C	A	A	T	G	G	1
HPV068	M73258	GCACAGGGACACAACAATGG	G	C	A	C	A	G	G	G	A	C	A	C	A	A	C	A	A	T	G	G	18
HPV068	JQ902131	GCACAAGGACACAACAATGG	G	C	A	C	A	A	G	G	A	C	A	C	A	A	C	A	A	T	G	G	5
HPV068	KC470275	GCACAAGGGCACAACAATGG	G	C	A	C	A	A	G	G	G	C	A	C	A	A	C	A	A	T	G	G	1
HPV068	AJ831567	GCCCAGGGTCATAAGAMTGG	G	C	C	C	A	G	G	G	T	C	A	T	A	A	G	A	M	T	G	G	1
HPV068	AJ831568	GCCCAGGGACATAAGAATGG	G	C	C	C	A	G	G	G	A	C	A	T	A	A	G	A	A	T	G	G	1

Table A.9: Variants of HPVs and their identifiers

Type	Sequence	Type	Sequence	Type	Sequence	Type	Sequence
HPV016	HQ644284	HPV016	GQ479010	HPV045	KC470259	HPV053	EF546474
HPV016	HQ644275	HPV016	DQ680078	HPV045	U45907	HPV056	X74483
HPV016	FJ610149	HPV016	HQ644238	HPV045	U45908	HPV056	EF177181
HPV016	FJ610146	HPV016	HQ644240	HPV045	U45909	HPV056	EF177180
HPV016	HQ644234	HPV016	AF472508	HPV045	U45910	HPV056	EF177178
HPV016	JQ004094	HPV016	GQ479011	HPV045	EF202162	HPV056	EF177179
HPV016	AF125673	HPV016	HQ644298	HPV045	EF202163	HPV056	EF177176
HPV016	JQ004099	HPV016	AY098926	HPV045	EF202161	HPV056	JQ902126
HPV016	HQ644267	HPV016	AF003031	HPV045	KC470257	HPV056	EF177177
HPV016	HQ644264	HPV016	AF084952	HPV045	U45911	HPV056	KC815982
HPV016	HQ644283	HPV016	KP161010	HPV045	U45914	HPV056	KC815983
HPV016	HQ644297	HPV016	KP160988	HPV045	U45916	HPV056	KC815981
HPV016	KP212152	HPV016	KM058644	HPV045	JQ902121	HPV058	D90400
HPV016	AY686581	HPV016	EU477377	HPV051	KF436866	HPV058	HM639578
HPV016	KP212153	HPV016	AF043287	HPV051	KF436868	HPV058	HM639600
HPV016	KP212150	HPV016	HM057182	HPV051	KP090027	HPV058	HM639595
HPV016	HQ644280	HPV016	JF728170	HPV051	KP090024	HPV058	AB819279
HPV016	AB818687	HPV016	FJ797044	HPV051	KP090012	HPV058	HQ537760
HPV016	HQ644236	HPV018	NC001357	HPV051	KP090011	HPV058	HM639588
HPV016	KF954093	HPV018	GQ180786	HPV051	KP090009	HPV058	HM639508
HPV016	AF534061	HPV018	EF202145	HPV051	KP090008	HPV058	HM639614
HPV016	HQ644235	HPV018	GQ180784	HPV051	KP090002	HPV058	HM639617
HPV016	FJ610150	HPV018	GQ180788	HPV051	KP090004	HPV058	HM639585
HPV016	AB889494	HPV018	GQ180789	HPV051	KP090031	HPV058	HM639591
HPV016	HQ644287	HPV018	GQ180790	HPV051	KP090007	HPV058	HM639621
HPV016	KP874716	HPV018	GQ180792	HPV051	KP090000	HPV058	HQ537756
HPV016	HQ644261	HPV018	KC470208	HPV051	KP090020	HPV058	HM639581
HPV016	AB818691	HPV018	KC470209	HPV051	M62877	HPV058	HM639612
HPV016	AY686580	HPV018	EF202147	HPV051	KP090013	HPV058	HM639618
HPV016	KP212155	HPV018	EF202148	HPV051	KP090003	HPV058	HM639619
HPV016	AB818688	HPV018	EF202149	HPV051	KF436867	HPV058	JX313761
HPV016	FJ610151	HPV018	KC470212	HPV051	KP090006	HPV058	JX313771
HPV016	AB889489	HPV018	AY262282	HPV051	KP090016	HPV058	HM639606
HPV016	JQ004093	HPV018	GQ180785	HPV051	KF707617	HPV058	HM639620
