Identification and characterization of novel human papillomaviruses in oral rinse samples.

Juliet Dang

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

2015

Reading Committee:
Nancy B. Kiviat, Chair
Qinghua Feng
Dolphine Oda
Richard Presland

Program Authorized to Offer Degree:
School of Dentistry, Oral Health Sciences

Abstract

Background

With close to 200 different human papillomavirus (HPV) types in the papilloma virus genome database (http://pave.niaid.nih.gov), it is not surprising that there are still uncharacterized novel HPVs present in the oral cavity and oropharynx.

In our previous study we discovered and fully cloned three novel types of HPVs in healthy patients (Martin et al., J Clinical Virology. 2014 Jan:59(1):30). We hypothesize that there are new, as yet unidentified oncogenic HPVs present in the oral cavity and oropharynx of head and neck cancer patients, which could be identified using next generation sequencing (NGS) technology.

Objectives

The objectives of this study were to: i) Discover novel HPVs using Next Generation Sequencing (NGS) technology in oral rinse samples collected from oral squamous cell carcinoma (OSCC) and oropharyngeal squamous cell carcinoma (OPSCC) patients; ii) Determine prevalence of novel HPVs in archived OSCC/OPSC tissue samples; and iii) Examine frequency of novel oncogenic HPVs in cancer and non-cancer oral rinse samples using real-time PCR.

Methods

We collected 110 oral rinse samples from healthy patients and 100 oral rinse samples from patients with OSCC/OPSCC. Enrichment of HPV DNA was completed using multiply-primed rolling-circular amplification (MP-RCA) techniques. Fluorescent arbitrarily primed (FAP) PCR methods were used to isolate the L1 region of potential novel HPVs. NGS was used to detect for HPVs from 7-pooled samples that consisted of samples that underwent enrichment and FAP PCR. Potential novel HPVs were identified through cloning and Sanger sequencing methods. BLASTn and PaVE databases were used for nucleotide

searches. Phylogenetic trees were created to determine related HPVs and genus.

New primers and probes were created for the novel HPVs in order to test prevalence in 221 archived tissue biopsies and 210 oral rinse samples.

Results

We discovered three potential novel HPVs: NV14.4, NV69.1, and NV95. NV14.4 has 89% homology to HPV76; NV69.1 has 85% homology to HPV152; and NV95 has 77% homology to HPV147. From the archived tissue biopsy samples, only 0.8% of the OSCC patients were positive for NV14.4; NV69.1 and NV95 were not detected in the samples. Of the OPSCC oral rinse samples: 1% was positive for NV14.4; 13% was positive for NV69.1; and 1% was positive for NV95. Of the OSCC oral rinse samples: 6% was positive for NV14.4; 12.5% was positive for NV69.1; and 6% was positive for NV95. Of the other head and neck cancer oral rinse samples 12.5% was positive for NV69.1; NV14.4 and NV95 were not detected. None of the non-cancer samples in the tissue biopsy set and the oral rinse sample set were positive for the three novel HPVs.

Conclusions

Novel, potentially oncogenic, HPVs can be detected in oral rinse samples using NGS technology in conjunction with cloning and Sanger sequencing.

Table of Contents

List of Tables

List of Figures

Chapter I – Background & significance

The human papillomavirus, the oral cavity, and oropharynx

HPV is a circular, double-stranded DNA molecule consisting of approximately 8,000 base pairs [1]. More than 200 types of HPVs have been identified [2] and novel types continue to be added to the list [3]. HPVs are classified into mucosal and cutaneous categories, and into low and high-risk types according to their presence in malignant lesions of the uterine cervix [4]. HPV types 6 and 11 are examples of low-risk mucosal HPVs and cause genital warts, where as high-risk types such as HPV 16 and18 are considered oncogenic [4]. The E6 and E7 oncoproteins in HPV are responsible for deactivating p53 and Rb, respectively which are tumor suppressor proteins, and thus are essential for carcinogenesis [5-9]. Loss of p53 and Rb disrupts normal cell cycle regulation, which causes uncontrolled cell proliferation, inhibition of apoptosis, and genetic instability, thus resulting in formation of epithelial lesions of the skin or mucosa [5, 6]. Recent evidence shows that low risk types of HPV are not able to cause malignancy due to weak binding of E6 and E7 proteins to their cellular targets, as well as dissimilarities in mRNA splicing patterns and promoter positioning and regulation [10, 11].

The oropharynx includes the base of tongue, soft palate, tonsils, uvula, and posterior pharyngeal wall [12]. This area of the mouth is separate from the oral cavity, which includes the floor of the mouth, gingiva, buccal mucosa, hard palate, mobile part of the tongue, and lips [12]. When studies refer to oral cancer this usually includes both OSCC and OPSCC. In our research, we distinguish between the two cancers.

The oropharynx and cervix are similar in terms of easy access for infection, as well being both derived from endoderm [13]. As a result of microabrasion or foreign body infiltration, HPV will invade keratinocytes via the exposed basement membrane [14]. It is theorized that in the oropharynx,

infection of exposed crypt cells by HPV occurs by introduction to the basal layer of the tonsillar epithelium and basement membrane that has been disrupted [6].

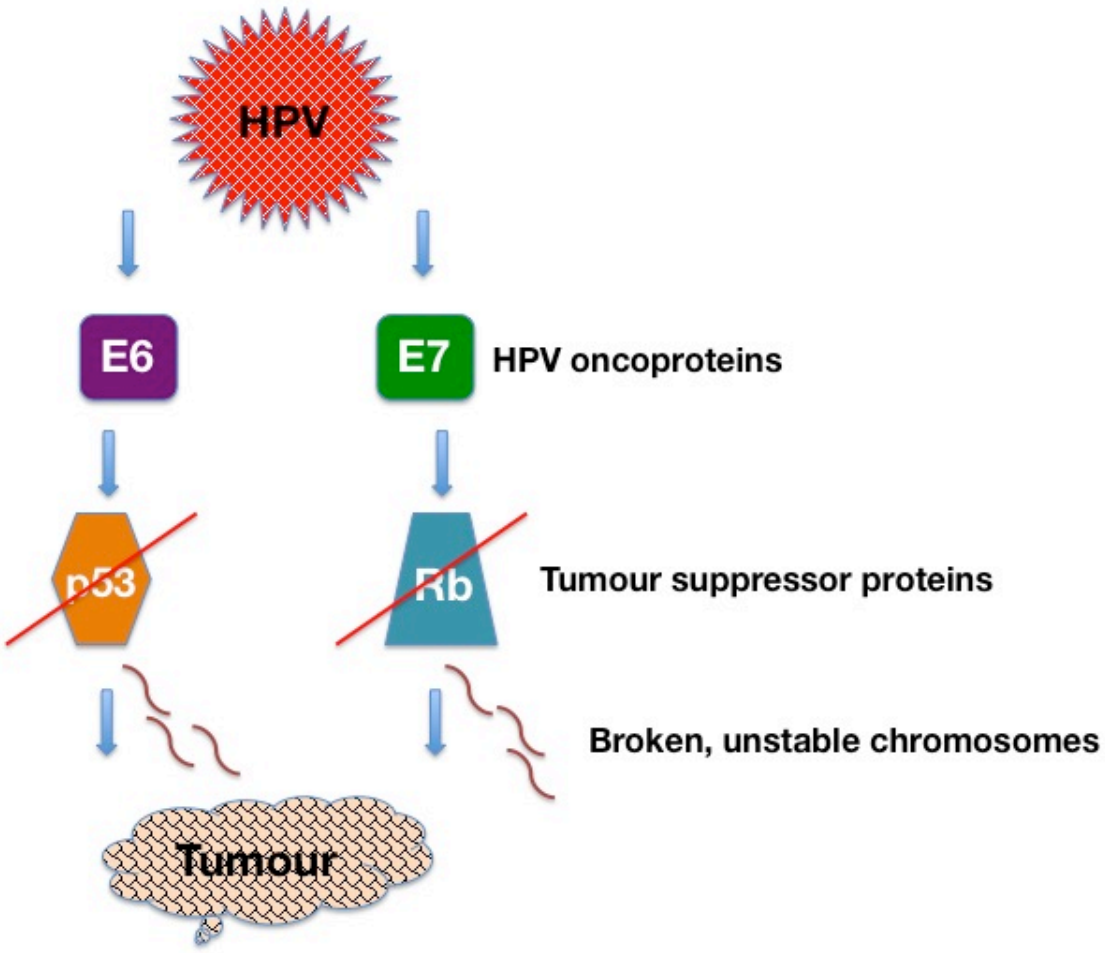

Figure 1.0 HPV infection and progression to oropharyngeal cancer. E6 will bind to p53 and inactivate its function; E7 will bind to retinoblastoma (Rb) and inactivate its function. The result are unstable, broken chromosomes, which could progress to oropharyngeal cancer.

Epidemiology

Worldwide, over 300,000 people will be diagnosed with OSCC and OPSCC – a disease that has a rate increase of >50% mortality rate each year [15, 16]. Even with some improvements in scientific efforts and screening, the mortality rate of OSCC remains high and the 5-year disease-free survival rate remains quite poor [17]. Well-established risk factors for OSCC include tobacco use and alcohol consumption [18]. However, recently, there has been an increase in the incidence of cancers arising in the oropharyngeal area [19] especially among younger individuals without the typical risk factors such as tobacco and alcohol use [19, 20]. HPV infection now has been identified as an etiologic agent for many such OSCCs, especially for OPSCCs [21-23].

A recent International Agency for Research on Cancer (IARC) review estimated that 25.6% of OPSCCs worldwide were associated with HPV infection, which is just over 21,000 people (~17,000 males, ~4,400 females) [24]. In a paper by Gillison et al., the proportion of HPV-positive OPSCCs for specific geographical regions was presented - 56% in North America, 52% in Japan, 45% in Australia, 39% in Northern and Western Europe, 38% in Eastern Europe, 17% in Southern Europe, and 13% for rest of the world [12]. Studies done within the US reported 29% of case patients with HPV-positive OPSCC [25], and 3.9% from OSCC cases [21]. From our most recent study we detected HPV, specifically type 16, in 25% (19/76) of OPSCC case patients, and 12.5% (2/16) of OSCC patients.

Chaturvedi et al. [26] looked at data between 1988 and 2004 in three different states. They reported that the population-level incidence of HPV-positive OPSCCs increased by 225%, while the incidence for HPV-negative cancers declined by 50% (n=271). The authors suggested that by 2020 the number of HPV-positive OPSCCs will surpass the annual number of cervical cancers. Complementary to these findings, data from Sweden also shows a parallel increase in both the incidence of tonsillar or base of the tongue SCC and the percentage of HPV-positive cases [27, 28]. HPV-positive tonsillar cancers

increased from 23% of the total in 1970-1979 [27] to 79% in 2007 [29]. And, for base of tongue SCCs an increased prevalence from 58% to 84% was seen, for HPV-associated cases, between 1998 and 2007 [28]. Similar trends of increasing incidence and prevalence of HPV-positive OSCCs have been noted in Australia and Canada [30, 31].

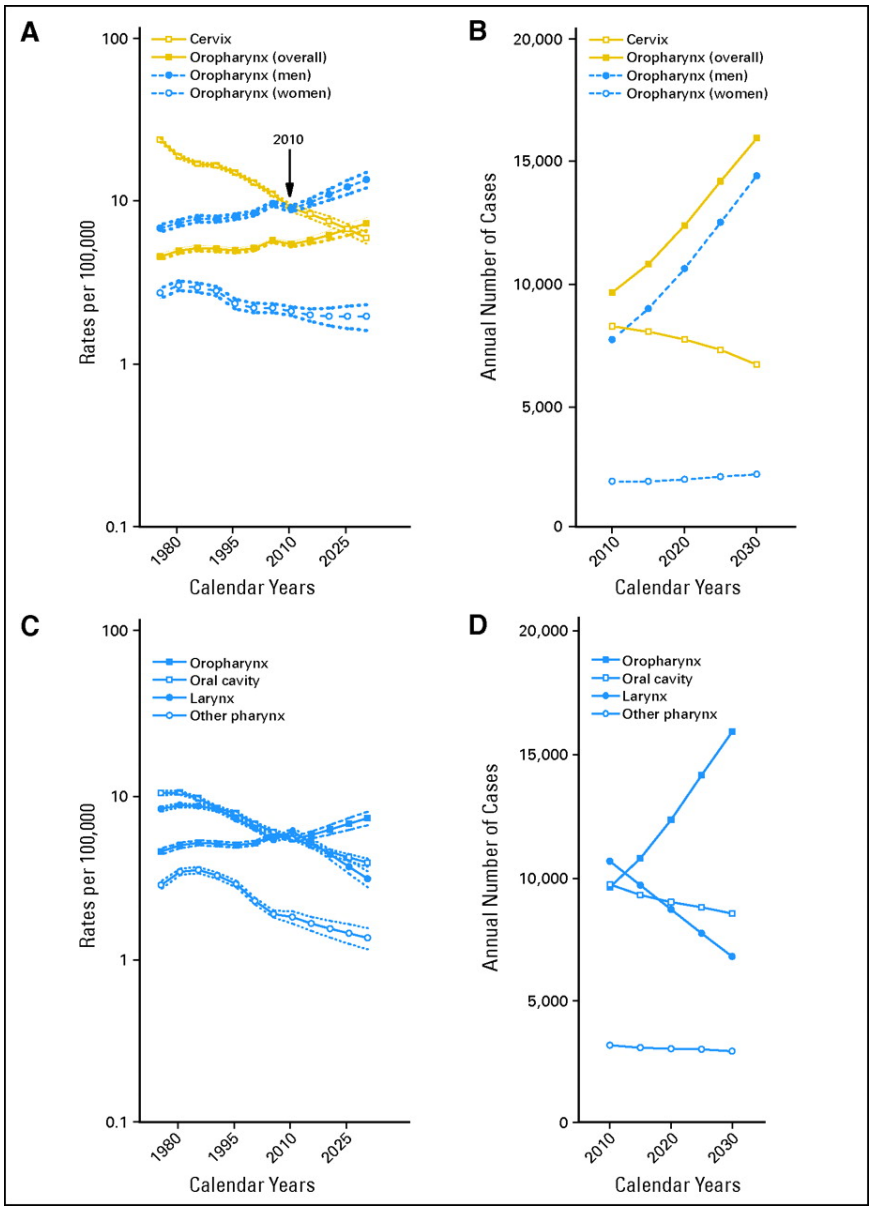

Figure 1.1 (Chaturvedi et al. 2011, *J CLin Oncol*). (A) Observed and projected incidence rates and bootstrap 95% CIs (ages 30 to 84 years) for oropharyngeal cancers overall (solid squares), oropharyngeal cancers among men (solid circles), oropharyngeal cancers among women (open circles), and cervical cancers (open squares). (B) Projected annual number of patients (ages 30 to 84 years) of oropharyngeal cancers overall, oropharyngeal cancers among men, oropharyngeal cancers among women, and cervical cancers through the year 2030. (C) Observed and projected incidence rates for oropharyngeal (solid squares), oral cavity (open squares), larynx (solid circles), and other pharynx (open circles) cancers. (D) Projected annual number of patients with oropharyngeal, oral cavity, laryngeal, and other pharynx cancers through the year 2030.

Risk factors

Internationally, over time, there have been changes in sexual behaviors with the age of sexual interaction beginning much younger, along with the increasing number of sexual partners one has [32-34]. As well, sexual practices have changed with oral sex being performed more by men and women within the 30-49 year age range compared with older adults [33]. As such, it is not surprising that HPV infection of the oral cavity is strongly associated with sexual behavior, but what is still perplexing is that it is twice as predominant in males than in females [12]. A multicenter study observed that the frequency of HPV detection in oral cancer biopsy specimens was higher among individuals who reported having more than one sexual partner or having oral sexual contact [21]. Furthermore, an increased risk of oral cancer has been reported both in women with cervical cancer and in spouses of women with cervical cancer [21, 35]. These results may collectively suggest a common mode of HPV transmission between the oral and genital area, reinforcing the possibility of sexual transmission of the virus in the oral cavity.

Recent studies demonstrate a non-sexual contact, which includes salivary transmission via deep kissing being linked to HPV infection [36, 37]. However, data on cervical HPV infection shows that 90% of HPV infections should clear within 1-2 years, with the other 10% having continual infection and an increase risk of developing SCC [38]. Potential reservoirs for the virus are proposed to be the tonsillar crypts [27, 39] as well as periodontal pockets [40]. There has even been an association seen between long-standing periodontitis and the risk of tongue cancers [41]. Thus, poor oral health remains a risk factor for HPV infection. As well, there seems to be a distinct risk profile for HPV-positive OPSCC compared with HPV-negative cancers, which includes factors such as being a young, white male, with a high number of sexual partners, and history of marijuana use [23, 42].

One study showed that former smokers had over two times higher a risk of acquiring oral HPV infection, and current smokers had almost three times the

risk [43]. A very recent large cross-sectional study by Fakhry et al. presented a statistically significant dose-response relationship between current tobacco use and oral HPV16 infection [44]. Their results showed that oral HPV16 prevalence was higher in current tobacco users compared with never or former tobacco users (n=6,887, p=0.004).

HPV association

The most prevalent type of HPV associated with oral infection is type 16 [45, 46]. Data from studies that included at least 30 cases, found HPV 16 in 80-100% of HPV-positive OPSCC patients [47], and type-specific carcinogenicity has been evidenced for HPV 16 [48]. In contrast, the relative contribution of HPV 16 in cervical cancer is ~61% [49]. HPV types 18, 31, 33, 35, 52, and 58 are additional HPV types hypothesized to be associated with OPSCC [26, 50].

Natural history

Compared with head and neck squamous cell carcinoma (HNSCC), little is known about the natural history of oral HPV infections [51]. Fakhry et al. showed that sampling carried out at 6-month intervals was deemed to be suitable for future natural history studies of oral HPV infection [52]. OPSCC involves transformation of benign tumors where normal epithelium progresses to dysplasia, in situ carcinoma, and then to invasive carcinoma [51]. In 10 different studies, HPV status was analyzed in benign, dysplastic, and invasive carcinoma lesions. Overall, the odds of detection of HPV in both dysplastic lesions and invasive carcinoma were 3 times as high compared to tissue without dysplasia or cancer [51]. No difference was seen in the odds of detection between mild, moderate, and severe dysplasia [53]. Thus, what the authors are proposing is that HPV is found in early precancerous lesions of the oral cavity, which is comparable to cervical cancer.

Pathogenesis

The information that we have on the pathogenesis of cervical cancer is what is used to help create a model for OSCC and OPSCC. Studies on the epidemiology and molecular biology of OSCC/OPSCC assist in providing evidence for the causal relationship of oncogenic HPVs [51]. For example, the observation that short hairpin RNA-mediated inhibition of HPV 16 E6 and E7 expression leads to the restoration of p53 and Rb tumor suppressor pathways, and thus results in apoptosis, provides proof that HPV is directly oncogenic in oral cancer [54, 55]. In HPV-associated oral cancers, the HPV genome can be found integrated in the host genome or in its episomal form [56]. In tonsillar cancer specifically, the HPV genome has been found in its episomal form in 40-100% of cases [56, 57]. An explanation for how HPV remains in the tissue in episomal form and stays transcriptionally active is not available, but could involve the E2 protein binding episomal HPV to cellular mitotic spindles [58].

p16 is a cellular tumor suppressor protein that is often overexpressed in OPSCC due to HPV activity [51]. It has been shown that HPV-positive oral cancers are associated with wild-type p53, low Rb levels, and p16 overexpression, whereas in tobacco- and alcohol-associated oral cancer, mutated p53 and high Rb levels are present; and as a result of point mutations, promoter methylation, homozygous deletion, and low expression of p16 are also found [55, 59-61].

Prevention, detection, and treatment

There is currently not a gold-standard for oral HPV detection [62]. The most common methods for HPV detection within the mouth and oropharynx begin with collection of cells with a cotton swab, cytobrush, or a mouth rinse [63], followed by the use of PCR-based assays or DNA in situ hybridization [64]. However, there are challenges with certain techniques. For example, the use of a swab/brush limits the amount of mucosa that is sampled, and obtaining a sample from a non-visible lesion within the tonsillar crypt may not be feasible

[65]. The base of the tongue is not entirely accessible either as there is both flat mucosa and tonsillar tissue, thus increasing the risk for false negatives [66]. We chose to use a mouth rinse technique for sample collection as it is non-invasive, quick, and simple for the patient.