HPV016	JQ067943	HPV018	GQ180791	HPV051	U45917	HPV058	HM639709
HPV016	FJ610147	HPV018	GQ180787	HPV051	KF707621	HPV058	HM639710
HPV016	HQ644271	HPV018	KC470213	HPV051	KP090033	HPV058	HM639692
HPV016	KF880690	HPV018	EF202144	HPV051	KF436879	HPV058	HM639716
HPV016	HQ644274	HPV018	FJ528600	HPV051	KF436878	HPV058	HM639613
HPV016	JQ067944	HPV018	EF202150	HPV051	KF436877	HPV058	HM639708
HPV016	HQ644260	HPV018	X05015	HPV051	KF436876	HPV058	FJ407208

Type	Sequence	Type	Sequence	Type	Sequence	Type	Sequence
HPV016	JN565302	HPV018	EU834744	HPV051	KF436875	HPV058	FJ407201
HPV016	HQ644286	HPV018	KC456642	HPV051	KP090001	HPV058	FJ407194
HPV016	HQ644259	HPV018	EF202143	HPV051	GQ487711	HPV058	FJ385267
HPV016	AY686583	HPV018	AY863160	HPV051	KF436873	HPV058	FJ385262
HPV016	HQ644256	HPV018	AY863157	HPV051	KF436874	HPV058	FJ385265
HPV016	HQ644258	HPV018	AY863156	HPV051	KF436869	HPV058	HM639699
HPV016	AF536179	HPV018	AY863161	HPV051	KP090023	HPV058	HM639717
HPV016	HQ644262	HPV018	AY863158	HPV051	KP090026	HPV058	HM639532
HPV016	HQ644282	HPV018	AY863159	HPV051	JQ902122	HPV058	HQ537757
HPV016	KP212157	HPV018	AY863162	HPV051	KF436871	HPV058	HQ537759
HPV016	HQ644272	HPV018	U89349	HPV051	KF436872	HPV058	HM639701
HPV016	AB889492	HPV018	JF728187	HPV051	KF707619	HPV058	HM639697
HPV016	HQ644266	HPV018	JF728184	HPV051	AB438955	HPV058	HM639610
HPV016	AB889488	HPV018	JF728185	HPV051	AB438954	HPV058	FJ385263
HPV016	HQ644248	HPV018	JF728186	HPV051	KP090028	HPV058	FJ385261
HPV016	HQ644268	HPV018	DQ486472	HPV051	KF436884	HPV058	HM639573
HPV016	FJ610148	HPV018	JQ917454	HPV051	KF436883	HPV058	HM639712
HPV016	EU918764	HPV018	JQ902111	HPV051	KF436882	HPV058	FJ407216
HPV016	KC935953	HPV018	U45889	HPV051	KF436881	HPV058	HM639715
HPV016	FJ610152	HPV018	EF202146	HPV051	KF436880	HPV058	EU999960
HPV016	FJ006723	HPV018	KC470210	HPV051	KF707626	HPV058	EU999961
HPV016	KP212151	HPV018	KC470211	HPV051	KF436886	HPV058	JX313768
HPV016	HQ644245	HPV018	EF202151	HPV051	KF436885	HPV058	JX313772
HPV016	EU118173	HPV018	U45890	HPV051	KF436887	HPV058	HM639616
HPV016	JF728173	HPV018	U45891	HPV051	KF707624	HPV058	HM639714
HPV016	JF728179	HPV018	KF225496	HPV051	KF707618	HPV058	HM639706
HPV016	JF728176	HPV018	KC470214	HPV052	X74481	HPV058	HM639703
HPV016	JF728159	HPV018	KC470216	HPV052	EU077193	HPV058	HM639684
HPV016	JF728164	HPV018	KC470217	HPV052	EU077195	HPV058	FJ407192
HPV016	JF728181	HPV018	KC470219	HPV052	EU077196	HPV058	FJ407195
HPV016	JF728182	HPV018	KC470222	HPV052	EU077197	HPV058	FJ385268
HPV016	JF728155	HPV018	KC470218	HPV052	EU077198	HPV058	JX313759
HPV016	JF728183	HPV018	EF202152	HPV052	EU077199	HPV058	JX313758
HPV016	JF728175	HPV018	KC470226	HPV052	EU077200	HPV058	KC860270
HPV016	JF728180	HPV018	EF202155	HPV052	EU077201	HPV058	HM639711
HPV016	JF728174	HPV018	KC470228	HPV052	EU077202	HPV058	HM639707
HPV016	JF728168	HPV018	KC470224	HPV052	EU077203	HPV058	EU918765
HPV016	JF728167	HPV018	KC470227	HPV052	EU077204	HPV058	HM639457
HPV016	KJ467225	HPV018	KC470225	HPV052	EU077206	HPV058	HM639529
HPV016	KJ467230	HPV018	KC470223	HPV052	EU077207	HPV058	HM639522
HPV016	KJ467234	HPV018	AY863165	HPV052	EU077208	HPV058	HM639524
HPV016	U89348	HPV018	JF728188	HPV052	EU077209	HPV058	HM639318
HPV016	EU430680	HPV018	U45894	HPV052	EU077211	HPV058	HM639509

Type	Sequence	Type	Sequence	Type	Sequence	Type	Sequence
HPV016	GQ465886	HPV018	U45892	HPV052	EU077212	HPV058	HM639525
HPV016	GQ465880	HPV018	KC470221	HPV052	EU077213	HPV058	HQ537752
HPV016	JX313720	HPV018	U45893	HPV052	EU077215	HPV058	HQ537753
HPV016	GQ465881	HPV018	EF202153	HPV052	EU077216	HPV058	HQ537755
HPV016	GQ465897	HPV018	KC470215	HPV052	EU077220	HPV058	FJ407200
HPV016	JX313709	HPV018	KC470229	HPV052	HQ537731	HPV058	FJ407199
HPV016	JX313708	HPV018	KC470230	HPV052	HQ537732	HPV058	AB819278
HPV016	JX313699	HPV018	AY863164	HPV052	HQ537734	HPV058	AB819276
HPV016	KC166220	HPV018	AY863163	HPV052	HQ537735	HPV058	AB819275
HPV016	GQ465891	HPV018	EF202154	HPV052	HQ537736	HPV058	HM639486
HPV016	EU430672	HPV031	J04353	HPV052	HQ537737	HPV058	HM639528
HPV016	GQ465896	HPV031	HQ537668	HPV052	HQ537738	HPV058	HM639458
HPV016	GQ465878	HPV031	HQ537666	HPV052	HQ537739	HPV058	HM639530
HPV016	GQ465893	HPV031	HQ537675	HPV052	HQ537740	HPV058	HM639485
HPV016	GQ465885	HPV031	KJ754573	HPV052	HQ537741	HPV058	HM639504
HPV016	GQ465894	HPV031	KJ754572	HPV052	HQ537745	HPV058	HM639500
HPV016	GQ465877	HPV031	KJ754569	HPV052	HQ537746	HPV058	HM639488
HPV016	JX313700	HPV031	KJ754576	HPV052	JN874419	HPV058	HM639337
HPV016	GQ465889	HPV031	HQ537686	HPV052	JN874422	HPV058	HM639459
HPV016	GQ465883	HPV031	HQ537685	HPV052	JN874424	HPV058	HM639359
HPV016	GQ465884	HPV031	U37410	HPV052	KJ675998	HPV058	HM639428
HPV016	GQ465882	HPV031	KJ754579	HPV052	KJ676004	HPV058	HM639464
HPV016	GQ465895	HPV031	KJ754563	HPV052	KJ676012	HPV058	HM639454
HPV016	JX313703	HPV031	HQ537687	HPV052	KJ676022	HPV058	HM639521
HPV016	GQ465887	HPV031	KJ754565	HPV052	KJ676031	HPV058	AB819277
HPV016	GQ465890	HPV031	KJ754577	HPV052	KJ676035	HPV058	JX313765
HPV016	JX313697	HPV031	KJ754566	HPV052	KJ676041	HPV058	HM639481
HPV016	JX313695	HPV031	KJ754568	HPV052	KJ676043	HPV058	HM639513
HPV016	GQ465892	HPV031	HQ537680	HPV052	KJ676045	HPV058	HM639489
HPV016	JX313705	HPV031	KJ754562	HPV052	KJ676046	HPV058	HM639474
HPV016	GQ465888	HPV031	KJ754564	HPV052	KJ676047	HPV058	HM639492