During the diagnosis of OPSCC/OSCC, it is rare for patients to receive HPV testing even when they do not have the traditional risk factors (ie. smoking and drinking). With the rising numbers of OPSCC it is imperative that we have oral HPV infection information, which would assist in providing a more accurate picture for all patients as smokers could also have HPV infection, which may be increasing their risk of cancer.

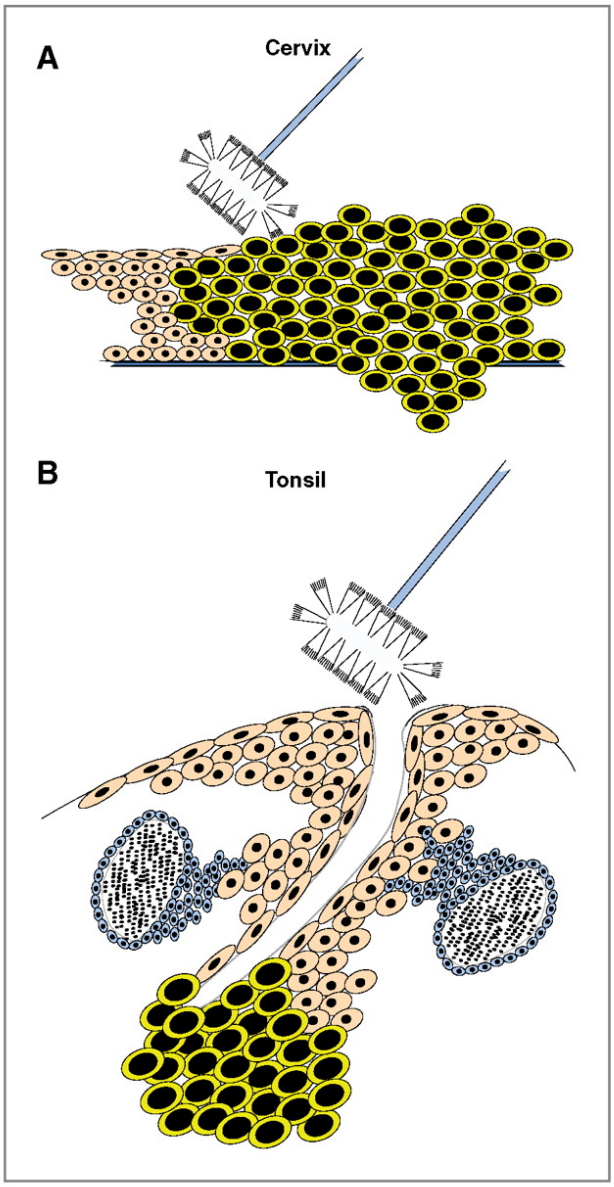

Figure 1.2 (Lingen, *Cancer Prev Res*. 2011. 4(9):1350-1352). Anatomic differences in cervical and tonsillar mucosa affect the ability of the cytobrush to collect premalignant or malignant epithelial cells. A, the cytobrush is able to obtain an adequate and representative collection of atypical or cancer cells from a relatively flat and uniform cervical mucosa. B, the cytobrush has difficulty in obtaining an adequate and representative sample of dysplastic or malignant tonsillar mucosa, particularly when these lesions are lurking deep within a tonsillar crypt (running diagonally from top of tonsil to lower left), which would not be sampled using a conventional cytobrush. The large, oval structures to the left or right of the crypt are lymphoid follicles, with lymphoid cells (blue) emanating from them.

Analyzing p16 expression has been used as a biomarker for HPV-associated OSCC/OPSCC, but studies have reported that p16 overexpression is not always present in cases involving oncogenic HPVs [67-70]. In fact, a recent study concluded that p16 should not be used as a surrogate marker for HPV infection in oral cancers due to poor concordance between the two [71]. These findings support an earlier study which indicated that p16 immunohistochemistry (IHC) alone is not a reliable method for HPV detection in OSCC/OPSCC cases [17].

Though there are differences in anatomy, the standard treatments for OPSCC and cervical cancer share similarities [12]. Clinical stage at diagnosis is what determines primary therapy. For early stages in either cancer, surgical resection is common, and the prerequisite for adjuvant postoperative radiotherapy [72, 73] or chemoradiotherapy [74-76] is based on observations of histology from the resection. In confined OPSCCs, primary radiotherapy by itself is comparable to early stage surgery, but this method is not effective in cervical cancers [77]. For both OPSCC and cervical cancer cases involving large primary tumors or regional nodal metastasis, the preferred treatment is primary chemoradiotherapy. In recurrent or metastatic cases, palliative platinum-based combination chemotherapy (ie. cisplatin and paclitaxel) is the standard of care. Even with these available treatments, there still is not a specific therapy for HPV-associated malignancies, which would target viral oncoproteins. Nor, is there a primary preventative or validated screening method in place [12].

A study published recently provides evidence that HPV vaccines used to prevent cervical cancer can be effective in preventing infection of oral HPVs. In a randomized clinical trial in Costa Rica, 7,466 women between the ages of 18-25 years were given the HPV16/18 vaccine or hepatitis A as a control [78]. During the last visit of the blinded 4-year study, 5,840 subjects gave oral specimens in order to evaluate vaccine efficacy (VE) against oral infections. The results demonstrated a VE of 93.3%. The authors theorized that if protection was seen

in females the same should occur in males, and that this could be a primary preventative measure for HPV-related oral cancers.

Hypothesis

We hypothesize there are unidentified HPVs in OSCC/OPSCC patients, and oral rinse sample collection is an effective method for identifying novel HPV types.

Specific Aims

1. Discover novel HPVs using NGS technology in oral rinse samples collected from newly diagnosed and untreated OSCC and OPSCC patients.

2. Determine prevalence of novel HPVs in archived OSCC/OPSCC tissue samples.

3. Examine the frequency of novel oncogenic HPVs in cancer and non-cancerous oral rinse samples.

Chapter II – Review of methods

Sample collection and purification protocols

Enrolling participants for studies where collection of bodily fluids is involved can be a daunting task. The healthy population seem to be less willing to participate in studies especially when they are asked to produce an oral rinse sample and answer health questions. Patients in general do not have a lot of time available when visiting the doctor. Thus, it is imperative to have a sampling technique that is easy, quick, non-invasive, and inexpensive with a simple questionnaire. And, one huge factor is being able to obtain high quality DNA from such samples.

Heath et al. compared 5 different mouthwashes (Cepacol, FreshBurst Listerine, Listermint, Scope, and Saline) for obtaining buccal cells for DNA clinical testing [79]. There were 5 factors used to evaluate the brands of mouthwash: 1) compatibility with the DNA purification chemistry, 2) DNA yield, 3) DNA quality, 4) DNA stability at room temperature, and 5) mouthwash taste. For the first aim, purification chemistry compatibility was determined by observing the protein precipitation step. The supernatant formed contains the DNA and it is best if there are no contaminating dyes as this could interfere with downstream analyses. Of all the mouthwashes saline ranked best with Scope at a close second, and Listerine third. When comparing DNA yield (aim 2), standard UV quantitation was used. Scope was ranked first followed by Cepacol and Listerine. The same results were observed when testing DNA yield by Quantitative PCR. An amplification assay was used to test DNA quality, and all 5 mouthwashes were given top scores as there were no significant differences. For aim 4, aliquots from each mouthwash were taken at 0, 4, and 7 days, and DNA was isolated and analyzed using gel electrophoresis. Scope demonstrated the most consistent high-molecular-weight DNA during the 7 days with no observed degradation.

How palatable a mouth rinse is, is actually quite a significant factor in an oral rinse study. If subjects do not find the oral rinse taste appealing they will be less willing to keep it in their mouth for the full amount of time requested. In aim 5, Listerine ranked best for taste followed by Scope, and Cepacol.

Overall, Scope had the highest ranking out of all the mouthwashes, thus the authors used Scope for further investigation to determine an ideal technique for DNA yield. Proteinase K is used to increase DNA yields in tissue samples and other high-protein specimens, and glycogen is often used as a co-precipitant to enhance the precipitation of DNA when alcohol is present. Four treatment groups were created with interchanging absence and presence of proteinase K and glycogen. The results showed that combining proteinase K and glycogen produced a considerably higher yield than any of the other treatment groups. Thus, this study presented very good evidence that Scope brand mouthwash, a commercially available product, can be used in clinical studies for analysis of human DNA that is both effective and can produce valuable results.

When comparing levels of inhibition in DNA detection, viral nucleic acid has a tendency to be affected more when compared to human DNA, which is only slightly affected; and DNA yields may be variable as well [80]. The variability in DNA yield is dependent on degree of infection and shedding, which can affect the sensitivity of detection. There is quite a bit of inconsistency between studies on the extent and whether or not oral HPV infection is associated with OSCC/OPSCC, and these differences could be due to sample processing methodology. D'Souza et al. performed a study where they compared and analyzed 5 different DNA purification methods used for detection of HPV in oral rinse samples [80]. With regards to the materials and methods, an oral rinse sample was collected by rinsing and gargling for 30s with 10ml of Original Mint Scope® mouthwash, and was stored at 4°C until further processing (48h maximum). A second oral rinse was collected 7 days after the first visit to measure concordance. Participants of the study were HIV-positive as this population has an increased risk for tonsillar cancer [81] and tonsillar HPV

infection [82]. The oral rinse specimens were transferred into 15ml tubes and centrifuged at 3,000 x g for 10 min. at 4°C. The supernatant was decanted and resuspended in 10 ml phosphate-buffered saline (PBS) and centrifuged again. PBS was chosen due to its compatibility with all subsequent DNA purification methods. The pellet was resuspended in 6ml of PBS and aliquoted equally into four tubes, and stored at -80°C until further processing. The 5 different DNA purification methods used in the study were: i) Puregene DNA purification kit, ii) phenol/chloroform extraction, iii) QIAamp DNA blood midi kit, iv) proteinase K digestion, and v) proteinase K digestion with ethanol. The last two methods listed were included because they are often used to extract DNA from cervical vaginal rinse specimens for HPV detection.

The authors' data proposes that the DNA purification protocol used on oral rinse samples can significantly affect results when detecting for HPV genomic DNA using PCR. It was also seen that PCR inhibition is a common problem for oral rinse samples. When comparing each of the purification techniques, the Puregene protocol demonstrated a much greater human cell yield. In addition, the number of subjects who were found to have oral HPV infection was considerably higher. The Qiagen kit and phenol-chloroform methods both produced a substantial loss of human DNA, which seemed to affect the ability to detect HPV DNA. In contrast, Puregene maintained high DNA purity while preserving human DNA yield, thus providing better results. This study showed that HPV prevalence could be underestimated in studies reporting results for unpurified or chemically contaminated oral exfoliate samples.

The results for samples collected on day 1 and on day 7 had a low concordance, which may be due to sampling difference factors such as the time the subject last brushed his teeth or ate, or could indicate that oral HPV infection can be dynamic in an immunocompromised patient population. Another finding of interest was that infections with multiple types of HPV was seen to decrease multiplex assay sensitivity, which may be due to competition for primers by other HPV types with higher viral loads.

The authors propose that oral HPV prevalence may have been significantly underreported by as much as six-fold for studies not including further DNA purification after protein removal, and studies only using ethanol precipitation or phenol-chloroform extraction could have underestimated prevalence by 40-75%. Thus, this study emphasizes the significance of DNA purification in order to prevent the misclassification of HPV status (ie. false-negative results) in oral rinse samples, and emphasizes that misclassification is very much dependent on the purification techniques used.

PCR and MP-RCA

PCR is considered to be of the highest sensitivity and can detect even a single copy of viral DNA per infected cell [83]. It is the most established method today for HPV detection in oral rinse samples [84]. However, for discovery of novel viruses, PCR relies on prior knowledge of a sequence and produces short amplicons [85] thus only having the ability to find closely related viruses [86]. HPV being a circular DNA molecule can be amplified using a rolling-circle mechanism, which makes rolling-circle amplification (RCA) a suitable technique [86]. In multiply primed (MP)-RCA the polymerization process is primed by exonuclease-resistant random hexamers that bind at multiple locations on the circular template DNA, thus creating multiple replication forks [87]. With the use of random hexamer primers, custom primers and information of the sequence are not required for amplification. The bacteriophage Φ29 DNA polymerase, a high-fidelity enzyme with a strong strand-displacing capability, high processivity (>70,000 bases per binding event), with great stability and proofreading activity, is the enzyme used in MP-RCA [88]. MP-RCA has been demonstrated to amplify circular DNA templates up to 10^7-fold [89].

A major drawback in the process of discovering novel HPVs is the lack of a conventional cell culture system for in vitro viral reproduction [86] . In the past new cutaneous HPVs were discovered with the use of degenerate-primer PCR

methods where primers were created using previous knowledge on the L1 gene since it is the most conserved region of the papillomavirus genome [86]. Because this technique is restricted to discovering HPVs that are related to previous types, Rector et al. tested to see if complete circular double-stranded DNAs of papillomaviruses could be amplified without any need for prior knowledge of their sequences using MP-RCA. HPV 16 genomic DNA was extracted from the human cervical keratinocyte cell line W12, which had ~100 copies of HPV 16 mostly in episomal form [90]. For the unknown papillomavirus genotype, a biopsy of a fibropapillomatous wart from a bovine udder was the source of the sample. For the W12 cells, DNA was extracted using the QIAamp DNA blood minikit (Qiagen), and DNA from the bovine tissue was isolated using the phenol-chloroform-isoamyl alcohol extraction method. MP-RCA was performed with the TempliPhi 100 amplification kit (Amersham Biosciences). The complete papillomavirus genomes were isolated, cloned, and analyzed by agarose gel electrophoresis. Results showed that using the unmodified TempliPhi kit, according to the manufacturer's protocol, produced quite a low amplification efficiency (172-fold). The use of a more diluted and larger circular DNA molecule as input material is apparently less efficient than the pUC18 DNA control. The authors added an extra 450μM of dNTPs and were able to achieve an amplification efficiency of up to 2.4×10^4-fold. This was the optimal amount found to improve amplification efficiency as too many dNTPs could interfere with the sequencing reaction. The unknown papillomavirus was found to have a 99% homology with bovine papillomavirus-1 (BPV-1) isolate 307. The authors concluded that their study presented a promising technique for amplification and isolation of novel human and animal papillomaviruses without the prior knowledge of their sequences.

FAP PCR is a method in which degenerated primers, created from conserved L1 regions of HPV, are used to detect for a broad range of HPVs [91]. A band seen at approximately 480bp indicates a positive sample.

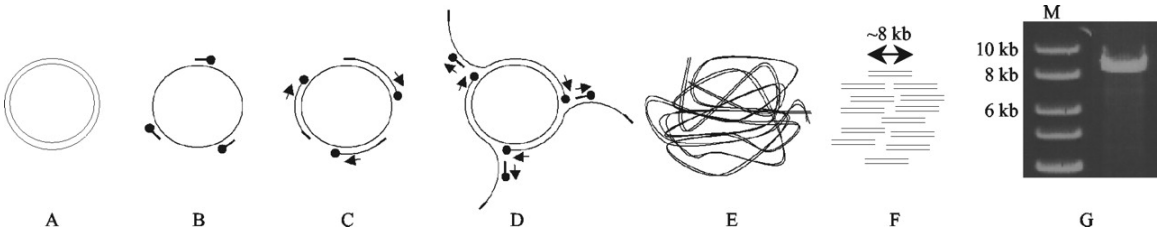

Figure 2.0 (Rector et al. 2004, *J Virology*). Schematic representation of the MP-RCA method for amplification of the complete circular double-stranded DNA genomes of PVs (A). In a first denaturation step, the PV episomal DNA is rendered single stranded. Exonuclease-protected random hexamer primers (—) can now anneal to multiple sites on this template DNA, after which the ϕ29 DNA polymerase (•) binds (B) and isothermally extends these primers at the 3′ end (→) (C). Strand displacement synthesis occurs when the DNA polymerase reaches a downstream extended primer, and hexamer primers can anneal to the displaced single-stranded product strands and will again be elongated by the ϕ29 DNA polymerase (D). Continuation of this process results in exponential amplification of the template DNA, generating linear double-stranded, high-molecular-weight repeated copies of the complete PV genome (E). Digestion of this multiply primed RCA product with a restriction enzyme which has only a single recognition site in the PV genome will result in multiple double-stranded, linear copies of the PV complete genomic DNA (F), which can be visualized as a single band of ca. 8 kb by agarose gel electrophoresis (G). Lane M, DNA molecular size marker (Fermentas).

Whole genome sequencing

Sanger sequencing technology was introduced in 1977 and was the 1st generation of sequencing technology available [92]. It was also the technology used for the human genome project [93], which was completed in 2001. The next generation sequencing (NGS) platforms developed after the Sanger technology, perform better in terms of read lengths, applications, manpower, and efficiency. Three major NGS platforms are Roche 454, AB SOLiD, and Illumina GA/HiSeq. The 454 GS FLX from Roche carries out a pyrosequencing mechanism which is fast with long read lengths, however the error rate is higher compared with the other two platforms; also, cost is high and throughput is low. The HiSeq2000 (Illumina) sequences via synthesis and has a high throughput and a low reagent cost, but creates short assembly reads. SOLiDv4 (AB) uses a ligation and two-base coding mechanism for sequencing, which is very accurate, but also has a short read assembly. Thus, the type of NGS platform that is chosen is dependent on a variety of factors.

The molecular diagnostic tools available for HPV detection are mainly PCR-based, which are targeted towards known HPV types and those prevalent in the developed world [94]. Such tests are not useful for detecting rare or novel HPVs. NGS technologies have made it possible to study the variety of viruses in clinical samples without knowledge of the sequence beforehand [95]. Meiring et al. used the Illumina sequencing platform to detect and genotype the HPV types present in a complex multiple infection of a cervical specimen taken from an HIV-infected South African woman [94]. The Roche Linear Array (LA) HPV genotyping detection kit was used on the same specimen in order to compare if certain HPVs were detectable with a readily available diagnostic tool. The Roche LA detected twelve HPV types in one sample, which was selected to undergo NGS. The sample was enriched using a randomly primed RCA technique before being sequenced by the Illumina GAII system. A total of 9,818,116 short sequence reads (SSRs) of 76 nucleotides (nt) in length were obtained and were trimmed with FASTx-Toolkit to 41 nt. Following the trim, *de novo* assembly with

Velvet and CLC Genomics Workbench was carried out where human DNA was mapped and excluded from further analyses. Four contigs were the result of the *de novo* assembly, and full-length HPV genomes were characterized with top BLAST hits to HPV types 39, 40, 16, and 56 with type 39 having the highest copy number. The genomes for HPV types 30, 39, 40, 16, and 56 were completely assembled, where the less abundant types did not have the ability to undergo *de novo* assembly. Sixteen HPV types were found through Illumina sequencing compared to the twelve types with Roche LA. The authors concluded that the use of RCA and Illumina sequencing decreases many of the difficulties connected to PCR-based HPV detection methods, which require prior knowledge of the genome. Other problems such as false positives due to cross-reactivity between types, false negatives related to low viral loads, and biased amplification, which creates difficulty in identifying all the types in multiple infections [96-98]. Since the study was completed there has been much improvement with NGS, which include paired-end sequencing, multiplexing and increased sequence output.