HPV016	GQ465879	HPV031	HQ537681	HPV052	KJ676048	HPV058	GQ472850
HPV016	JX313702	HPV031	HQ537679	HPV052	KJ676050	HPV058	HM639317
HPV016	JX313715	HPV031	HQ537678	HPV052	KJ676051	HPV058	HM639351
HPV016	JX313696	HPV031	KJ754571	HPV052	KJ676053	HPV058	HM639487
HPV016	KM058636	HPV031	HQ537677	HPV052	KJ676055	HPV058	HM639493
HPV016	KP161014	HPV031	HQ537676	HPV052	KJ676058	HPV058	HM639506
HPV016	KP160999	HPV031	KJ754567	HPV052	KJ676059	HPV058	HM639497
HPV016	KM058645	HPV031	KJ754561	HPV052	KJ676061	HPV058	JX313766
HPV016	KM058642	HPV031	HQ537667	HPV052	KJ676062	HPV058	HM639518
HPV016	FJ797033	HPV031	KJ754575	HPV052	KJ676064	HPV058	HM639490
HPV016	KM058665	HPV031	HQ537671	HPV052	KJ676065	HPV058	KC860271
HPV016	DQ155283	HPV031	KJ754570	HPV052	KJ676066	HPV058	HM639442

Type	Sequence	Type	Sequence	Type	Sequence	Type	Sequence
HPV016	AF140365	HPV031	HQ537672	HPV052	KJ676070	HPV058	HM639678
HPV016	FJ797051	HPV031	HQ537673	HPV052	KJ676071	HPV058	FJ385264
HPV016	FJ797050	HPV031	HQ537670	HPV052	KJ676072	HPV058	HM639615
HPV016	FJ797035	HPV031	HQ537682	HPV052	KJ676078	HPV058	HM639611
HPV016	FJ797043	HPV031	HQ537683	HPV052	KJ676084	HPV058	JX313770
HPV016	KM058666	HPV031	KJ754580	HPV052	KJ676085	HPV058	HM639582
HPV016	DQ469930	HPV031	KJ754578	HPV052	KJ676087	HPV058	HM639527
HPV016	AF134177	HPV031	HQ537674	HPV052	KJ676095	HPV058	HM639471
HPV016	KM058637	HPV031	HQ537669	HPV052	AB819272	HPV058	JX313752
HPV016	FJ797048	HPV031	AJ831566	HPV052	AB819273	HPV058	HM639688
HPV016	AY177679	HPV033	A12360	HPV052	AB819274	HPV058	HM639695
HPV016	KM058654	HPV033	HQ537694	HPV052	GQ472848	HPV058	HM639690
HPV016	FJ797040	HPV033	HQ537693	HPV052	HQ537742	HPV058	HM639667
HPV016	FJ797036	HPV033	HQ537692	HPV052	HQ537743	HPV058	HM639656
HPV016	FJ797034	HPV033	HQ537691	HPV052	JN874416	HPV058	HM639671
HPV016	FJ797056	HPV033	HQ537690	HPV052	JN874417	HPV058	HM639664
HPV016	U34173	HPV033	HQ537688	HPV052	JN874420	HPV058	HM639520
HPV016	U34187	HPV033	GQ479013	HPV052	JN874423	HPV058	HM639531
HPV016	U34167	HPV033	HQ537689	HPV052	JN874425	HPV058	HM639472
HPV016	U34181	HPV033	GQ479014	HPV052	JN874426	HPV058	HM639516
HPV016	U34179	HPV033	HQ537697	HPV052	JN874427	HPV058	HM639515
HPV016	U34193	HPV033	HQ537700	HPV052	JN874428	HPV058	HM639466
HPV016	U34171	HPV033	HQ537699	HPV052	JN874430	HPV058	HM639519
HPV016	U34185	HPV033	HQ537698	HPV052	JN874434	HPV058	HM639526
HPV016	U34169	HPV033	GQ479015	HPV052	JN874436	HPV058	HM639477
HPV016	U34168	HPV033	HQ537701	HPV052	KF225497	HPV058	HM639511
HPV016	JX313707	HPV033	HQ537706	HPV052	KJ675965	HPV058	HM639467
HPV016	JX313701	HPV033	GQ479018	HPV052	KJ675967	HPV058	HM639470
HPV016	JX313723	HPV033	HQ537704	HPV052	KJ675969	HPV058	JX313764