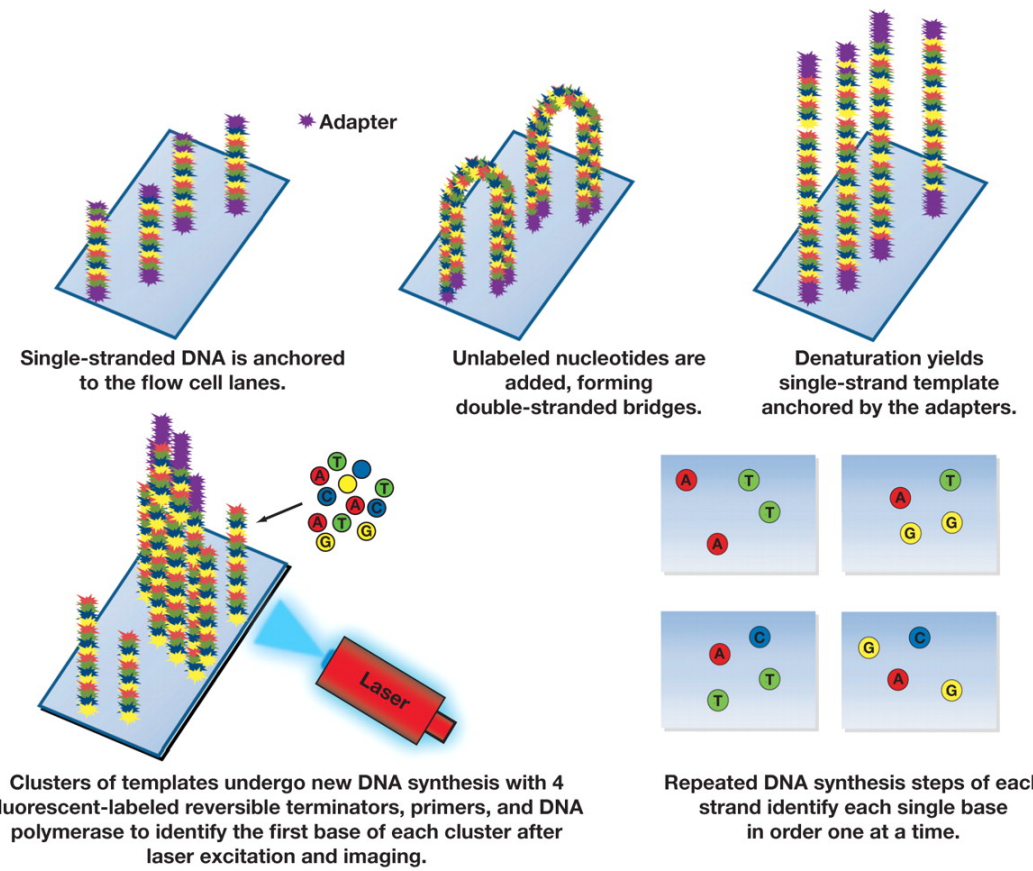

Figure 2.1 (Ross and Cronin, *American Journal of Clinical Pathology.* 2011. 136(4):527-39). Whole cancer genome sequencing on the Illumina platform.

Pilot study on oral rinse samples

We performed a pilot study in order to determine whether HPV types 16 and 18 could be detected in 19 oral rinse samples collected from 15 HIV positive Senegal women. 11 were positive for HIV-1 and 4 were positive for HIV-2. Of matched cervical swab samples, eight were positive for HPV 16, one was positive for HPV 18, and ten were negative for both HPVs. There were 4 patients that contributed 2 samples each for a total of 19 samples. Quantitative Taqman real-time (RT)-PCR assays with primers specific to the E6/E7 region of the HPV genes was used for detection. The total volume was a 5μl reaction containing 1μl DNA template, 2.5μl Taqman Universal Master Mix, 1μl reverse and forward primers (3μM), 0.1μl probe (10μM), and 1.4μl dH2O. RT-PCR was run on the ABI Prism 7900 Sequence Detection System with ~40 cycles in a reaction (denaturation at 95°C, annealing at 50°C, and extension at 72°C).

Absolute quantification was determined on the oral rinse samples for both HPVs expressed as HPV copy number/μl, and the Alu gene. Serial dilutions of human genomic DNA, and the E6/E7 regions of HPV 16 and 18 with concentrations of 10^1-10^6 copies/μl, were used as standard curves.

Only one sample came out positive for HPV, which was type 18. From our results we concluded that oral rinse samples used for the detection of oncogenic HPVs was a non-invasive and effective method. Limitations to our study include small sample size, archived samples were used (1995-1998) in which DNA degradation may have occurred, and we only detected for HPVs 16 and 18 but not other high-risk types being detected. Finally, HPV prevalence difference may represent the difference of different study populations.

Characterization of four novel HPVs from healthy individuals

In our previous study, we isolated four novel HPVs from oral rinse samples collected from healthy individuals using RCA coupled with degenerated PCR assay [99]. Full-length HPV DNA was cloned for three of the four novel HPV types using long range PCR. Our samples were obtained from a longitudinal

study investigating the natural history of HPV infection in the male population. 41/48 archived oral rinse samples were selected for isolation of novel HPV types. Participants were asked to rinse and gargle with 10ml of Scope® mouthwash (Proctor&Gamble) for 30 seconds, the sample was centrifuged and the cell pellet was resuspended in 1ml of STM [100]. Genomic DNA was isolated using the QIAamp DNA mini kit (Qiagen), and MP-RCA was used on each sample using the TempliPhi 100 amplification kit (GE healthcare). We modified the MP-RCA technique in order to optimize papillomavirus amplification, which used the following conditions: 1µl of DNA (~50-100ng of sample DNA) was mixed with 5µl of sample buffer, heated at 95°C for 3 min., then cooled to 4°C; a 5µl mixture was added to each cooled sample containing approximately 5µl reaction buffer, 0.2µl enzyme mix and 0.047µl of 25mM dNTPs; incubation of the samples took place at 30°C for 16 hours followed by heating for 10 min. at 65°C and cool down to 4°C; samples were stored at -20°C.

Four consensus PCR assays [91, 101-103] were performed on the MP-RCA amplified samples, and agarose gel electrophoresis was used to determine correct size of PCR product. The FAP PCR protocol detected the four potential new HPV types. The PCR product was cloned using the TA cloning kit (Invitrogen) and at least two clones from each PCR product were sequenced. BLASTn searches were used to determine the presence of HPV sequences. Primers were created using the cloned HPV fragments for long range PCR using the Stratagene kit. Three new HPV types underwent cloning, and sequencing of the HPV PCR product (~6-7kb) was then carried out. A Bayesian phylogenetic tree was created by BEAST v1.6.2 [104]. The three fully cloned new types are as follows: HPV 171, which was most homologous to HPV 169 (88%); HPV 172 with homology to HPV 156 (70%); and HPV 173 with most homology to HPV 4 (73%). OSL 37, which was not fully cloned had a 69% homology to HPV 144.

The prevalence of the novel HPV types was also determined in archived oral tissue blocks. 158 blocks were selected from the Department of Pathology at The University of Washington, which included 76 normal oral tissue blocks (56

from oral cavity and 20 from oropharynx), 82 malignant oral tissue blocks (66 from patients with OSCC with biopsies from the oral cavity, 16 from patients with OPSCC with biopsies from the oropharynx). DNA was isolated from 80μm tissue block sections using RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE Tissues (Applied Biosystems) using the manufacturer's protocol. DNA concentration was determined by UV spectrometer. Type-specific Taqman assays were created for the novel HPVs based on the E6/E7 region. RT-PCR was performed with the ABI Prism 7900 Sequence Detection System. Absolute quantification was performed on the archived tissue samples for each new HPV and Alu, and expressed as copy number per 100 cells. Serial dilutions of human genomic DNA as well as full-length HPV plasmids of known concentration were used as standard curves.

All analyses were conducted using SAS version 13.1. In our previous study, we found that HPV 171 was detected slightly more in malignant tissue samples (21% vs 13%, $p=0.21$), HPV 172 was only detected in normal tissue samples (33% vs 0%, $p<0.0001$), and HPV 173 was rarely detected in the samples [99]. The prevalence of HPV 16 and 18 was also determined. HPV 16 had similar prevalence in both normal and malignant tissue samples (11% vs 13%, $p=0.67$), but the average viral load in malignant samples was notably higher than in normal samples ($p<0.0001$), even after adjusting for age ($p=0.0005$). This same trend was seen in HPV 171 in malignant samples ($p=0.01$), but not after adjusting for age ($p=0.12$). We concluded that our study further confirms that the oral cavity harbors many unknown HPV types, mainly of the gammapapillomavirus genus.

Chapter III – Materials & Methods

Specific Aim 1a. Discovery of novel HPVs using high throughput sequencing technology in oral rinse samples collected from newly diagnosed and untreated OSCC/OPSCC patients:

We collected 100 samples from OSCC/OPSCC and 110 samples from the normal healthy population. Each patient answered a simple health questionnaire, and rinsed and gargled for 30 seconds with Original Mint Scope® mouthwash. Four control patients requested to use Crest® Alcohol Free mouthwash due to a history of alcoholism. Oral rinse samples were centrifuged for 15 min. at 4°C to form a pellet, the supernatant was discarded, and the pellet was placed in -80°C until further processing. The Puregene® DNA Purification Kit for purification from buccal cells mouthwash was used to isolate genomic DNA and manufacturer's protocol was followed. Absolute quantification was determined for the isolated DNA samples for HPVs 16 and 18 expressed as HPV copy number/µl, and the Alu gene by RT-PCR. Serial dilutions of human genomic DNA, and the E6/E7 regions of HPV 16 and 18 with concentrations of 10^1-10^6 copies/µl, were used as standard curves.

HPV16 E7 Primers
Forward: CGGACAGAGCCCATTACAATATT
Reverse: CGCACAACCGAAGCGTAGA

HPV16 E7 Probe: TAACCTT(T/C)TGTTGCAAGTGT

HPV18 E7 Primers:
Forward: CCGACGAGCCGAACCA
Reverse: TGGCTTCACACTTACAACACATACA

HPV18 E7 Probe: AACGTCACACAATGTT

Alu Primers

Forward: GGCCAACACGGTGAAACC

Reverse: CCACGCCCGGCTAATTTT

Alu Probe: CGTCTCTACTAAAAATAC

Taqman Assay:

2.5µl Taqman master mix

1µl Primer (forward + reverse, 3µM)

0.1µl Probe (10µM)

0.4µl $MgCl_2$ (50µM)

1µl DNA template

Total = 5µl reaction

From the RT-PCR results, specific samples were selected to undergo MP-RCA in order to preferentially amplify unknown HPV DNA. In the MP-RCA protocol, which was optimized for papillomavirus amplification [86], we used 1µl of sample and added 5µl of sample buffer; using a thermocycler samples were heated for 3min at 95ºC then cooled to 4ºC; 5µl of TempliPhi premix (5µl reaction buffer, 0.094µl of 25µM dNTPs, 0.2µl enzyme mix) was then added to the sample. Reactions were incubated for 16hrs at 30ºC and then for 10min at 65ºC followed by cooling to 4ºC. HPV DNA was preferentially amplified by MP-RCA and RT-PCR Taqman assays were used to verify this using the HPV16 gene and Alu house-keeping gene.

FAP PCR was performed on the amplified samples and PCR products determined by agarose gel electrophoresis. All 100 of our oral

rinse samples underwent this process of validation. Samples positive for HPV DNA and ones suspected were pooled and sent in two different rounds of NGS. First round: 1) FAP PCR positive, 2) HPV 16 positive RCA, 3) FAP PCR positive RCA, 4) OPSCC RCA. Second round: 1) FAP PCR RCA, 2) FAP PCR RCA2, 3) FAP PCR RCA16. The Illumina HiSeq2500 platform was used to perform whole genome sequencing.

FAP primers:
NBK1026 (FAP64 without 5' tag)
CCWATATCWVHCATITCICCATC

NBK1025 (FAP59 without 5' tag)
TAACWGTIGGICAYCCWTATT

FAP PCR master mix:
5µl 10x PCR buffer
7µl 25µM $MgCl_2$
1µl 25 µM dNTPs
3.8µl NBK1025 (10µM)
3.8µl NBK1026 (10µM)
0.25µl AmpliTaq Gold enzyme
27.15µl dH2O
2µl DNA template
Total = 50µl

FAP PCR thermocycler protocol:
1. 94ºC for 10min
2. 94ºC for 1.5min
3. 50ºC for 1.5min
4. 72ºC for 1.5min
5. Go to step 2 for 44 cycles

6. 4ºC forever

Samples were run on a 1.5% agarose gel (1.0g agarose + 0.5g Nusieve) with 3µl ethidium bromide for 1hr 10min at 110V. Expected band size is ~480bp.

As for supervised assembly, a total of 189 HPV L1 gene sequences and whole genome sequences were downloaded from the papillomavirus knowledge source at http://pave.niaid.nih.gov/. This data set includes a number of Non-reference genomes. Shot reads were aligned to the L1 gene region using BWA (http://bio-bwa.sourceforge.net/, version 0.7.12) with default settings. Subsequently, the whole genome sequences were used as the reference database if majority of the short reads did not map to the specific L1 gene region.

De novo assembly strategy was applied for cases where the majority of the short reads did not align to HPV genome sequences at all. The short reads were first aligned to the human genome UCSC hg19 (http://support.illumina.com/sequencing/sequencing_software/igenome.html), and then unmapped reads were selected to run Velvet (https://www.ebi.ac.uk/~zerbino/velvet/, version 1.2.10) to construct high quality unique contigs. Contigs (Fig. 3.0) with 200 bp or larger in length and with a minimum 100 coverage were aligned (blastn) to both HPV Specific database at http://pave.niaid.nih.gov/#search/pv_specific_blast, and to NCBI nucleotide collection (nt) database at http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome.

BLASTn and PaVE were used to determine homology of the sequences to HPV. A homology of <90% of the L1 region indicates a novel HPV [2, 105]. Figure 3.1 demonstrates the process of how to submit a novel HPV.

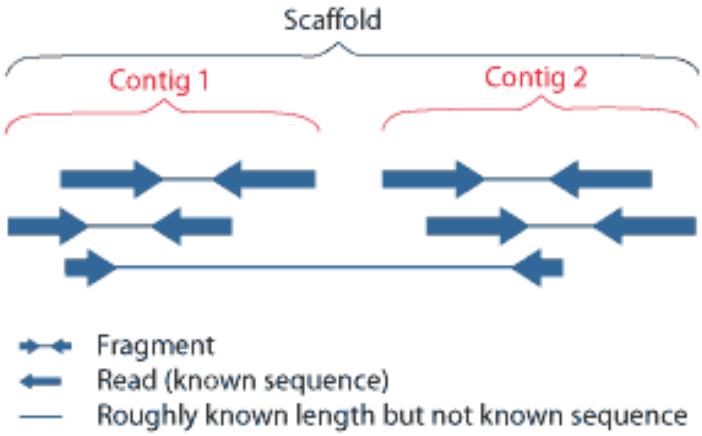

Figure 3.0 Schematic of contig formation (http://genome.jgi.doe.gov/help/scaffolds.html).

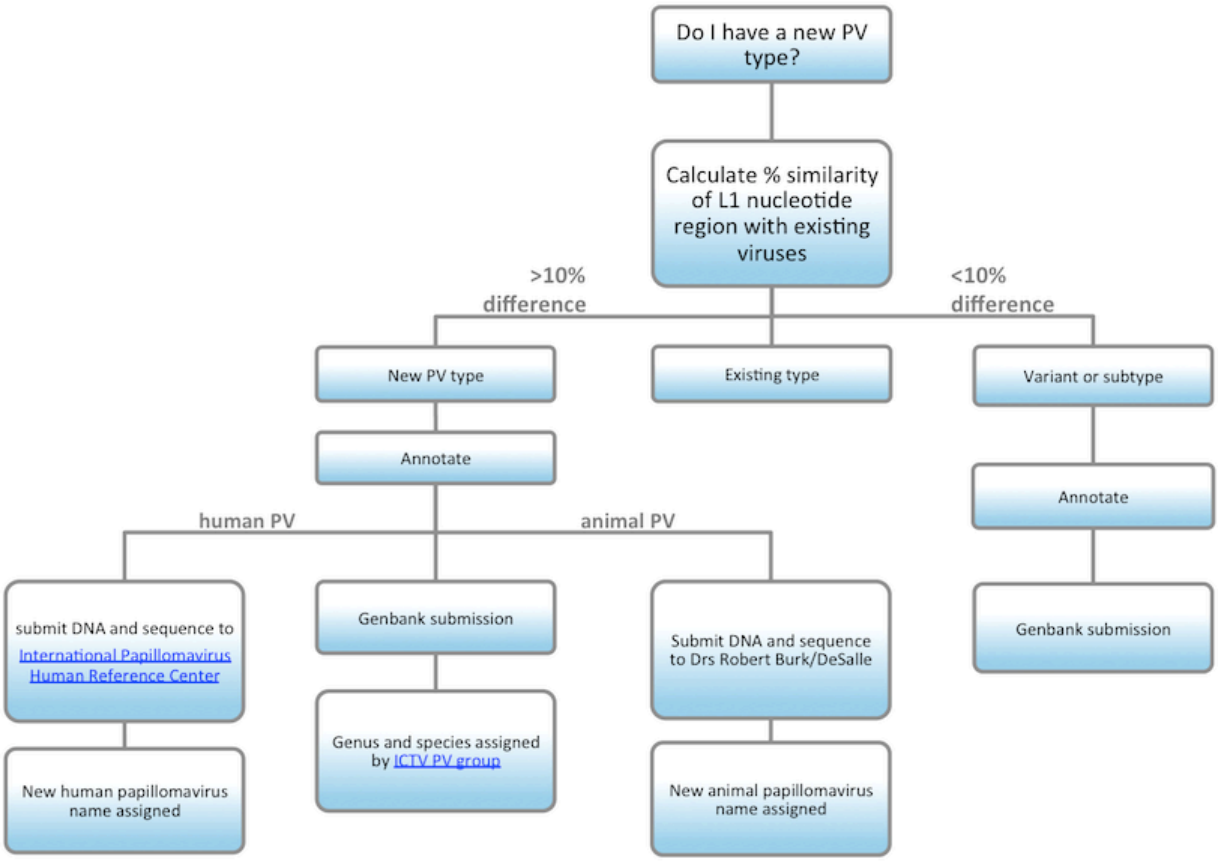

Figure 3.1 Novel HPV submission process (http://pave.niaid.nih.gov/#explore/taxonomy/submission_process).

Specific Aim 1b. Cloning and sequencing of FAP PCR positive samples:

Ten samples that were FAP PCR positive but not HPV16 positive, underwent regular sequencing to determine if they were novel HPVs. Five samples had ambiguous and poor quality results, thus cloning was performed using CloneJet PCR Cloning Kit by Thermo Scientific. Sequencing by Genewiz was executed after cloning followed by BLASTn (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins/) and PaVE (http://pave.niaid.nih.gov/) database search for similar HPVs. Chen et al. recommend using both databases in order to determine homology [106].

Clustal Omega was used to align multiple sequences to produce a phylogenetic tree for observation of evolutionary relationships (http://www.ebi.ac.uk/Tools/msa/clustalo/). Novel viruses were compared to all 175 HPVs in the PaVE database and sequences from the L1 region were obtained.

Specific Aim 2. Prevalence of novel HPVs in archived OSCC/OPSCC tissue samples:

With the identification of novel types of HPV, the presence of the new viral sequences in 106 normal and 115 malignant (OSCC/OPSCC) oral tissues was determined by HPV-type specific quantitative RT-PCR Taqman assays. Archived oral tissue blocks were accessed through the Department of Pathology's repository. A total of 80 µm of tissue was cut with a microtome and a new blade for each tissue block to eliminate contamination. Genomic DNA was isolated using RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE Tissues (Applied Biosystems) according to manufacturer's protocol. Primers and probes specific for the novel HPVs were created and used for detection.

OligoArchitect by Sigma (http://www.sigmaaldrich.com/technical-documents/articles/biology/oligoarchitect-online.html) was used to create custom primers and probes for three potential novel HPVs.

Stata version 13.1 was used to analyze data. Chi-square or Fisher's exact tests was used to determine the novel viral frequency differences between the case and control populations. CART (Classification and Regression Tree) analysis (Salford Systems) was used to identify potential viral-viral or viral-genetic interactions associated with OSCC/OPSCC. The potential role and association of these novel HPVs was determined by viewing the relationships of related HPVs.

Given a sample size of 100 cases and 100 controls without cancer, we had sufficient power to detect meaningful differences in HPV prevalence (Figure 1). For example, if novel HPV types are detected in 5% of control samples (p0), we have 80% power to detect differences in HPV prevalence if at least 19% of case samples have novel HPV detected (red line). Similarly, if 15% of controls have novel HPV detected, we have 80% power to detect differences if at least 33% of cases have novel HPV types (pink line).

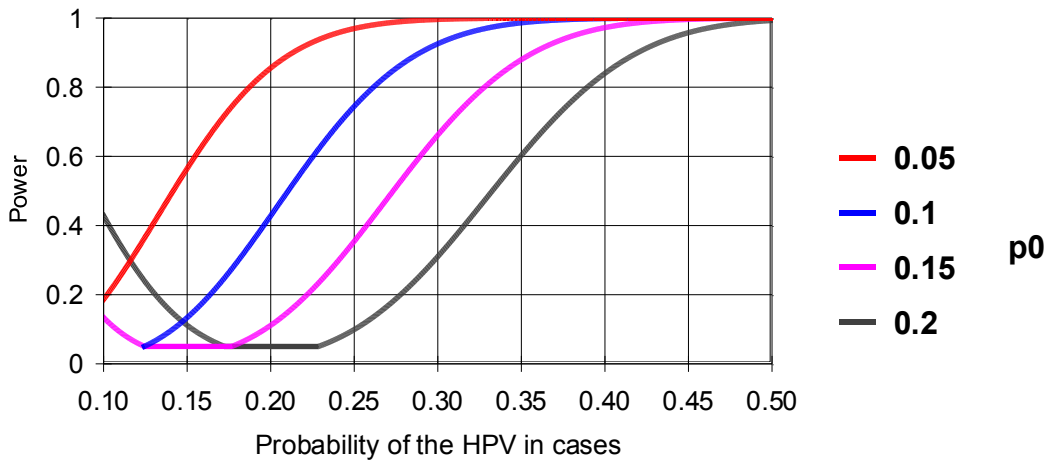

Figure 3.2 Power curves for detection of HPV in cases, with HPV detection in controls varying from 5% to 20%.