HPV016	JX313693	HPV033	HQ537703	HPV052	KJ675970	HPV058	JX313763
HPV016	NC001526	HPV033	GQ479016	HPV052	KJ675971	HPV058	HM639523
HPV016	AY098925	HPV033	HQ537705	HPV052	KJ675973	HPV058	HM639478
HPV016	JF728163	HPV033	GQ479019	HPV052	KJ675976	HPV058	HM639483
HPV016	JQ004092	HPV033	HQ537707	HPV052	KJ675977	HPV058	KC860269
HPV016	JQ004096	HPV033	HQ537702	HPV052	KJ675979	HPV058	HQ537758
HPV016	AB818693	HPV033	GQ479017	HPV052	KJ675980	HPV058	HM639608
HPV016	AB818692	HPV033	U45896	HPV052	KJ675981	HPV058	HQ537754
HPV016	JN565303	HPV033	GQ479012	HPV052	KJ675982	HPV058	HM639705
HPV016	AB889491	HPV033	DQ486473	HPV052	KJ675985	HPV058	HM639650
HPV016	AY686584	HPV033	FJ202004	HPV052	KJ675987	HPV058	HM639674
HPV016	HQ644246	HPV033	FJ202006	HPV052	KJ675988	HPV058	HQ537771
HPV016	AB889490	HPV033	EF626588	HPV052	KJ675989	HPV058	AY101598
HPV016	AB889493	HPV033	EU779744	HPV052	KJ675990	HPV058	HQ537769

Type	Sequence	Type	Sequence	Type	Sequence	Type	Sequence
HPV016	JF728177	HPV033	EU918766	HPV052	KJ675993	HPV058	HM639622
HPV016	JF728178	HPV033	HQ537695	HPV052	KJ675994	HPV058	HQ537766
HPV016	EU430688	HPV033	HQ537696	HPV052	KJ675995	HPV058	HQ537767
HPV016	JX313706	HPV033	KF436865	HPV052	KJ675996	HPV058	HQ537768
HPV016	JX313698	HPV035	M74117	HPV052	KJ675997	HPV058	HQ537770
HPV016	JX313722	HPV035	HQ537722	HPV052	KJ675999	HPV058	HM639672
HPV016	JX313694	HPV035	GQ479029	HPV052	KJ676000	HPV058	HQ537763
HPV016	EF547252	HPV035	HQ537725	HPV052	KJ676001	HPV058	HQ537762
HPV016	GQ465899	HPV035	GQ479024	HPV052	KJ676005	HPV058	HM639662
HPV016	GQ465900	HPV035	HQ537726	HPV052	KJ676006	HPV058	HQ537761
HPV016	FJ797047	HPV035	GQ479036	HPV052	KJ676007	HPV058	HM639655
HPV016	AF393502	HPV035	GQ479025	HPV052	KJ676008	HPV058	HM639675
HPV016	AF043286	HPV035	HQ537711	HPV052	KJ676009	HPV058	HQ537764
HPV016	KJ571159	HPV035	GQ479033	HPV052	KJ676010	HPV058	HQ537765
HPV016	HQ644296	HPV035	HQ537708	HPV052	KJ676011	HPV058	HM639624
HPV016	HQ644299	HPV035	GQ479023	HPV052	KJ676014	HPV058	HQ537774
HPV016	HQ644290	HPV035	HQ537709	HPV052	KJ676015	HPV058	HQ537775
HPV016	HQ644293	HPV035	HQ537710	HPV052	KJ676016	HPV058	HQ537776
HPV016	AF536180	HPV035	GQ479037	HPV052	KJ676017	HPV058	HM639657
HPV016	U34183	HPV035	X74477	HPV052	KJ676018	HPV058	HM639629
HPV016	GQ465902	HPV035	GQ479020	HPV052	KJ676019	HPV058	HQ537772
HPV016	GQ479009	HPV035	GQ479028	HPV052	KJ676020	HPV058	HM639659
HPV016	GQ465901	HPV035	GQ479021	HPV052	KJ676021	HPV058	HM639641
HPV016	GQ479007	HPV035	HQ537712	HPV052	KJ676023	HPV058	HM639669
HPV016	U34189	HPV035	GQ479027	HPV052	KJ676024	HPV058	HQ537777
HPV016	U34172	HPV035	HQ537728	HPV052	KJ676025	HPV058	HM639668
HPV016	U34188	HPV035	HQ537729	HPV052	KJ676026	HPV058	HM639677