Specific Aim 3. Determine frequency of novel oncogenic HPVs in cancer and non-cancer oral rinse samples:

Using RT-PCR Taqman assays we determined the frequency of the novel oncogenic HPVs in cancer and non-cancer oral rinse samples. Specific primers and probes for the novel HPVs was used for detection.

OligoArchitect by Sigma (http://www.sigmaaldrich.com/technical-documents/articles/biology/oligoarchitect-online.html) was used to create custom primers and probes for three potential novel HPVs.

Data analysis will be done using Stata version 13.1.

Chapter IV - Results

Specific aim 1a results

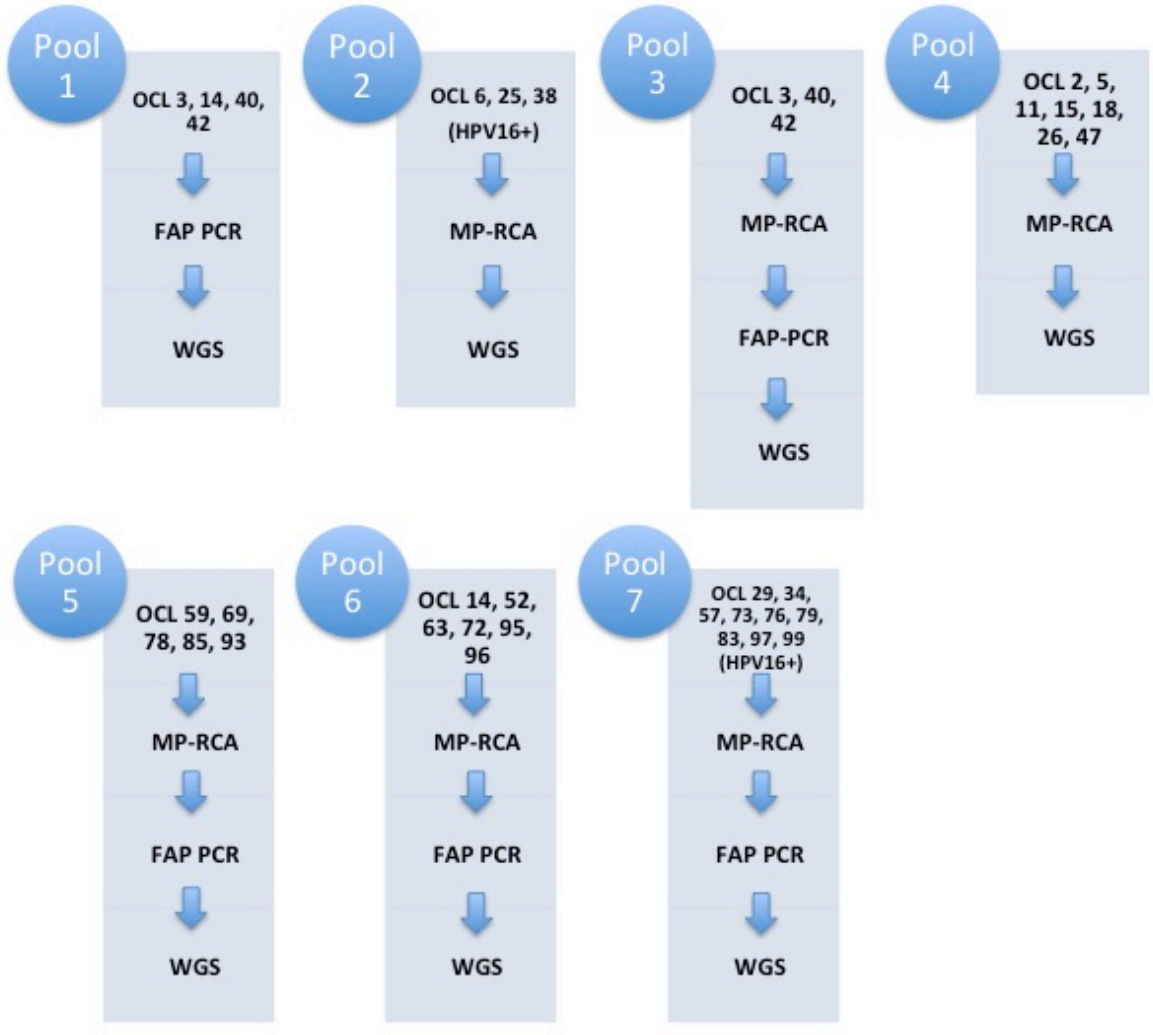

Figure 4.0. Details for process of 7 pooled samples during whole genome sequencing (WGS). OCL represents patient sample numbers.

Table 4.0 MP-RCA preferential amplification. HPV16 acted as a positive control, and Siha a negative control.

	HPV16 gene		Alu gene		
Sample	Before MP-RCA (Ct)	After MP-RCA (Ct)	Before MP-RCA (Ct)	After MP-RCA (Ct)	Enrichment*
HPV16	32	16	34	35	786,431
Siha	29	30	26	24	0.5

*Enrichment was calculated using the $2^{-\Delta\Delta Ct}$ method, where $\Delta\Delta Ct = [Ct_{HPV}(\text{after RCA}) - Ct_{HPV}(\text{before RCA})] - [Ct_{ALU}(\text{after RCA}) - Ct_{ALU}(\text{before RCA})]$

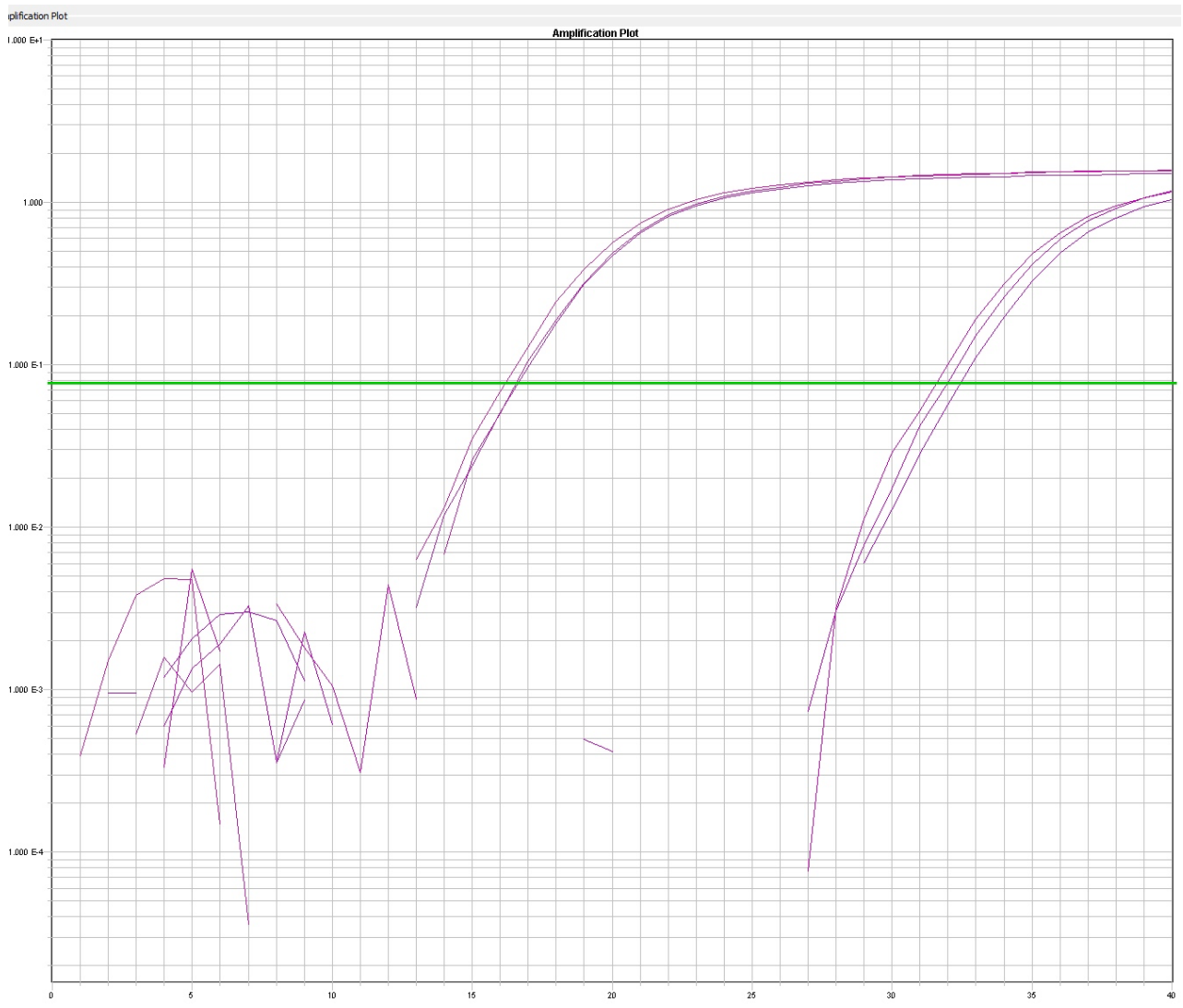

Figure 4.1 MP-RCA results for HPV16 plasmid acting as a positive control. The enrichment was almost 800,000-fold thus indicating a large amount of preferential amplification.

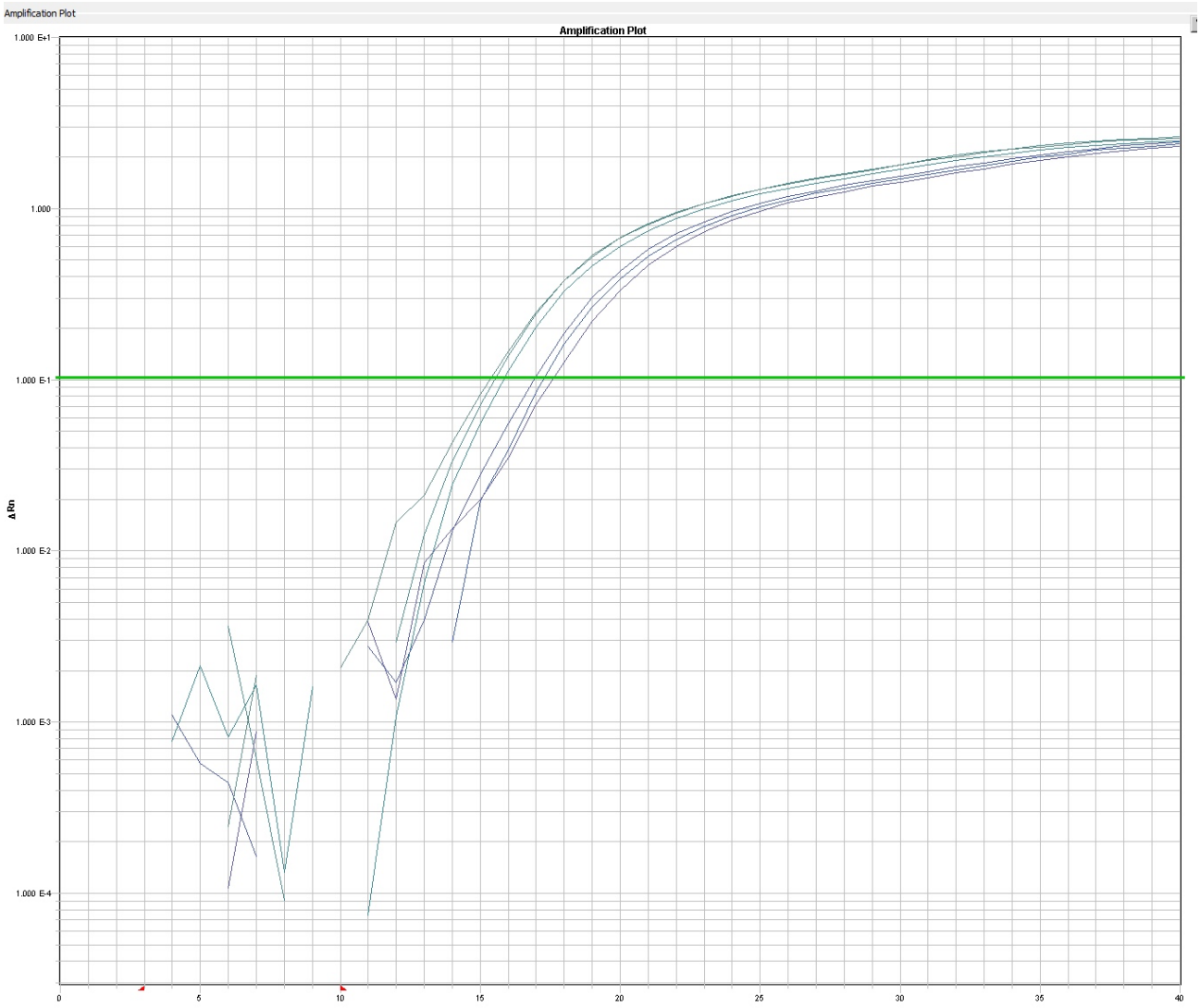

Figure 4.2 MP-RCA results using the Alu house-keeping gene to show that preferential amplification occurred mainly in circular HPV DNA and minimally in genomic linear DNA.

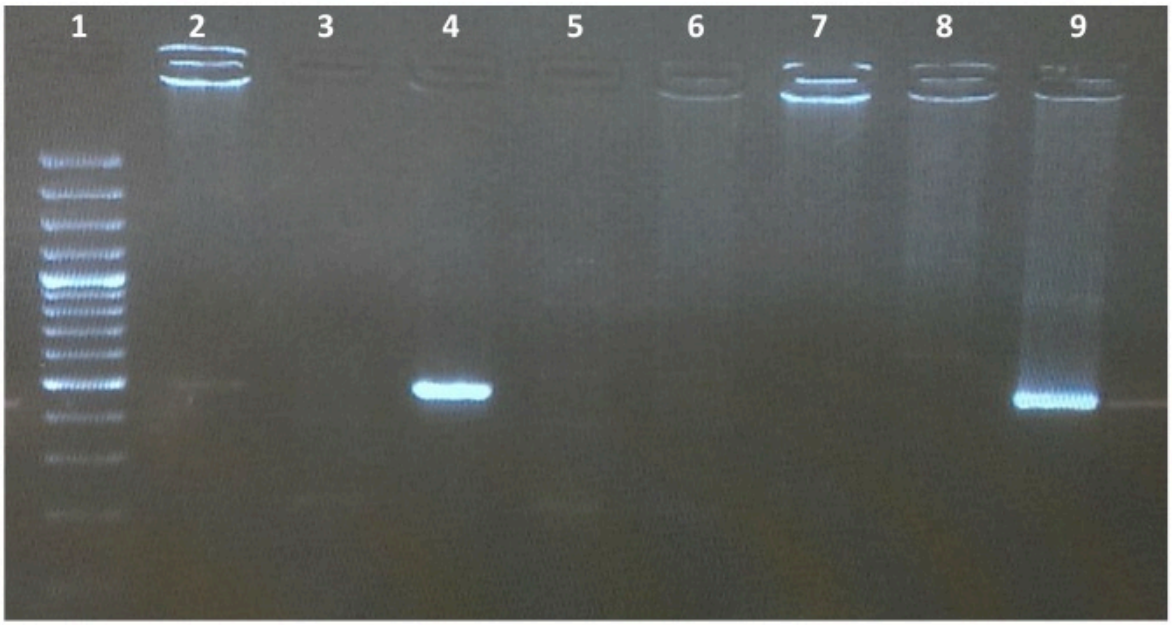

Figure 4.3 FAP PCR gel electrophoresis results. Lane 1 represents the DNA base pair ladder; lanes 2-7 represent HPV negative samples with the exception of lane 4, which was positive for the L1 region of HPV at ~480bp. Lane 8 was the negative control and lane 9 the positive control.

Table 4.1 NGS data analysis.

1st round of NGS	**Total Reads**	**Mapped Reads**	**% Mapped Reads to L1 gene**
I) FAP PCR	338374994	277598421	82.04
II) HPV16 RCA	394607770	9	0
III) FAP PCR RCA	380204724	67	0
IV) RCA	382376134	87	0
2nd round of NGS			
V) FAP PCR RCA	196273326	61518757	31.34
VI) FAP PCR RCA2	229994340	126833119	55.15
VII) FAP PCR RCA16	199750516	101856187	50.99

Table 4.2 FAP PCR (1st round of NGS) analysis. "% counts" refers to the number of reads that were mapped onto the L1 region of the specific HPV type indicated. "OCL sample" refers to the oral rinse sample that was positive for that specific HPV type after sequencing.

HPV type	Counts	% Counts	OCL sample match
172	109936978	39.6	3, 40, 42
171	64423530	23.21	
8	39473771	14.22	14
23	37762624	13.6	14
113	15793984	5.69	
49	4303896	1.55	
76	4134700	1.49	14
5	922729	0.33	
20	378380	0.14	
105	354884	0.13	
159	67672	0.02	
15	14461	0.01	
12	9517	0	
100	6985	0	
169	6941	0	
111	3009	0	
122	1202	0	
21	1085	0	
14	639	0	
143	322	0	
155	229	0	
24	199	0	
104	186	0	
75	153	0	
99	80	0	
93	62	0	
109	58	0	
38	32	0	
96	27	0	
36	21	0	
145	15	0	

98	12	0	
37	9	0	
151	9	0	
110	4	0	
120	2	0	
92	2	0	
80	2	0	
107	1	0	
126	1	0	
17	1	0	
170	1	0	

Table 4.3 FAP PCR RCA (2nd round of NGS) L1 sequence analysis.

HPV type	Counts	% Counts	OCL sample match
62	26223379	42.63	93
152	15993736	26	69
32	6799264	11.05	59
123	5170048	8.4	
80	2428865	3.95	
20	2419595	3.93	
76	871476	1.42	
22	340537	0.55	
24	328677	0.53	
124	240741	0.39	
145	221765	0.36	
23	141373	0.23	
8	93737	0.15	
96	82494	0.13	
36	44651	0.07	
161	43667	0.07	
92	29308	0.05	
17	22404	0.04	
37	6923	0.01	
15	4382	0.01	
107	4266	0.01	
150	3075	0	
12	937	0	
14	501	0	
111	474	0	
122	461	0	93
81	450	0	
9	425	0	
174	199	0	
109	196	0	
120	149	0	
21	125	0	
143	94	0	
149	80	0	

101	60	0	
110	54	0	
151	51	0	
104	33	0	
5	24	0	
93	18	0	
138	15	0	
100	14	0	
19	13	0	
75	4	0	
113	3	0	
105	2	0	
35	2	0	
38	1	0	
115	1	0	
119	1	0	
25	1	0	
102	1	0	

Table 4.4 FAP PCR RCA2 (2nd round of NGS) L1 sequence analysis.

HPV type	Counts	% Counts	OCL sample match
90	35679275	28.13	52
23	26865855	21.18	14
11	26393016	20.81	96
8	23967471	18.9	14
49	10632486	8.38	
33	1390467	1.1	72
105	1112699	0.88	63
76	490440	0.39	14
50	61770	0.05	
5	60111	0.05	
21	39472	0.03	
100	30253	0.02	
124	21997	0.02	
15	19466	0.02	
52	15955	0.01	
147	15910	0.01	95
14	12461	0.01	
32	9110	0.01	
20	6693	0.01	
98	2102	0	
104	1579	0	
113	1226	0	
111	1064	0	
36	591	0	
142	555	0	
122	515	0	
96	157	0	
24	118	0	
143	66	0	
75	49	0	
145	48	0	
37	20	0	
139	18	0	
80	15	0	

123	15	0	
134	14	0	
151	10	0	
48	8	0	
150	5	0	
135	5	0	
6	4	0	
152	3	0	
93	3	0	
12	3	0	
110	2	0	
58	2	0	
66	2	0	
92	2	0	
160	2	0	
148	1	0	
136	1	0	
19	1	0	
164	1	0	
99	1	0	
9	1	0	

Table 4.5 FAP PCR RCA16 (2nd round of NGS) L1 sequence analysis. Nine pooled samples (OCL 29, 34, 57, 73, 76, 79, 83, 97, 99) that were positive for HPV16. Infection with multiple HPVs is evident.