HPV016	JQ004097	HPV035	HQ537730	HPV052	KJ676027	HPV058	HM639670
HPV016	JQ004098	HPV035	JX129488	HPV052	KJ676028	HPV058	HM639673
HPV016	HQ644263	HPV035	JX129487	HPV052	KJ676029	HPV058	HM639658
HPV016	AY686582	HPV035	GQ479022	HPV052	KJ676032	HPV058	HM639666
HPV016	HQ644252	HPV035	HQ537713	HPV052	KJ676034	HPV058	HQ537773
HPV016	AY686579	HPV035	HQ537714	HPV052	KJ676036	HPV058	HM639676
HPV016	HQ644239	HPV035	HQ537715	HPV052	KJ676037	HPV058	HM639609
HPV016	HQ644237	HPV035	HQ537716	HPV052	KJ676038	HPV058	HM639484
HPV016	HQ644273	HPV035	HQ537718	HPV052	KJ676040	HPV059	KC470263
HPV016	HQ644279	HPV035	GQ479026	HPV052	KJ676074	HPV059	KC470264
HPV016	HQ644244	HPV035	HQ537717	HPV052	KJ676075	HPV059	KC470265
HPV016	AB818689	HPV035	GQ479039	HPV052	KJ676076	HPV059	KC470266
HPV016	HQ644291	HPV035	GQ479038	HPV052	KJ676079	HPV059	X77858
HPV016	HQ644250	HPV035	GQ479034	HPV052	KJ676080	HPV059	EU918767
HPV016	AF472509	HPV035	JX129485	HPV052	KJ676081	HPV059	U12496
HPV016	HQ644257	HPV035	JN104064	HPV052	KJ676089	HPV059	KC470261

Type	Sequence	Type	Sequence	Type	Sequence	Type	Sequence
HPV016	HQ644281	HPV035	JQ902115	HPV052	KJ676091	HPV059	KC470262
HPV016	HQ644277	HPV035	HQ537721	HPV052	KJ676093	HPV059	AF374230
HPV016	HQ644254	HPV035	GQ479035	HPV052	HQ537744	HPV059	DQ486471
HPV016	HQ644242	HPV035	JN104062	HPV052	EU077194	HPV059	U45932
HPV016	HQ644241	HPV035	HQ537723	HPV052	EU077214	HPV059	JN104073
HPV016	HQ644249	HPV035	HQ537724	HPV052	KJ676088	HPV059	JN104074
HPV016	HQ644270	HPV035	GQ479030	HPV052	EU077218	HPV059	JQ902128
HPV016	HQ644243	HPV035	HQ537727	HPV052	KJ676030	HPV059	AB437933
HPV016	KP161006	HPV035	HQ537719	HPV052	EU077219	HPV059	AB437934
HPV016	KP161001	HPV035	JX129486	HPV052	HQ537733	HPV059	AJ617548
HPV016	U37217	HPV035	GQ479031	HPV052	KJ676042	HPV066	EF177189
HPV016	FJ797046	HPV035	GQ479032	HPV052	KJ676054	HPV066	EF177184
HPV016	KJ467228	HPV035	JN104063	HPV052	KJ676063	HPV066	EF177191
HPV016	KJ467238	HPV039	JN104068	HPV052	KJ676067	HPV066	EF177183
HPV016	GQ479008	HPV039	KC470233	HPV052	KJ676068	HPV066	EF177188
HPV016	KP161013	HPV039	KC470234	HPV052	KJ676083	HPV066	EF177187
HPV016	KP161012	HPV039	KC470235	HPV052	KJ676094	HPV066	EF177190
HPV016	KP161011	HPV039	KC470236	HPV052	JN874421	HPV066	DQ486474
HPV016	KP161009	HPV039	KC470232	HPV052	KJ675983	HPV066	AY147908
HPV016	KP161008	HPV039	KC470237	HPV052	KJ676013	HPV066	EF177182
HPV016	KP161007	HPV039	M62849	HPV052	KJ676033	HPV066	U31794
HPV016	KP161005	HPV039	KC470231	HPV052	KJ676039	HPV066	EF177186
HPV016	KP161004	HPV039	KC470239	HPV052	KJ676077	HPV066	EF177185
HPV016	KP161003	HPV039	KC470240	HPV052	KJ676090	HPV066	U12498
HPV016	KP161002	HPV039	KC470241	HPV052	KJ676092	HPV066	HM585478
HPV016	KP161000	HPV039	KC470242	HPV052	EU077205	HPV067	D21208
HPV016	KP160998	HPV039	KC470243	HPV052	EU077217	HPV067	HQ537778
HPV016	KP160997	HPV039	KC470245	HPV052	KJ675991	HPV067	HQ537781
HPV016	KP160996	HPV039	KC470244	HPV052	EU077221	HPV067	HQ537780
HPV016	KP160995	HPV039	JQ902116	HPV052	EU077222	HPV067	HQ537779
HPV016	KP160994	HPV039	KC470238	HPV052	EU077223	HPV067	HQ537784
HPV016	KP160993	HPV039	KC470248	HPV052	EU077224	HPV067	HQ537783
HPV016	KP160992	HPV039	KC470249	HPV052	EU077225	HPV067	HQ537782
HPV016	KP160991	HPV039	JN104070	HPV052	HQ537747	HPV067	U12492
HPV016	KP160990	HPV039	KC470247	HPV052	HQ537748	HPV068	M73258
HPV016	KM058638	HPV039	JN104069	HPV052	HQ537749	HPV068	KC470273
HPV016	KP160989	HPV039	U45899	HPV052	HQ537750	HPV068	KC470272
HPV016	FJ797045	HPV039	U45901	HPV052	KJ675978	HPV068	KC470274
HPV016	AF134178	HPV039	U45900	HPV052	KJ675984	HPV068	FR751040
HPV016	U34174	HPV039	U45904	HPV052	KJ676073	HPV068	JQ902131
HPV016	U34176	HPV039	U45905	HPV052	EU077226	HPV068	KC470279
HPV016	U34175	HPV039	U45902	HPV052	HQ537751	HPV068	U45934
HPV016	KP212156	HPV039	U45903	HPV053	KF436825	HPV068	KC470281

Type	Sequence	Type	Sequence	Type	Sequence	Type	Sequence
HPV016	KP212159	HPV045	EF202157	HPV053	KF436824	HPV068	KC470280
HPV016	KP212158	HPV045	EF202158	HPV053	EF546479	HPV068	KC470283
HPV016	JQ004095	HPV045	EF202159	HPV053	GQ472849	HPV068	KC470282
HPV016	HQ644247	HPV045	EF202160	HPV053	JQ902124	HPV068	KC470277
HPV016	HQ644289	HPV045	KC470253	HPV053	EF546482	HPV068	KC470278
HPV016	KP212154	HPV045	KC470254	HPV053	EF546481	HPV068	GQ472851
HPV016	HQ644285	HPV045	KC470255	HPV053	DQ486475	HPV068	KC470275
HPV016	HQ644278	HPV045	U45912	HPV053	EF546480	HPV068	KC470276
HPV016	HQ644276	HPV045	X74479	HPV053	EF546476	HPV068	Y17206
HPV016	HQ644269	HPV045	EF202156	HPV053	EF546478	HPV068	AJ831567
HPV016	HQ644253	HPV045	KC470250	HPV053	KF436823	HPV068	AJ831568
HPV016	HQ644294	HPV045	KC470251	HPV053	KF436822	HPV068	DQ080079
HPV016	HQ644255	HPV045	DQ080002	HPV053	KF436818	HPV068	KC470267
HPV016	HQ644295	HPV045	KC470256	HPV053	KF436819	HPV068	KC470268
HPV016	HQ644288	HPV045	U45906	HPV053	KF436821	HPV068	KC470269
HPV016	AF402678	HPV045	U45913	HPV053	KF436820	HPV068	X67161
HPV016	KJ467232	HPV045	U45915	HPV053	NC001593	HPV068	KC470270
HPV016	GQ479006	HPV045	KC470252	HPV053	EF546473	HPV068	KC470271
HPV016	AF134175	HPV045	EF202164	HPV053	EF546470		
HPV016	FJ797037	HPV045	EF202165	HPV053	EF546472		
HPV016	U34170	HPV045	EF202166	HPV053	EF546471		
HPV016	U34165	HPV045	EF202167	HPV053	X74482		
HPV016	KM058656	HPV045	KC470258	HPV053	EF546469		
HPV016	AB818690	HPV045	KC470260	HPV053	EF546475		