HPV type	Counts	% Counts
32	25896574	25.42
24	23472631	23.04
80	16758587	16.45
111	14702114	14.43
143	8484485	8.33
16	4781520	4.69
92	4685275	4.6
138	660990	0.65
135	582723	0.57
134	552801	0.54
76	137580	0.14
12	39851	0.04
172	39768	0.04
121	6491	0.01
175	4013	0
105	3947	0
120	3441	0
180	3077	0
133	2262	0
23	1666	0
151	1245	0
87	946	0
152	846	0
98	766	0
166	560	0
167	537	0
92	517	0
36	467	0
15	462	0
113	389	0
122	282	0
104	222	0

20	221	0
8	211	0
67	207	0
5	181	0
96	173	0
25	119	0
66	108	0
40	92	0
14	72	0
9	58	0
154	55	0
37	47	0
110	28	0
124	25	0
21	19	0
30	18	0
75	14	0
22	12	0
118	11	0
137	10	0
62	9	0
197	9	0
174	8	0
161	5	0
38	4	0
115	4	0
73	4	0
18	4	0
100	3	0
44	3	0
150	2	0
47	2	0
93	2	0
35	2	0
139	1	0
171	1	0
107	1	0

123	1	0
159	1	0
19	1	0
162	1	0
7	1	0

Summary of results for aim 1a

It is evident that we had substantial amplification present after MP-RCA as our HPV16 positive control had enrichment of almost 800,000-fold compared to only a 0.5-fold amplification with Siha (a cervical cancer cell line with integrated HPV16), which was our negative control (Table 4.0). Figure 4.1 and 4.2 demonstrates that preferential amplification occurred with circular HPV DNA as there was not much difference in Ct values for the negative control sample Siha before and after MP-RCA, thus indicating linear genomic DNA was amplified minimally.

In the first round of NGS, we only had reads from sample 1) FAP PCR where 82% of the reads mapped onto the L1 gene of all 189 HPV types in the PaVE database (Table 4.1). For sample 2) HPV16 RCA, we should have had reads mapped onto the HPV L1 gene as this was our positive control, but no results were seen. It is also odd that we did not see any results in sample 3) FAP PCR RCA, which contained OCL 3, 40, and 42, which were also in sample 1). In the second round of NGS, sample 5) FAP PCR RCA had 31% of reads mapped onto the L1 gene; sample 6) FAP PCR RCA2 had 55%; and sample 7) had 51% of reads mapped onto the L1 gene (Table 4.1).

We compared the conventional sequence results from aim 1b to NGS results in aim 1a and found results to be in concordance. OCL 3, 40, and 42 all matched to HPV172; OCL14 had multiple HPV infection with matches to HPVs 8, 23, and 76 (Table 4.2). OCL93 matched to HPV62 and 122; OCL 69 to HPV152; and OCL 59 to 32 (Table 4.3). OCL52 matched to HPV90; OCL96 to HPV11; OCL 72 to HPV33; OCL63 to HPV105; OCL95 to HPV147; and OCL14 was repeated in the second round of NGS with the same HPVs being mapped (types 8, 23, 76) thus acting as a positive control and confirmation of results from the first round of NGS.

In pooled sample VII) FAP PCR RCA 16 (Table 4.5), all samples were positive for HPV16. However, there is a high possibility from the mapped reads

that multiple infections are present as there were reads that matched to 15 other HPV types.

Specific Aim 1b results

Of the 10 samples that were FAP PCR positive and not HPV16 positive, we completed regular sequencing. Five samples had sequences with adequate quality for analysis. The other five had ambiguous/unreadable sequences and thus were cloned for additional DNA sequence analysis. Three potential novel viruses were seen in samples OCL 14.4, 69.2, and 95 as the difference in homology was >10%. OCL 14 and OCL 93 had multiple HPV infection present as at least two types of HPVs were seen.

Table 4.6 Samples that underwent cloning and sequencing. BLASTn and PaVE database search results and HPV types associated with the samples.

Sample	Clone	HPV type	% Homology BLASTn	% Homology PaVE
OCL 3.1	Yes	172	97.02	96.51
OCL 3.2	Yes	172	95.04	96.51
OCL 3.6	Yes	172	96.00	96.70
OCL 3.8	Yes	172	92.07	89.20
OCL 3.10	Yes	172	97.02	96.51
OCL 14	No	23	80.00	49.00
OCL 14.1	Yes	8	95.04	95.63
OCL 14.2	Yes	8	95.04	93.74
OCL 14.3	Yes	76	91.08	86.67
OCL 14.4	Yes	76	89.24	88.51
OCL 14.6	Yes	76	87.40	86.70
OCL 40	No	172	98.01	98.00
OCL 42	No	172	98.01	95.60
OCL 52	No	90	98.01	99.60
OCL 63.2	Yes	105	96.03	96.42
OCL 63.3	Yes	105	98.01	97.91
OCL 63.4	Yes	105	96.03	94.01
OCL 63.5	Yes	105	96.03	96.42
OCL 63.6	Yes	105	96.03	96.43
OCL 69.1	Yes	152	84.63	85.37
OCL 69.2	Yes	152	89.18	89.18
OCL 69.3	Yes	152	83.72	83.27
OCL 69.4	Yes	152	89.18	89.55
OCL 69.5	Yes	152	88.27	89.19
OCL 93.1	Yes	62	95.04	94.94
OCL 93.2	Yes	62	96.03	96.43
OCL 93.3	Yes	122	99.00	99.30
OCL 93.4	Yes	122	94.20	94.20
OCL 93.5	Yes	122	96.03	96.72
OCL 95	No	147	76.80	64.55
OCL 96	No	11	98.01	99.60

Table 4.7 Sequence results from Genewiz.

Sample	Sequence
OCL 3.1	ATTAACTGTGGGGCATCCATATTTTGATGTTATGGATGTTACCGATGAGTCTAA AGTAGCAATTCCAAAGGTTTCTGCCAATCAGTACAGGGTTATTAGACTACAATT TCCAGATCCAAACAAATTTGCTATCACAGATGCATGTGTTTATAATCCTGAAAAA GAGCGATTAGTATGGAGACTAGTAGGATTTCAAATGGATCGAGGTGGTCCATT AGGTATAGGAGCAACAGGCCATCCTTATTTTAATAAGTATGTGGATGCTGAGAA TCCTACAACATATCCCGAAAAGCAAGCAGAGGATGGGGATTATAGACAGGATA TGGCATTTGACCCTAAACAGGTTCAAATGTGTATTGTGGGCTGCACACCACCAA CAGGACAGTACTGGGATACCGCTGAATTTTGTCCAGGTCATAACAAAAATAATG GAGATTGTCCTCCAATAGAACTACACCACACTACAATTCAGGATGGCGA
OCL 3.2	TAACAGTGGGGCATCCATATTTTGATGTTATGGATGTTACCGATGAGTCTAAAG TAGCAATTCCAAAGGTTTCTGCCAATCAGTACAGGGTTATTAGACTACAATTTCC AGATCCAAACAAATTTGCTATCACAGATGCATGTGTTTATAATCCTGAAAAAGA GCGATTAGTATGGAGACTAGTAGGATTTCAAATGGATCGAGGTGGTCCATTAG GTATAGGAGCAACAGGCCATCCTTATTTTAATAAGTATGTGGATGCTGAGAATC CTACAACATATCCCGAAAAGCAAGCAGAGGATGGGGATTATAGACAGGATATG GCATTTGACCCTAAACAGGTTCAAATGTGTATTGTGGGCTGCACACCACCAACA GGACAGTACTGGGATACCGCTGAATTTTGTCCAGGTCATAACAAAAATAATGG AGATTGTCCTCCAATAGAACTACACCACACTACAATTCAGGATGGCGACA
OCL 3.6	TAACAGTGGGGCATCCTTATTTTGATGTTATGGATGTTACCGATGAGTCTAAAG TAGCAATTCCAAAGGTTTCTGCCAATCAGTACAGGGTTATTAGACTACAATTTCC AGATCCAAACAAATTTGCTATCACAGATGCATGTGTTTATAATCCTGAAAAAGA GCGATTAGTATGGAGACTAGTAGGATTTCAAATGGATCGAGGTGGTCCATTAG GTATAGGAGCAACAGGCCATCCTTATTTTAATAAGTATGTGGATGCTGAGAATC CTACAACATATCCCGAAAAGCAAGCAGAGGATGGGGATTATAGACAGGATATG GCATTTGACCCTAAACAGGTTCAAATGTGTATTGTGGGCTGCACACCACCAACA GGACAGTACTGGGATACCGCTGAATTTTGTCCAGGTCATAACAAAAATAATGG AGATTGTCCTCCAATAGAACTACACCACACTACAATTCAGGATGGCGACA
OCL 3.8	TCTAAAGTAGCAATTCCAAAGGTTTCTGCCAATCAGTACAGGGTTATTAGACTA CAATTTCCAGATCCAAACAAATTTGCTATCACAGATGCATGTGTTTATAATCCTG AAAAAGAGCGATTAGTATGGAGACTAGTAGGATTTCAAATGGATCGAGGTGGT CCATTAGGTATAGGAGCAACAGGCCATCCTTATTTTAATAAGTATGTGGATGCT GAGAATCCTACAACATATCCCGAAAAGCAAGCAGAGGATGGGGATTATAGACA GGATATGGCATTTGACCCTAAACAGGTTCAAATGTGTATTGTGGGCTGCACACC ACCAACAGGACAGTACTGGGATACCGCTGAATTTTGTCCAGGTCATAACAAAAA TAATGGAGATTGTCCTCCAATAGAACTACACCACACTACAATTCAGGATGGCGA CATGATAGATATAGGATCTTTCTAGAAGATCTCCTACAATATTCTCAGC

OCL 3.10	GTGGGGCATCCATATTTTGATGTTATGGATGTTACCGATGAGTCTAAAGTAGCA ATTCCAAAGGTTTCTGCCAATCAGTACAGGGTTATTAGACTACAATTTCCAGATC CAAACAAATTTGCTATCACAGATGCATGTGTTTATAATCCTGAAAAAGAGCGAT TAGTATGGAGACTAGTAGGATTTCAAATGGATCGAGGTGGTCCATTAGGTATA GGAGCAACAGGCCATCCTTATTTTAATAAGTATGTGGATGCTGAGAATCCTACA ACATATCCCGAAAAGCAAGCAGAGGATGGGGATTATAGACAGGATATGGCATT TGACCCTAAACAGGTTCAAATGTGTATTGTGGGCTGCACACCACCAACAGGACA GTACTGGGATACCGCTGAATTTTGTCCAGGTCATAACAAAAATAATGGAGATTG TCCTCCAATAGAACTACACCACACTACAATTCAGGATGGCGACATGGCA
OCL 14	AAGGACTTGAAATTGGACGCGGCCAACCTTTAGGGGTCGGAGCACCGGGACAC CCTCTTTTTAATAAGCTTCATAA
OCL 14.1	GATTAACAGTGGGGCATCCTTATTTCAATGTTTACAACAATAATGGTGACACATT ACAGGTTCCCAAAGTATCGGGAAATCAACACAGGGTCTTTCGCTTAAAGTTACC AGATCCAAATAGGTTTGCACTGGCAGATATGTCTGTGTACAATCCAGACAAGGA AAGGTTGGTATGGGCTTGCAGAGGCTTAGAAATCAGTAGGGGACAACCATTAG GTGTTGGGAGCACCGGCCATCCCTATTTTAATAAAGTGAAAGACACTGAAAACA GCAATTCATACACCACAACATCTACAGATGACAGACAAAATACTTCCTTTGATCC TAAGCAAATACAAATGTTCATTGTGGGTTGCACACCCTGCATTGGTGAGCATTG GGAAAAAGCCATTCCATGTGCAGAGGACCAACAGCAAGGTCTGTGCCCACCCA TTGAACTAAAAAATACAGTTATTGAAGATGGCGACATGTCAGATATAG
OCL 14.2	ATATCTATCATGTCGCCATCTTCAATAACTGTATTTTTTAGTTCAATGGGTGGGC ACAGACCTTGCTGTTGGTCCTCTGCACATGGAATGGCTTTTTCCCAATGCTCACC AATGCAGGGTGTGCAACCCACAATGAACATTTGTATTTGCTTAGGATCAAAGGA AGTATTTTGTCTGTCATCTGTAGATGTTGTGGTGTATGAATTGCTGTTTTCAGTG TCTTTCACTTTATTAAAATAGGGATGGCCGGTGCTCCCAACACCTAATGGTTGTC CCCTACTGATTTCTAAGCCTCTGCAAGCCCATACCAACCTTTCCTTGTCTGGATT GTACACAGACATATCTGCCAGTGCAAACCTATTTGGATCTGGTAACTTTAAGCG AAAGACCCTGTGTTGATTTCCCGATACTTTGGGAACCTGTAATGTGTCACCATTA TTGTTGTAAACATTGAAATATGGGTGCCCCACTGTTATCTTG
OCL 14.3	CCTATATCTCACATGTCGCCATCTTCAATTACAGTATTTACTAATTCTAAAGGTG GACATTTGCCTGCTCCTCTGTCAGCATCACAAGGTTTTGCTGCATCCCAGTGCTC TCCTTCACACGGTGTACAGCCAATAATAAACATTTGAACTTGTTTAGGATCAAAT GAGGTGTCCTGCCTATCATCCTTAGATGTTACTATGTAATTATTAGAATTTTCTG TATCCTTCACTTTATTGAATAGAGGGTGACCTGTAGATCCTACTCCCAGGGGTT GTCCGCGACCTATTTCCAAACCTCTACAGGCCCAAACCAGTCTTTCCTTTTCAGG ATTATAGACATTCATATCTACCAAGGCGAATCTATTGGGATCTGGTAGTAATAA TCTAAATGCTCTAAACTGATTACCTGATACTTTAGGAACTAGTATTTTTGTTTGAT CCACAGTGTCTCTAACATCAAAATATGGATGCCCCACAGTT

OCL 14.4	AACTGTGGGGCACCCATATTTTGATGTTAGAGACACTGTGGATCAAACAAAAAT ACTAGTTCCTAAAGTATCAGGTAATCAGTTTAGAGCATTTAGATTATTACTACCA GATCCCAATAGATTCGCCTTGGTAGATATGAATGTCTATAATCCTGAAAAGGAA AGGCTGGTTTGGGCCTGTAGAGGTTTGGAAATAGGTCGCGGACAACCCCTGGG AGTAGGATCTACAGGTCACCCTCTATTCAATAAAGTGAAGGATACAGAAAATTC TAATAATTACATAGTAACATCTAAGGATGATAGGCAGGACACCTCATTTGATCC TAAACAAGTTCAAATGTTTATTATTGGCTGTACACCGTGTGAAGGAGAGCACTG GGATGCAGCAAAACCTTGTGATGCTGACAGAGGAGCAGGCAAATGTCCACCTT TAGAATTAGTAAATACTGTAATTGAAGATGGCGACATGTCAGATATTGG
OCL 14.6	ATCCAATATCTGTCATGTCGCCATCTTCAATTACAGTATTTACTAATTCTAAAGGT GGACATTTGCCTGCTCCTCTGTCAGCATCACAAGGTTTTGCTGCATCCCAGTGCT CTCCTTCACACGGTGTACAGCCAATAATAAACATTTGAACTTGTTTAGGATCAAA TGAGGTGTCCTGCCTATCATCCTTAGATGTTACTATGTAATTATTAGAATTTTCT GTATCCTTCACTTTATTGAATAGAGGGTGACCTGTAGATCCTACTCCCAGGGGT TGTCCGCGACCTATTTCCAAACCTCTACAGGCCCAAACCAGTCTTTCCTTTTCAG GATTATAGACATTCATATCTACCAAGGCGAATCTATTGGGATCTGGTAGTAATA ATCTAAATGCTCTAAACTGATTACCTGATACTTTAGGAACTAGTATTTTTGTTTG ATCCACAGTGTCTCTAACATCAAAATATGGATGCCCCACTG
OCL 40	ATCAGTACAGGGTTATTAGACTACAATTTCCAGATCCAAACAAATTTGCTATCAC AGATGCATGTGTTTATAATCCTGAAAAAGAGCGATTAGTATGGAGACTAGTAG GATTTCAAATGGATCGAGGTGGTCCATTAGGTATAGGAGCAACAGGCCATCCTT ATTTTAATAAGTATGTGGATGCTGAGAATCCTACAACATATCCCGAAAAGCAAG CAGAGGATGGGGATTATAGACAGGATATGGCATTTGACCCTAAACAGGTTCAA ATGTGTATTGTGGGCTGCACACCACCAACAGGACAGTACTGGGATACCGCTGA ATTTTGTCCAGGTCATAACAAAAATAATGGAGATTGTCCTCCAATAGAACTACA CCACACTACAATTCAGGATGGCGACATGA
OCL 42	TTCCAAAGGTTTCTGCCAATCAGTACAGGGTTATTAGACTACAATTTCCAGATCC AAACAAATTTGCTATCACAGATGCATGTGTTTATAATCCTGAAAAAGAGCGATT AGTATGGAGACTAGTAGGATTTCAAATGGATCGAGGTGGTCCATTAGGTATAG GAGCAACAGGCCATCCTTATTTTAATAAGTATGTGGATGCTGAGAATCCTACAA CATATCCCGAAAAGCAAGCAGAGGATGGGGATTATAGACAGGATATGGCATTT GACCCTAAACAGGTTCAAATGTGTATTGTGGGCTGCACACCACCAACAGGACA GTACTGGGATACCGCTGAATTTTGTCCAGGTCATAACAAAAATAATGGAGATTG TCCTCCAATAGAACTACACCACACTACAATTCAGGATGGCGACATGACTGATAT AGGT
OCL 52	TAGTGGTTCCCAAGGTGTCTGGATATCAATATAGGGTGTTTAGGGTACGTTTGC CTGATCCCAATAAGTTTGGCCTTCCTGATGCATCGCTATACAATCCTGACTCGCA GCGCCTTGTATGGGCCTGTACAGGTGTTGAGGTTGGCAGGGGACAGCCTTTAG GCGTTGGGGTAAGTGGCCACCCGTTGTACAACCGCCTGTATGACACTGAAAAC ACCAATTTATATGATGTTGTGCCTGGCGATGACACCCGGGACAATCTTACTATG

	GACTATAAGCAAACCCAGCTATTTATTATAGGGTGCAAACCTCCTTTAGGCGAG CACTGGGCAAAGGGTACCCCATGCAATATGTCTAATGTACAGGCTGGGGATTG CCCTCCTATAGAACTTAAATCTTCCACAATTCAGGATGGCGACATGATG
OCL 63.2	CCTATATCTGTCATGTCGCCATCCTGGATAACAGTGTTCTTTAATTCTATAGGAG GACAGGCTCCATTGTCCTGCGCCTGTCCTGCACATGGCAGAGCCTTTTCCCAAT GTTCCCCAATACAAGGTGTACATCCAACTATGAACATCTGTATTTGTTTTGGGTC AAAAGAGGTATTTTGTCTGTCATCTTTGGAAGTGGTGGAATATTGATTACTGTTT TCAGTATCCTTCAGTTTATTAAAGTAAGGGTGACCTGTGCTGCCGACACCTAAG GGTTGCCCTCTACTTATTTCCAAACCTCTACAGGCCCACACCAAACGCTCCTTAT CTGGATTATAAACTGACATGTCAGCTAATGCAAATCTGTTTGGATCAGGCAACT TCAGACGAAACACCCTGTGTTGATTGCCTGATACCTTAGGAACCTGTAATGTCT CACCAGTGTTGTTATATACATTAAAATATGGATGCCCCACTGTT
OCL 63.3	TAACTGTGGGGCATCCATATTTTAATGTATATAACAACACTGGTGAGACATTAC AGGTTCCTAAGGTATCAGGCAATCAACACAGGGTGTTTCGTCTGAAGTTGCCTG ATCCAAACAGATTTGCATTAGCTGACATGTCAGTTTATAATCCAGATAAGGAGC GTTTGGTGTGGGCCTGTAGAGGTTTGGAAATAAGTAGAGGGCAACCCTTAGGT GTCGGCAGCACAGGTCACCCTTACTTTAATAAACTGAAGGATACTGAAAACAGT AATCAATATTCCACCACTTCCAAAGATGACAGACAAAATACCTCTTTTGACCCAA AACAAATACAGATGTTCATAGTTGGATGTACACCTTGTATTGGGGAACATTGGG AAAAGGCTCTGCCATGTGCAGGACAGGCGCAGGACAATGGAGCCTGTCCTCCT ATAGAATTAAAGAACACTGTTATCCAGGATGGCGACATGTCAGATATAG
OCL 63.4	GGGGCATCCATATTTTAATGTATATAACAACACTGGTGAGACATTACAGGTTCC TAAGGTATCAGGCAATCAACACAGGGTGTTTCGTCTGAAGTTGCCTGATCCAAA CAGATTTGCATTAGCTGACATGTCAGTTTATAATCCAGATAAGGAGCGTTTGGT GTGGGCCTGTAGAGGTTTGGAAATAAGTAGAGGGCAACCCTTAGGTGTCGGCA GCACAGGTCACCCTTACTTTAATAAACTGAAGGATACTGAAAACAGTAATCAAT ATTCCACCACTTCCAAAGATGACAGACAAAATACCTCTTTTGACCCAAAACAAAT ACAGATGTTCATAGTTGGATGTACACCTTGTATTGGGGAACATTGGGAAAAGG CTCTGCCATGTGCAGGACAGGCGCAGGACAATGGAGCCTGTCCTCCTATAGAA TTAAAGAACACTGTTATCCAGGATGGCGACATGTTAGATATAGGATCTTT
OCL 63.5	CCTATATCTGTCATGTCGCCATCCTGGATAACAGTGTTCTTTAATTCTATAGGAG GACAGGCTCCATTGTCCTGCGCCTGTCCTGCACATGGCAGAGCCTTTTCCCAAT GTTCCCCAATACAAGGTGTACATCCAACTATGAACATCTGTATTTGTTTTGGGTC AAAAGAGGTATTTTGTCTGTCATCTTTGGAAGTGGTGGAATATTGATTACTGTTT TCAGTATCCTTCAGTTTATTAAAGTAAGGGTGACCTGTGCTGCCGACACCTAAG GGTTGCCCTCTACTTATTTCCAAACCTCTACAGGCCCACACCAAACGCTCCTTAT CTGGATTATAAACTGACATGTCAGCTAATGCAAATCTGTTTGGATCAGGCAACT TCAGACGAAACACCCTGTGTTGATTGCCTGATACCTTAGGAACCTGTAATGTCT CACCAGTGTTGTTATATACATTAAAATATGGATGCCCCACTGTT

OCL 63.6	TCTAGAAGATCCAATATCTGTCATGTCGCCATCCTGGATAACAGTGTTCTTTAAT TCTATAGGAGGACGGGCTCCATTGTCCTGCGCCTGTCCTGCACATGGCAGAGCC TTTTCCCAATGTTCCCCAATACAAGGTGTACATCCAACTATGAACATCTGTATTT GTTTTGGGTCAAAAGAGGTATTTTGTCTGTCATCTTTGGAAGTGGTGGAATATT GATTACTGTTTTCAGTATCCTTCAGTTTATTAAAGTAAGGGTGACCTGTGCTGCC GACACCTAAGGGTTGCCCTCTACTTATTTCCAAACCTCTACAGGCCCACACCAAA CGCTCCTTATCTGGATTATAAACTGACATGTCAGCTAATGCAAATCTGTTTGGAT CAGGCATCTTCAGACGAAACACCCTGTGTTGATTGCCTGATACCTTAGGAACCT GTAATGTCTCACCAGTGTTGTTATATACATTAAAATAAGGATG
OCL 69.1	TCCTGAATATAGGAGTTTACTAGTTCTATTGGAGGACATCTACCAGGTTGATCA TCAGCAACGCAGGCAGGGGCTTTGTCCCAATGTTCGCCAATGCAAGGGGTACA GCCAATAATAAACATTTGCAGTTGTTTAGGATCAAATGAGATATTTTGTCTATCA TCTTTAGAGGTATTTCTATATGTATTTCCATTTTCTGTGTCATTCACCTTATTAAAT AAAGGATGTCCACTACTGCCTATACCTAATGGTTGTCCTCTGCCTATTTCAATAC CCTTTAGGCCCCATACTAATCTTTCCTTGTCTGGATTATAGACAGACATGTCAGC TAATGCAAATCTATTGGGGTCAGGTAATTTTAATCTAAAGACCCTGTGTTGATTT CCTGATACCTTAGGTACCTCCAAACGTGTACCTGCATTGTTGTAAATGTTAAAAT ATGGATGCCCCACAGTTAATCTTGCTGAAAAACTCGAGCCA
OCL 69.2	TAACTGTGGGGCATCCATATTTTAACATTTACAACAATGCAGGTACACGTTTGG AGGTACCTAAGGTATCAGGAAATCAACACAGGGTCTTTAGATTAAAATTACCTG ACCCCAATAGATTTGCATTAGCTGACATGTCTGTCTATAATCCAGACAAGGAAA GATTAGTATGGGGCCTAAAGGGTATTGAAATAGGCAGAGGACAACCATTAGGT ATAGGCAGTAGTGGACATCCTTTATTTAATAAGGTGAATGACACAGAAAATGG AAATACATATAGAAATACCTCTAAAGATGATAGACAAAATATCTCATTTGATCCT AAACAACTGCAAATGTTTATTATTGGCTGTACCCCTTGCATTGGCGAACATTGG GACAAAGCCCCTGCCTGCGTTGCTGATGATCAACCTGGTAGATGTCCTCCAATA GAACTAGTAAACTCCTATATTCAGGATGGCGACATGGCAGATATTGGAT
OCL 69.3	AGGAGTTTACTAGTTCTATTGGAGGACATCTACCAGGTTGATCATCAGCAACGC AGGCAGGGGCTTTGTCCCAATGTTCGCCAATGCAAGGGGTACAGCCAATAATA AACATTTGCAGTTGTTTAGGATCAAATGAGATATTTTGTCTATCATCTTTAGAGG TATTTCTATATGTATTTCCATTTTCTGTGTCATTCACCTTATTAAATAAAGGATGT CCACTACTGCCTATACCTAATGGTTGTCCTCTGCCTATTTCAATACCCTTTAGGCC CCATACTAATCTTTCCTTGTCTGGATTATAGACAGACATGTCAGCTAATGCAAAT CTATTGGGGTCAGGTAATTTTAATCTAAAGACCCTGTGTTGATTTCCTGATACCT TAGGTACCTCCAAACGTGTACCTGCATTGTTGTAAATGTTAAAATATGGGTGCC CCACAGTTATCTTGCTGAAAAACTCGAGCCATCCGGAAGATC
OCL 69.4	TAACTGTGGGGCATCCATATTTTAACATTTACAACAATGCAGGTACACGTTTGG AGGTACCTAAGGTATCAGGAAATCAACACAGGGTCTTTAGATTAAAATTACCTG ACCCCAATAGATTTGCATTAGCTGACATGTCTGTCTATAATCCAGACAAGGAAA GATTAGTATGGGGCCTAAAGGGTATTGAAATAGGCAGAGGACAACCATTAGGT

	ATAGGCAGTAGTGGACATCCTTTATTTAATAAGGTGAATGACACAGAAAATGG AAATACATATAGAAATACCTCTAAAGATGATAGACAAAATATCTCATTTGATCCT AAACAACTGCAAATGTTTATTATTGGCTGTACCCCTTGCATTGGCGAACATTGG GACAAAGCCCCTGCCTGCGTTGCTGATGATCAACCTGGTAGATGTCCTCCAATA GAACTAGTAAACTCCTATATTCAGGATGGCGACATGGCAGATATTGGAT
OCL 69.5	CTTCTAGAAGATCCAATATCTGCCATGTCGCCATCCTGAATATAGGAGTTTACTA GTTCTATTGGAGGACATCTACCAGGTTGATCATCAGCAACGCAGGCAGGGGCT TTGTCCCAATGTTCGCCAATGCAAGGGGTACAGCCAATAATAAACATTTGCAGT TGTTTAGGATCAAATGAGATATTTTGTCTATCATCTTTAGAGGTATTTCTATATG TATTTCCATTTTCTGTGTCATTCACCTTATTAAATAAAGGATGTCCACTACTGCCT ATACCTAATGGTTGTCCTCTGCCTATTTCAATACCCTTTAGGCCCCATACTAATCT TTCCTTGTCTGGATTATAGACAGACATGTCAGCTAATGCAAATCTATTGGGGTC AGGTAATTTTAATCTAAAGACCCTGTGTTGATTTCCTGATACCTTAGGTACCTCC AAACGTGTACCTGCATTGTTGTAAATGTTAAAATAAGGGTGC
OCL 93.1	CCATCCTGAATAGTTGTATTTTTAAATTCCAACGGAGGGCATTCCGTGGGGGCC GGGGCAGCATTGGGGCATAAGGTACCTTTGGTCCAGTGCTCACCTATAGGGGG CTTACACCCCACAATTAACAACTGGGTCTGCTTATAATCCACAGAGATATTGTCC CGACTGTCATCATTAGCAGCAGCCAACAAAGAGGTATTTTCTGTATCATCCAAC CTGTTATATAACGGGTGGCCACTGGTGCCAACACCCAGTGGCTGCCCACGGCC GACCTCAATGCCCCTGCAGGCCCATACCATGCGTTCCGTGTCTGGATTATATAA GGTTCCATCAGGTAAAGCAAATTTATTAGGGTCTGGTAGTTTCACACGAAACAC CCTGTACTGATACCCAGACACCTTAGGAATGGTGGCCCGTTTACCCTGGCCAAC CTGTAAAGTACAATATGGATGCCCCACAGTTAATCTTGCTGAAAAACTC
OCL 93.2	TAACTGTGGGGCATCCATATTGTACTTTACAGGTTGGCCAGGGTAAACGGGCC ACCATTCCTAAGGTGTCTGGGTATCAGTACAGGGTGTTTCGTGTGAAATTACCA GACCCTAATAAATTTGCTTTACCTGATGGAACCTTATATAATCCAGACACGGAAC GCATGGTATGGGCCTGCAGGGGCATTGAGGTCGGCCGTGGGCAGCCACTGGG TGTTGGCACCAGTGGCCACCCGTTATATAACAGGTTGGATGATACAGAAAATAC CTCTTTGTTGGCTGCTGCTAATGATGACAGTCGGGACAATATCTCTGTGGATTAT AAGCAGACCCAGTTGTTAATTGTGGGGTGTAAGCCCCCTATAGGTGAGCACTG GACCAAAGGTACCTTATGCCCCAATGCTGCCCCGGCCCCCACGGAATGCCCTCC GTTGGAATTTAAAAATACAACTATTCAGGATGGCGACATGACAGATATTG
OCL 93.3	TAACAGTGGGGCATCCATATTTCGATGTCCGATCTCAAGATGGGCAACGTATAG AGGTCCCTAAGGTGTCTGGCAATTAGTATAGATCATTTAGAATAACATTTCCGG ATCCTAATAGATTTGCTTTAGCAGATATGTCTGTGTACAATCCTGAAAAGGAAA GATTAGTGTGGGCCTGTAGAGGCCTGGAGATAGGCAGGGGTCAGCCTTTGGG TGTAGGAACATCAGGTCATCCTTTATTTAACAAAGTCAGGGATACTGAAAACTC AGGTAACTATCAAGCAGTTTCTCAGGATGACAGACAAAATACATCTTTTGATCC TAAACAAGTGCAAATGTTTGTCATTGGCTGTGTGCCGTGTATGGGTGAACATTG GGACAAAGCTAAGGTTTGTGAATCAGAAGCAAATAATCAACAAGGCTTATGTC

	CACCCATAGAG
OCL 93.4	CCAATATCTGTCATGTCGCCATCTTCAATTACTGAATTTTTTAACTCTATGGGTG GACATAAGCCTTGTTGATTATTTGCTTCTGATTCACAAACCTTAGCTTTGTCCCA ATGTTCACCCATACACGGCACACAGCCAATGACAAACATTTGCACTTGTTTAGG ATCAAAAGATGTATTTTGTCTGTCATCCTGAGAAACTGCTTGATAGTTACCTGAG TTTTCAGTATCCCTGACTTTGTTAAATAAAGGATGACCTGATGTTCCTACACCCA AAGGCTGACCCCTGCCTATCTCCAGGCCTCTACAGGCCCACACTAATCTTTCCTT TTCAGGATTGTACACAGACATATCTGCTAAAGCAAATCTATTAGGATCCGGAAA TGTTATTCTAAATGATCTATACTGATTGCCAGACACCTTAGGGACCTCTATACGT TGCCCATCTTGAGATCGGACATCGAAATATGGGTGCCCCACT
OCL 93.5	GTCGCCATCTTCAATTACTGAATTTTTTAACTCTATGGGTGGACATAAGCCTTGT TGATTATTTGCTTCTGATTCACAAACCTTAGCTTTGTCCCAATGTTCACCCATACA CGGCACACAGCCAATGACAAACATTTGCACTTGTTTAGGATCAAAAGATGTATT TTGTCTGTCATCCTGAGAAACTGCTTGATAGTTACCTGAGTTTTCAGTATCCCTG ACTTTGTTAAATAAAGGATGACCTGATGTTCCTACACCCAAAGGCTGACCCCTG CCTATCTCCAGGCCTCTACAGGCCCACACTAATCTTTCCTTTTCAGGATTGTACA CAGACATATCTGCTAAAGCAAATCTATTAGGATCCGGAAATGTTATTCTAAATG ATCTATACTGATTGCCAGACACCTTAGGGACCTCTATACGTTGCCCATCTTGAGA TCGGACATCGAAATATGGGTGCCCCACTGTTAATCTTGCTGA
OCL 95	TTGCAGTTCCTAAAGTATCAGGTTCTCAATACCGAGTATTTAGATGTAAATTACC AGATCCTAATAAATTTGCTCTTATAGAGAGAACGGTGTATAATTCAGACAGTGA GCGGTTAGTGTGGAAACTTCGAGGATTACAATTAGGAAGAGGGGGTCCATTAG GATTAGGAACTAGTGGGCATCCTTTATTTAATAAAGTGTTAGATACAGAAAATC CTAACTCCTACCCACCAAAACAGACCGATGAACAACGACTGGATGTTAGCATGG ATCCTAAACAGGTTCAAATGCTAATTGTTGGTTGTGAACCTGCTATCGGTGAAC ATTGGGACATAGCTAAACCTTGCTCGGATGAACAGCCTGAAAATGGCGATTGTC CTCCTATTCAATTGCTAAACACAGTGATTGAGGATGGCGACATGACAGATATAG GT
OCL 96	GTACCAAAGGTGTCTGGATATCAATATAGAGTGTTTAAGGTAGTGTTGCCAGAT CCTAACAAGTTTGCATTACCTGATTCATCCCTGTTTGACCCCACTACACAGCGTT TAGTATGGGCGTGCACAGGGTTGGAGGTAGGCAGGGGTCAACCTTTAGGCGTT GGTGTTAGTGGGCATCCATTGCTAAACAAATATGATGATGTAGAAAATAGTGG TGGGTATGGTGGTAATCCTGGTCAGGATAATAGGGTTAATGTAGGTATGGATT ATAAACAAACCCAGCTATGTATGGTGGGCTGTGCTCCACCGTTAGGTGAACATT GGGGTAAGGGTACACAATGTTCAAATACCTCTGTACAAAATGGTGACTGCCCCC CGTTGGAACTTATTACCAGTGTTATACAGGATGGCGACATGC

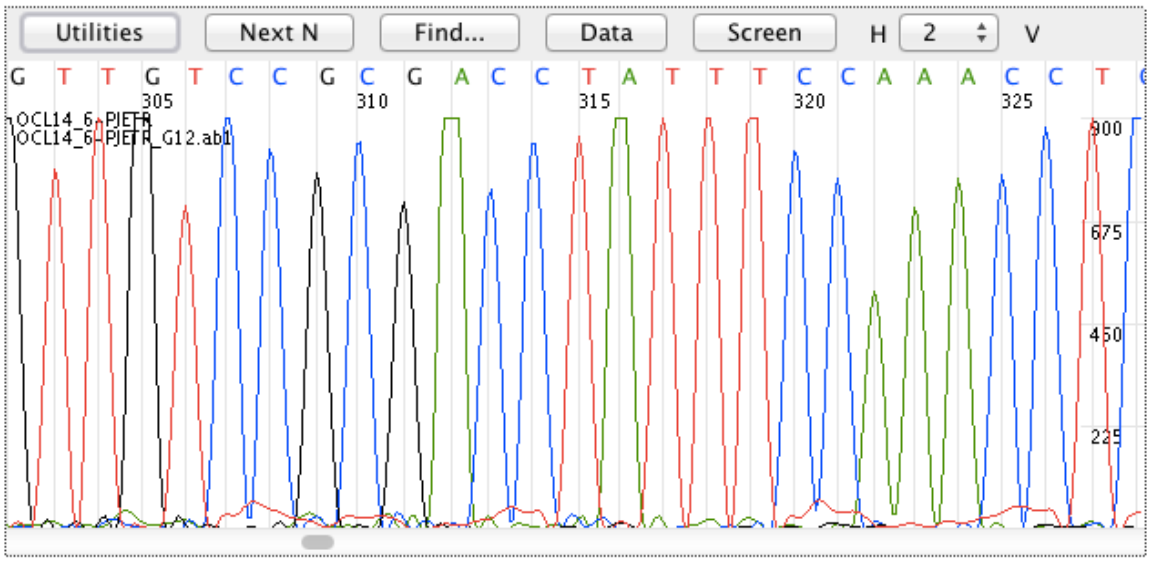

Figure 4.4 Trace file of high quality Sanger sequence from OCL14.6 clone.

Table 4.8 Medical history for patients where potential novel HPVs were isolated.

Sample	Cancer location	Gender	Age	Smoking history	Alcohol history
OCL 14	Larynx	Male	60	Heavy	Rarely
OCL 69	Base of tongue	Male	70	Light	Light
OCL 95	Lateral tongue	Male	63	Heavy	Heavy

Phylogenetic Tree Analysis

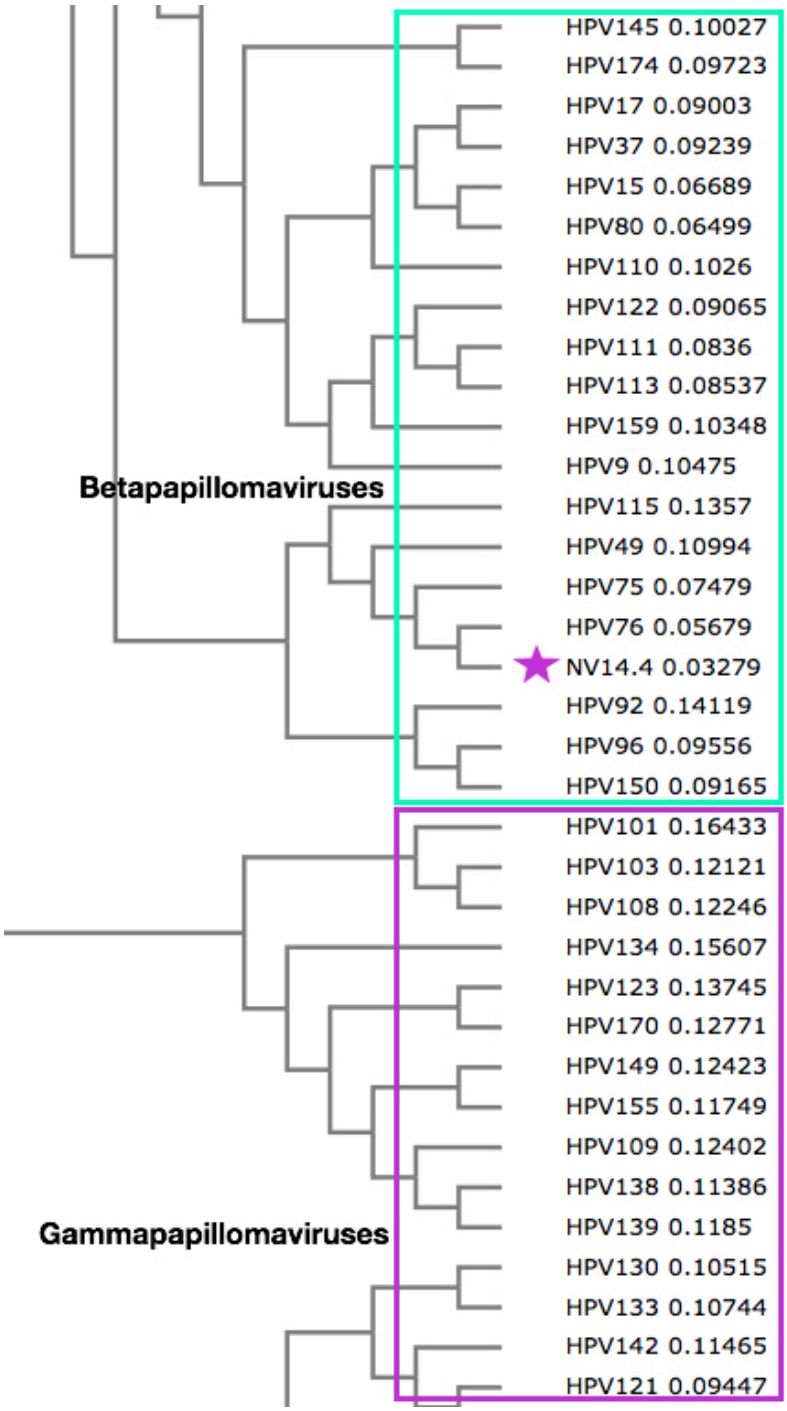

Figure 4.5 NV 14.4 (purple star) is related to HPV76, which belongs to the betapapillomavirus genus. The % homology between the L1 region of NV14.4 and HPV76 is 89%.

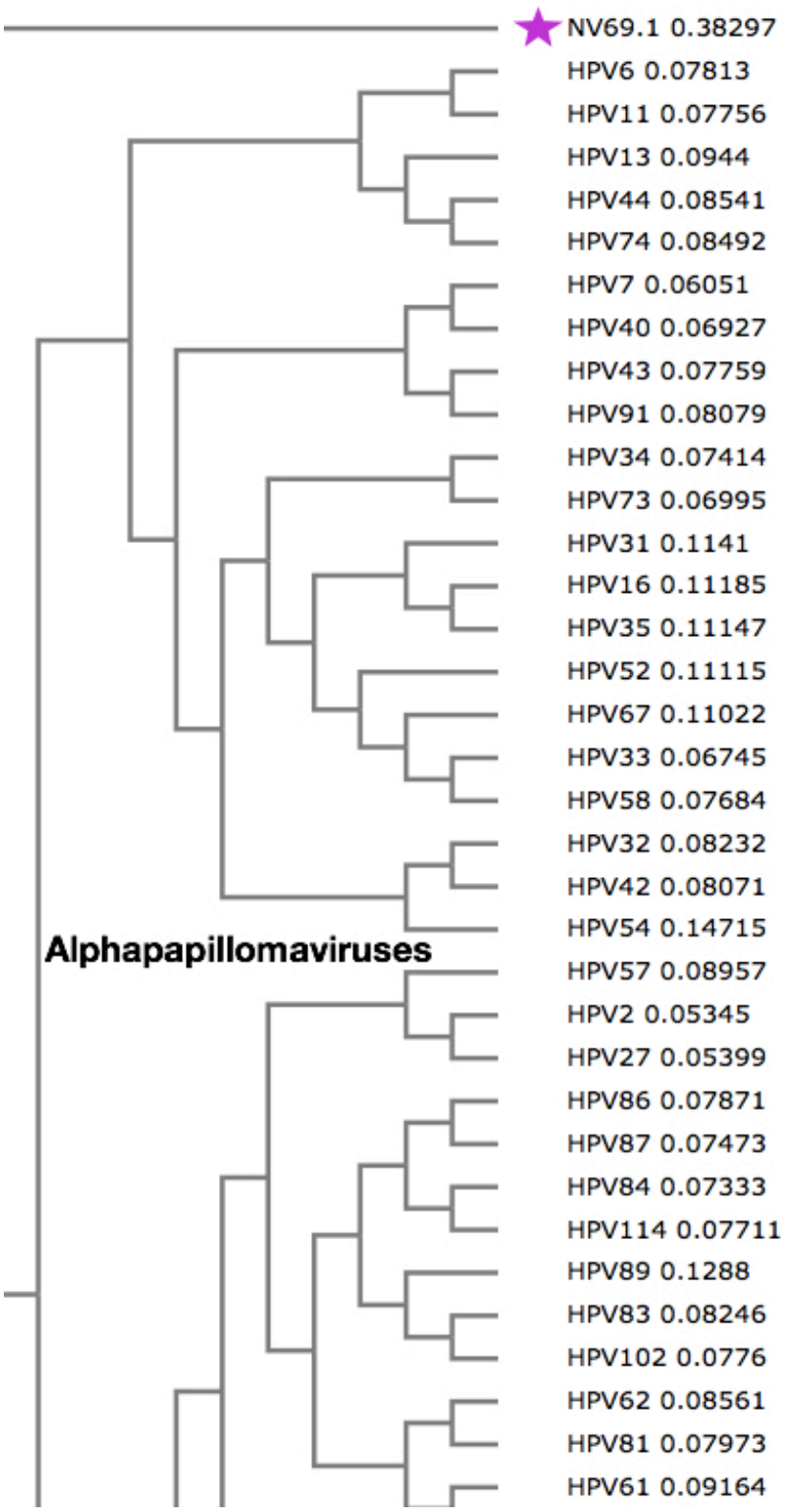

Figure 4.6 NV69.1 (purple star) is distantly related to the alphapapillomavirus genus.

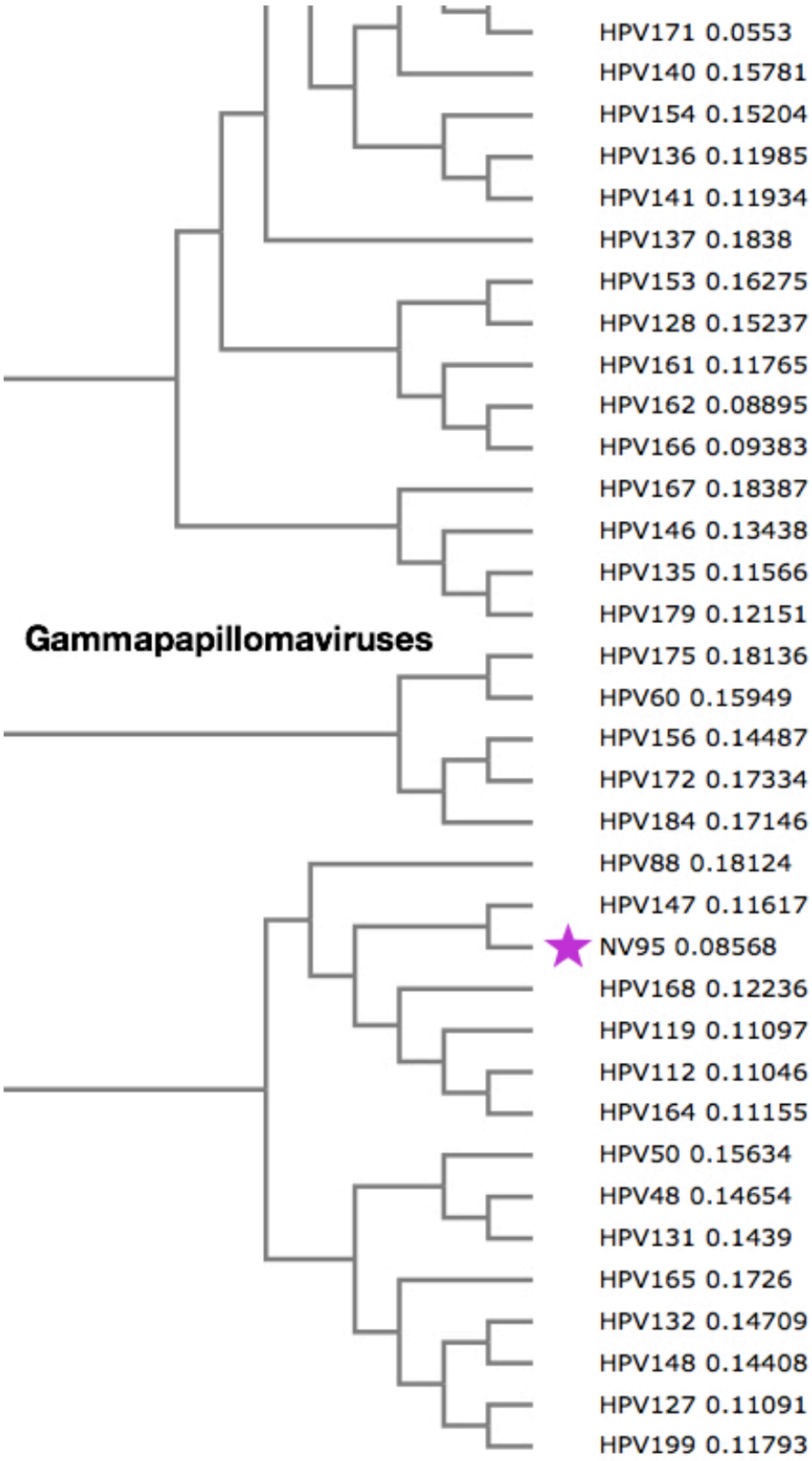
HPV171 0.0553
HPV140 0.15781
HPV154 0.15204
HPV136 0.11985
HPV141 0.11934
HPV137 0.1838
HPV153 0.16275
HPV128 0.15237
HPV161 0.11765
HPV162 0.08895
HPV166 0.09383
HPV167 0.18387
HPV146 0.13438
HPV135 0.11566
HPV179 0.12151
Gammapapillomaviruses
HPV175 0.18136
HPV60 0.15949
HPV156 0.14487
HPV172 0.17334
HPV184 0.17146
HPV88 0.18124
HPV147 0.11617
NV95 0.08568
HPV168 0.12236
HPV119 0.11097
HPV112 0.11046
HPV164 0.11155
HPV50 0.15634
HPV48 0.14654
HPV131 0.1439
HPV165 0.1726
HPV132 0.14709
HPV148 0.14408
HPV127 0.11091
HPV199 0.11793

Figure 4.7 NV 95 (purple star) is related to HPV147, which is part of the gammapapillomavirus genus. The % homology between NV95 and HPV147 ranged between 65-77%.

Summary of results for aim 1b

With sequencing and cloning we were able to obtain ~480bp of the L1 region of unknown HPVs from our samples (Table 4.7). Our sequences demonstrated to be of high quality (Figure 4.4). Using the BLASTn and PaVE databases we were able to tentatively identify which HPV types were present in our samples (Table 4.6); notably multiple HPV types were present in OCL14 and OCL93. Table 4.8 demonstrates the medical history of the patients where the potential novel HPVs were isolated. These patients had laryngeal cancer, OPSCC, and OSCC; they were 60 years or older; two were heavy smokers, one a light smoker; and alcohol use ranged from rare to heavy.

We used the phylogeny tool to view the relationship of our NVs to 175 fully characterized HPV types (Figure 4.5-4.7). NV14.4 is related to HPV76, which is of the betapapillomavirus family. NV69.1 is related to HPV152, which is also part of the betapapillomavirus family. However, the results from the phylogenic tree indicate that NV69.1 is quite distant from all HPVs, and may be distantly related to the alphapapillomavirus family. NV95 is related to HPV147, which belongs to the gammapapillomavirus family.

Specific aim 2 results

From the sequences in aim 1b, we created new primers and probes for the novel HPVs using Sigma's OligoArchitect online tool (http://www.sigmaaldrich.com/technical-documents/articles/biology/oligoarchitect-online.html).

Table 4.9 Primers and probe sequences for novel HPVs.

NV 14.4 forward primer	GACACTGTGGATCAAACA
NV 14.4 reverse primer	GCCTTTCCTTTTCAGGATTA
NV 14.4 probe (6FAM, BHQ-1)	AGACATTCATATCTACCAAGGCGAA
NV 69.1 forward primer	GGAGTTTACTAGTTCTATTGG
NV 69.1 reverse primer	ACTGCAAATGTTTATTATTGG
NV 69.1 probe (6FAM, BHQ-1)	ATCATCAGCAACGCAGGCAG
NV 95 forward primer	CAGACAGTGAGCGGTTAGTGTG
NV 95 reverse primer	CAATGTTCACCGATAGCAGGTT
NV 95 probe (6FAM, BHQ-1)	ATTTGAACCTGTTTAGGATCCATGCTA

NV = novel virus

Table 4.10 Novel HPV detection in archived tissue blocks.

NV type	OSCC	OPSCC
14.4	1/121 (0.8%)	0
69.1	0	0
95	0	0

Table 4.11 HPV16 status in oral cavity and oropharyngeal archived biopsy samples.

	Oral cavity			Oropharynx			
	total n	n HPV+	% HPV+	total n	n HPV+	% HPV+	p-value
Cancer+	65	6	9%	50	39	79%	<0.001
Cancer-	56	9	16%	50	1	2%	0.018*

p-value comparing % HPV positive in oral cavity vs. oropharynx within each subject group.
*Fisher's exact test

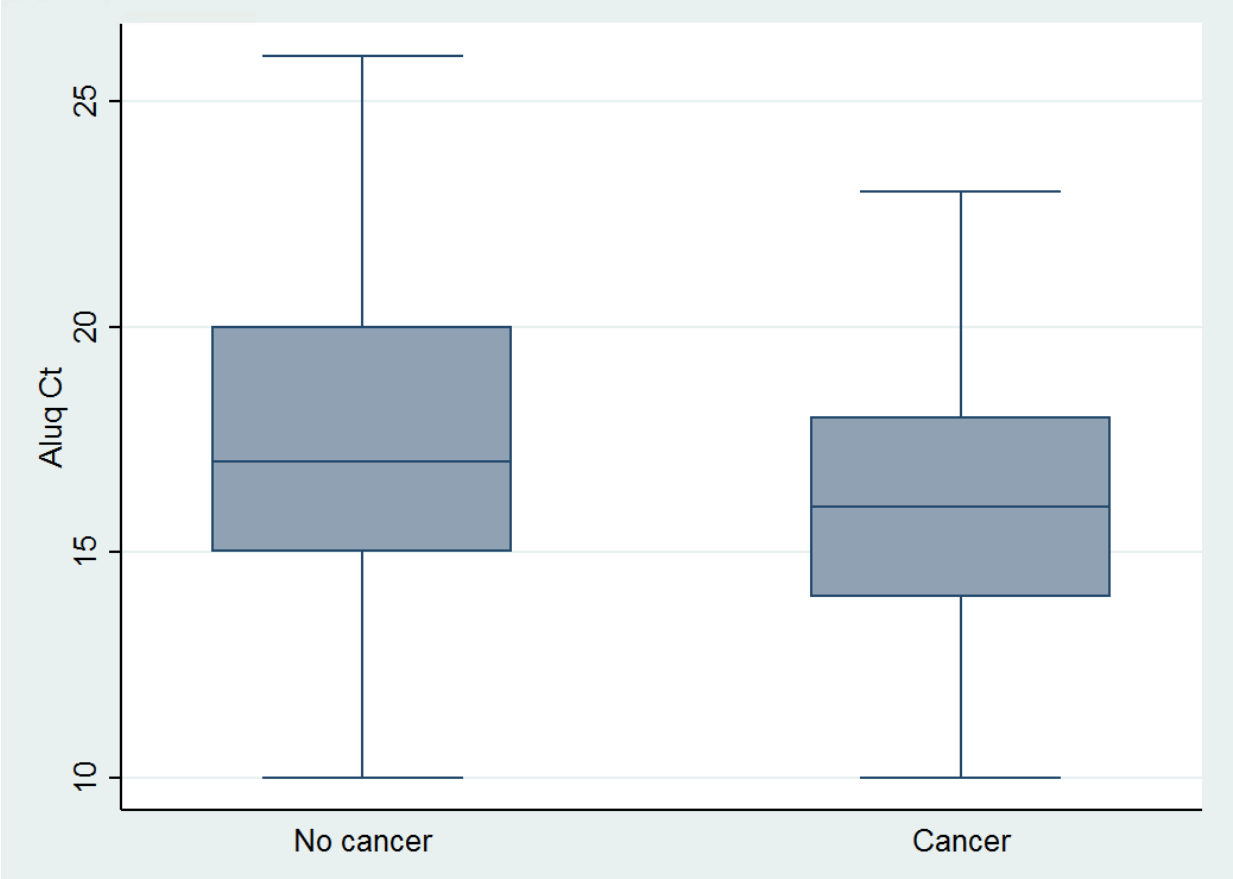

Figure 4.8 Aluq Ct values for archived tissue biopsy samples. Ct of <30 indicates sufficient amount of human DNA. No cancer: n=106; cancer: n=115.

Summary of results for aim 2

We created primers and probes for the three novel viruses and used RT-PCR for detection in archived tissue blocks. 39% of cancer patients had HPV16 compared to only 9% of healthy patients ($p<0.001$). No patients were positive for HPV18. From the Aluq Ct values we see that there was a sufficient amount of human DNA obtained from the samples for HPV detection as the Ct values remained under 30 (Figure 4.8).

Specific Aim 3 results

With the same primers and probes created for Aim 2, we tested three novel viruses in all of the oral rinse samples using RT-PCR Taqman assays (Table 4.12).

Table 4.12 HPV detection in oral rinse samples.

HPV type	No cancer	OPSCC	OSCC	HNSCC
NV14.4	0	1/76 (1%)	1/16 (6%)	0
NV69.1	0	10/76 (13%)	2/16 (12.5%)	1/8 (12.5%)
NV95	0	1/76 (1%)	1/16 (6%)	0
HPV16	0	19/76 (25%)	2/16 (12.5%)	2/8 (25%)
HPV18	1/109 (0.9%)	0	0	0

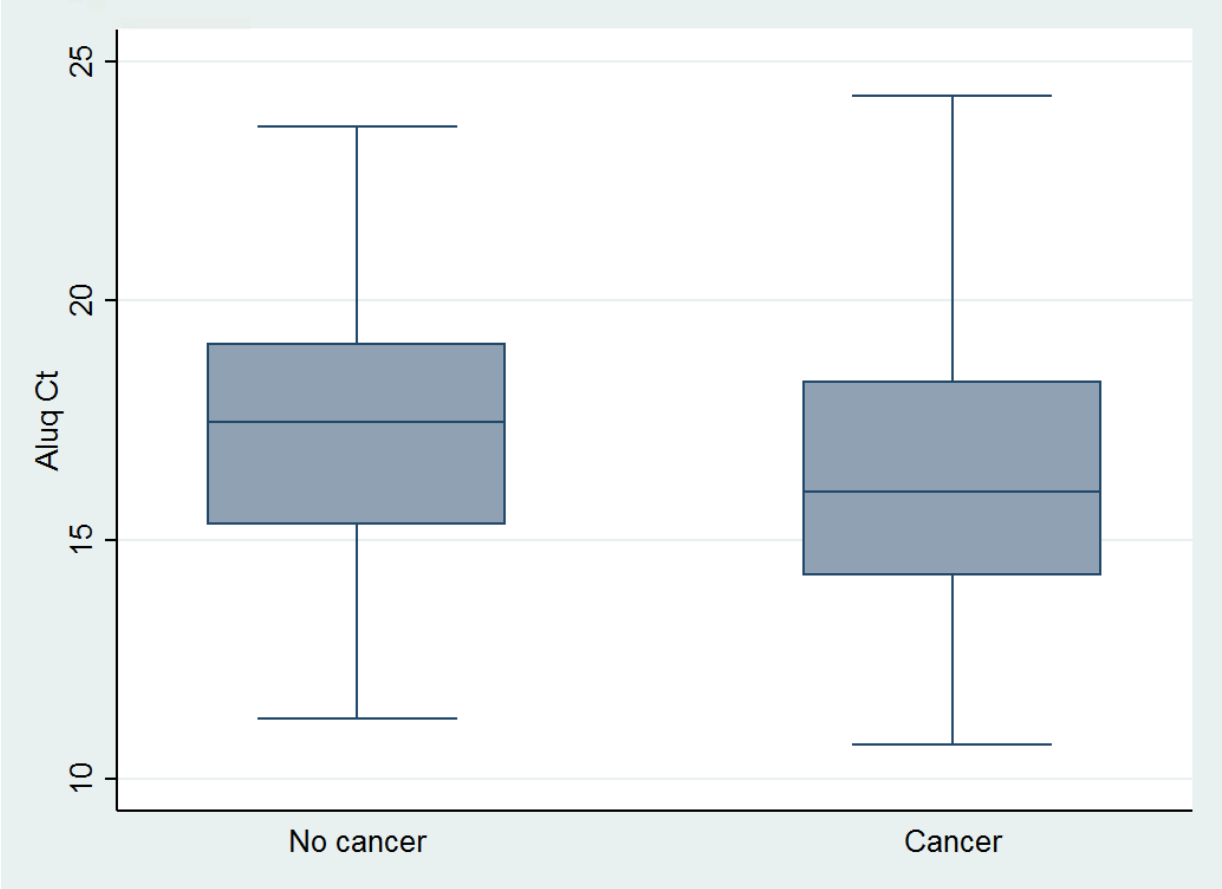

Figure 4.9 Aluq Ct values for oral rinse samples. Ct of <30 indicates sufficient amount of human DNA. No cancer: n=110; cancer: n=100.

Summary of results for aim 3

With the newly created primers and probes, we used RT-PCR to look at prevalence in all of the oral rinse samples. 2% of patients were positive for NV14.4; 13% were positive for 69.1; and 2% were positive for NV95. None of the non-cancer samples were positive for any of the novel viruses (Table 4.12). From the Aluq Ct values we see that there was a sufficient amount of human DNA obtained from the samples for HPV detection as the Ct values remained under 30 (Figure 4.9).

Chapter V

Discussion & Conclusions

Specific aim 1a discussion

Oral rinse sample collection is a good non-invasive method to use for the clinic. However, there are limitations as the patient may not swish/gargle as effectively as needed, thus an insufficient number of cells may be collected. However, from our Aluq detection results, which acted as a house-keeping gene for the normalization of human DNA, which seemed to be quite sufficient. Drawbacks of the mouthwash used include the mint flavor may be too strong for sensitive mouths, and the presence of alcohol in Scope, which is not suitable for recovering alcoholics.

NGS technology is a successful approach for detecting novel HPVs, but difficulties were encountered with the creation of contigs and removal of non-HPV sequences such as bacterial DNA. We pooled the samples together so as to stay within our budget, but it would have been best to only run one sample at a time so as to not receive a plethora of information that was difficult to analyze. From the read results it is evident that there was too much data in some samples to obtain clear results. Thus, we opted to perform DNA cloning and Sanger sequencing in conjunction to NGS in order to obtain clearer data. Upon reflection, NGS technology is a good method to have if an overall picture is needed, but if more specific information is needed, such as a clear sequence set, regular sequencing following cloning is more effective though more time consuming.

The MP-RCA method was effective in preferentially amplifying circular HPV16 DNA almost 800,000 fold, and very minimally amplifying linear DNA. Thus, we are confident that the unknown HPVs were amplified effectively.

Specific aim 1b discussion

The use of cloning and Sanger sequencing was an extremely useful method to determine clear and concise sequences. Our NGS results did not produce the results we wanted because we limited ourselves to the L1 region with FAP PCR. It is possible that we may have missed multiple infections in some samples due to limiting our sampling to five colonies for each sample, a drawback of the experimental design. FAP PCR led to some ambiguous results as faint bands seemed to indicate a positive sample, yet upon re-testing the sample a band was not observed. Thus, we excluded three samples that we once thought were HPV positive. Although, it is possible HPV DNA does exist, but perhaps in very low quantity.

The phylogenetic tree indicates that NV69.1 is not related to any of the 175 HPVs, although it may distantly be related to alphapapillomaviruses (α-PV). NV14.4 is related to the betapapillomavirus (β-PV) genus, and NV95 to the gammapapillomavirus (γ-PV) genus that are both common on human skin. β-PV has been suggested to play a role in the pathogenesis of SCC [107, 108], and γ-PV was reported to be isolated from the SCC of immunosuppressed patients [109, 110]. Thus, it would not be surprising if our novel HPVs were oncogenic in nature. This could be further studied through examination of E6/E7 regions of these novel HPVs.

Specific aim 2 discussion

Only one sample from the archived oral tissue biopsy samples tested positive for a novel HPVs, which was NV14.4. A possible problem with biopsy samples is if the HPV is not integrated with the host DNA (ie. it is in an episomal state) [111], this situation may make it difficult to collect the viral DNA as the biopsy may miss the area in which the HPV resides. The Alu results demonstrated sufficient amount of human DNA. However, older tissue blocks could be prone to DNA degradation.

We were able to make comparisons between non-cancer and cancer biopsy samples and found that of the OSCC cases only 9% had HPV16 infection compared to 79% of OPSCC cases ($p<0.001$), thus demonstrating a statistical difference. It should also be noted that of the oral cavity cases, the non-cancer patients had a slightly higher prevalence of HPV16 infection compared to cancer cases (16% vs. 9%; $p=0.018$). Within the literature this is usually not the case, thus perhaps it was our sampling method, which produced this result.

Specific aim 3 discussion

Out of the three novel viruses, NV69.1 was the most prevalent within the oral rinse samples. The genus to which it belongs is still not clear. From the phylogenetic tree it appears to be distantly related to α-PVs, which are mainly isolated from mucosa. HPV16 is an α-PV, and is the type that is mainly associated with OPSCC [46, 112]. Thus, it is possible that NV69.1 could be an oncogenic HPV. NV14.4 and NV95 may also be oncogenic, but further investigations are needed to determine this.

Conclusions

NGS is a useful tool for identifying novel HPVs in conjunction with MP-RCA and FAP PCR. Although, it is an expensive technique, thus precaution should be taken when choosing which samples to test. From our experience we do not recommend pooling of samples if possible as the data obtained is enormous, and is difficult to sift through if the concentration of HPV DNA is insufficient.

Through cloning and Sanger sequencing we were able to identify three potential novel HPVS. The prevalence of the NVs was very low in the archived tissue biopsies as only one NV was detected in one sample. In the oral rinse samples we were able to detect NV14.4 in two samples; NV69.1 in twelve samples; and NV95 in two samples.

Further investigations should include determining whether the potential novel HPVs are oncogenic, which can be accomplished by detecting for E6/E7 mRNA expression in tumour biopsies. Also, sequencing of the entire genome would be ideal in order to characterize it as a novel HPV.

In conclusion, oral rinse sample collection is effective for identifying new types of HPVs from the oral cavity and oropharynx.

References

1. Dell G, Gaston K: **Human papillomaviruses and their role in cervical cancer**. *Cellular and molecular life sciences : CMLS* 2001, **58**(12-13):1923-1942.
2. Bernard HU, Burk RD, Chen Z, van Doorslaer K, zur Hausen H, de Villiers EM: **Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments**. *Virology* 2010, **401**(1):70-79.
3. zur Hausen H: **Papillomaviruses in the causation of human cancers - a brief historical account**. *Virology* 2009, **384**(2):260-265.
4. Munger K, Baldwin A, Edwards KM, Hayakawa H, Nguyen CL, Owens M, Grace M, Huh K: **Mechanisms of human papillomavirus-induced oncogenesis**. *Journal of virology* 2004, **78**(21):11451-11460.
5. Chung CH, Gillison ML: **Human papillomavirus in head and neck cancer: its role in pathogenesis and clinical implications**. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2009, **15**(22):6758-6762.
6. Howard JD, Chung CH: **Biology of human papillomavirus-related oropharyngeal cancer**. *Seminars in radiation oncology* 2012, **22**(3):187-193.
7. Crook T, Tidy JA, Vousden KH: **Degradation of p53 can be targeted by HPV E6 sequences distinct from those required for p53 binding and trans-activation**. *Cell* 1991, **67**(3):547-556.
8. Slebos RJ, Lee MH, Plunkett BS, Kessis TD, Williams BO, Jacks T, Hedrick L, Kastan MB, Cho KR: **p53-dependent G1 arrest involves pRB-related proteins and is disrupted by the human papillomavirus 16 E7 oncoprotein**. *Proceedings of the National Academy of Sciences of the United States of America* 1994, **91**(12):5320-5324.
9. Munger K, Scheffner M, Huibregtse JM, Howley PM: **Interactions of HPV E6 and E7 oncoproteins with tumour suppressor gene products**. *Cancer surveys* 1992, **12**:197-217.
10. Klingelhutz AJ, Roman A: **Cellular transformation by human papillomaviruses: lessons learned by comparing high- and low-risk viruses**. *Virology* 2012, **424**(2):77-98.
11. Doorbar J, Quint W, Banks L, Bravo IG, Stoler M, Broker TR, Stanley MA: **The biology and life-cycle of human papillomaviruses**. *Vaccine* 2012, **30 Suppl 5**:F55-70.
12. Gillison ML, Castellsague X, Chaturvedi A, Goodman MT, Snijders P, Tommasino M, Arbyn M, Franceschi S: **Comparative epidemiology of HPV infection and associated cancers of the head and neck and cervix**. *International journal of cancer Journal international du cancer* 2013.
13. Haddad R, Crum C, Chen Z, Krane J, Posner M, Li Y, Burk R: **HPV16 transmission between a couple with HPV-related head and neck cancer**. *Oral oncology* 2008, **44**(8):812-815.
14. Woodman CB, Collins SI, Young LS: **The natural history of cervical HPV infection: unresolved issues**. *Nature reviews Cancer* 2007, **7**(1):11-22.

15. Parkin DM, Bray F, Ferlay J, Pisani P: **Global cancer statistics, 2002**. *CA: a cancer journal for clinicians* 2005, **55**(2):74-108.
16. Syrjanen S: **Human papillomavirus (HPV) in head and neck cancer**. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* 2005, **32 Suppl 1**:S59-66.
17. Pannone G, Rodolico V, Santoro A, Lo Muzio L, Franco R, Botti G, Aquino G, Pedicillo MC, Cagiano S, Campisi G *et al*: **Evaluation of a combined triple method to detect causative HPV in oral and oropharyngeal squamous cell carcinomas: p16 Immunohistochemistry, Consensus PCR HPV-DNA, and In Situ Hybridization**. *Infectious agents and cancer* 2012, **7**:4.
18. Marur S, D'Souza G, Westra WH, Forastiere AA: **HPV-associated head and neck cancer: a virus-related cancer epidemic**. *The lancet oncology* 2010, **11**(8):781-789.
19. Ryerson AB, Peters ES, Coughlin SS, Chen VW, Gillison ML, Reichman ME, Wu X, Chaturvedi AK, Kawaoka K: **Burden of potentially human papillomavirus-associated cancers of the oropharynx and oral cavity in the US, 1998-2003**. *Cancer* 2008, **113**(10 Suppl):2901-2909.
20. Chaturvedi AK, Engels EA, Anderson WF, Gillison ML: **Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States**. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2008, **26**(4):612-619.
21. Herrero R, Castellsague X, Pawlita M, Lissowska J, Kee F, Balaram P, Rajkumar T, Sridhar H, Rose B, Pintos J *et al*: **Human papillomavirus and oral cancer: the International Agency for Research on Cancer multicenter study**. *Journal of the National Cancer Institute* 2003, **95**(23):1772-1783.
22. Hobbs CG, Sterne JA, Bailey M, Heyderman RS, Birchall MA, Thomas SJ: **Human papillomavirus and head and neck cancer: a systematic review and meta-analysis**. *Clinical otolaryngology : official journal of ENT-UK ; official journal of Netherlands Society for Oto-Rhino-Laryngology & Cervico-Facial Surgery* 2006, **31**(4):259-266.
23. D'Souza G, Kreimer AR, Viscidi R, Pawlita M, Fakhry C, Koch WM, Westra WH, Gillison ML: **Case-control study of human papillomavirus and oropharyngeal cancer**. *The New England journal of medicine* 2007, **356**(19):1944-1956.
24. de Martel C, Ferlay J, Franceschi S, Vignat J, Bray F, Forman D, Plummer M: **Global burden of cancers attributable to infections in 2008: a review and synthetic analysis**. *The lancet oncology* 2012, **13**(6):607-615.
25. Rietbergen MM, Leemans CR, Bloemena E, Heideman DA, Braakhuis BJ, Hesselink AT, Witte BI, Baatenburg de Jong RJ, Meijer CJ, Snijders PJ *et al*: **Increasing prevalence rates of HPV attributable oropharyngeal squamous cell carcinomas in the Netherlands as assessed by a validated test algorithm**. *International journal of cancer Journal international du cancer* 2013, **132**(7):1565-1571.
26. Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, Jiang B, Goodman MT, Sibug-Saber M, Cozen W *et al*: **Human papillomavirus and**

rising oropharyngeal cancer incidence in the United States. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2011, **29**(32):4294-4301.
27. Hammarstedt L, Lindquist D, Dahlstrand H, Romanitan M, Dahlgren LO, Joneberg J, Creson N, Lindholm J, Ye W, Dalianis T *et al*: **Human papillomavirus as a risk factor for the increase in incidence of tonsillar cancer**. *International journal of cancer Journal international du cancer* 2006, **119**(11):2620-2623.
28. Attner P, Du J, Nasman A, Hammarstedt L, Ramqvist T, Lindholm J, Marklund L, Dalianis T, Munck-Wikland E: **The role of human papillomavirus in the increased incidence of base of tongue cancer**. *International journal of cancer Journal international du cancer* 2010, **126**(12):2879-2884.
29. Nasman A, Attner P, Hammarstedt L, Du J, Eriksson M, Giraud G, Ahrlund-Richter S, Marklund L, Romanitan M, Lindquist D *et al*: **Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: an epidemic of viral-induced carcinoma?** *International journal of cancer Journal international du cancer* 2009, **125**(2):362-366.
30. Auluck A, Hislop G, Bajdik C, Poh C, Zhang L, Rosin M: **Trends in oropharyngeal and oral cavity cancer incidence of human papillomavirus (HPV)-related and HPV-unrelated sites in a multicultural population: the British Columbia experience**. *Cancer* 2010, **116**(11):2635-2644.
31. Hocking JS, Stein A, Conway EL, Regan D, Grulich A, Law M, Brotherton JM: **Head and neck cancer in Australia between 1982 and 2005 show increasing incidence of potentially HPV-associated oropharyngeal cancers**. *British journal of cancer* 2011, **104**(5):886-891.
32. Turner CF, Danella RD, Rogers SM: **Sexual behavior in the United States 1930-1990: trends and methodological problems**. *Sexually transmitted diseases* 1995, **22**(3):173-190.
33. Herbenick D, Reece M, Schick V, Sanders SA, Dodge B, Fortenberry JD: **Sexual behavior in the United States: results from a national probability sample of men and women ages 14-94**. *The journal of sexual medicine* 2010, **7 Suppl 5**:255-265.
34. Bajos N, Bozon M, Beltzer N, Laborde C, Andro A, Ferrand M, Goulet V, Laporte A, Le Van C, Leridon H *et al*: **Changes in sexual behaviours: from secular trends to public health policies**. *AIDS* 2010, **24**(8):1185-1191.
35. Rintala M, Grenman S, Puranen M, Syrjanen S: **Natural history of oral papillomavirus infections in spouses: a prospective Finnish HPV Family Study**. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* 2006, **35**(1):89-94.
36. Pickard RK, Xiao W, Broutian TR, He X, Gillison ML: **The prevalence and incidence of oral human papillomavirus infection among young men and women, aged 18-30 years**. *Sexually transmitted diseases* 2012, **39**(7):559-566.

37. D'Souza G, Agrawal Y, Halpern J, Bodison S, Gillison ML: **Oral sexual behaviors associated with prevalent oral human papillomavirus infection**. *The Journal of infectious diseases* 2009, **199**(9):1263-1269.
38. Schiffman M, Wentzensen N, Wacholder S, Kinney W, Gage JC, Castle PE: **Human papillomavirus testing in the prevention of cervical cancer**. *Journal of the National Cancer Institute* 2011, **103**(5):368-383.
39. Syrjanen S: **HPV infections and tonsillar carcinoma**. *Journal of clinical pathology* 2004, **57**(5):449-455.
40. Hormia M, Willberg J, Ruokonen H, Syrjanen S: **Marginal periodontium as a potential reservoir of human papillomavirus in oral mucosa**. *Journal of periodontology* 2005, **76**(3):358-363.
41. Tezal M, Sullivan MA, Reid ME, Marshall JR, Hyland A, Loree T, Lillis C, Hauck L, Wactawski-Wende J, Scannapieco FA: **Chronic periodontitis and the risk of tongue cancer**. *Archives of otolaryngology--head & neck surgery* 2007, **133**(5):450-454.
42. Gillison ML, D'Souza G, Westra W, Sugar E, Xiao W, Begum S, Viscidi R: **Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers**. *Journal of the National Cancer Institute* 2008, **100**(6):407-420.
43. Kreimer AR, Pierce Campbell CM, Lin HY, Fulp W, Papenfuss MR, Abrahamsen M, Hildesheim A, Villa LL, Salmeron JJ, Lazcano-Ponce E *et al*: **Incidence and clearance of oral human papillomavirus infection in men: the HIM cohort study**. *Lancet* 2013, **382**(9895):877-887.
44. Fakhry C, Gillison ML, D'Souza G: **Tobacco use and oral HPV-16 infection**. *Jama* 2014, **312**(14):1465-1467.
45. Gillison ML, Broutian T, Pickard RK, Tong ZY, Xiao W, Kahle L, Graubard BI, Chaturvedi AK: **Prevalence of oral HPV infection in the United States, 2009-2010**. *JAMA : the journal of the American Medical Association* 2012, **307**(7):693-703.
46. Hariri S, Unger ER, Sternberg M, Dunne EF, Swan D, Patel S, Markowitz LE: **Prevalence of genital human papillomavirus among females in the United States, the National Health And Nutrition Examination Survey, 2003-2006**. *The Journal of infectious diseases* 2011, **204**(4):566-573.
47. Kreimer AR, Clifford GM, Boyle P, Franceschi S: **Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review**. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2005, **14**(2):467-475.
48. Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L, Zahurak ML, Daniel RW, Viglione M, Symer DE *et al*: **Evidence for a causal association between human papillomavirus and a subset of head and neck cancers**. *Journal of the National Cancer Institute* 2000, **92**(9):709-720.
49. de Sanjose S, Quint WG, Alemany L, Geraets DT, Klaustermeier JE, Lloveras B, Tous S, Felix A, Bravo LE, Shin HR *et al*: **Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study**. *The lancet oncology* 2010, **11**(11):1048-1056.

50. St Guily JL, Jacquard AC, Pretet JL, Haesebaert J, Beby-Defaux A, Clavel C, Agius G, Birembaut P, Okais C, Leocmach Y *et al*: **Human papillomavirus genotype distribution in oropharynx and oral cavity cancer in France--The EDiTH VI study**. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* 2011, **51**(2):100-104.
51. Zandberg DP, Bhargava R, Badin S, Cullen KJ: **The role of human papillomavirus in nongenital cancers**. *CA: a cancer journal for clinicians* 2013, **63**(1):57-81.
52. Fakhry C, Sugar E, D'Souza G, Gillison M: **Two-week versus six-month sampling interval in a short-term natural history study of oral HPV infection in an HIV-positive cohort**. *PloS one* 2010, **5**(7):e11918.
53. Jayaprakash V, Reid M, Hatton E, Merzianu M, Rigual N, Marshall J, Gill S, Frustino J, Wilding G, Loree T *et al*: **Human papillomavirus types 16 and 18 in epithelial dysplasia of oral cavity and oropharynx: a meta-analysis, 1985-2010**. *Oral oncology* 2011, **47**(11):1048-1054.
54. Rampias T, Sasaki C, Weinberger P, Psyrri A: **E6 and e7 gene silencing and transformed phenotype of human papillomavirus 16-positive oropharyngeal cancer cells**. *Journal of the National Cancer Institute* 2009, **101**(6):412-423.
55. Hafkamp HC, Speel EJ, Haesevoets A, Bot FJ, Dinjens WN, Ramaekers FC, Hopman AH, Manni JJ: **A subset of head and neck squamous cell carcinomas exhibits integration of HPV 16/18 DNA and overexpression of p16INK4A and p53 in the absence of mutations in p53 exons 5-8**. *International journal of cancer Journal international du cancer* 2003, **107**(3):394-400.
56. Mellin H, Dahlgren L, Munck-Wikland E, Lindholm J, Rabbani H, Kalantari M, Dalianis T: **Human papillomavirus type 16 is episomal and a high viral load may be correlated to better prognosis in tonsillar cancer**. *International journal of cancer Journal international du cancer* 2002, **102**(2):152-158.
57. Koskinen WJ, Chen RW, Leivo I, Makitie A, Back L, Kontio R, Suuronen R, Lindqvist C, Auvinen E, Molijn A *et al*: **Prevalence and physical status of human papillomavirus in squamous cell carcinomas of the head and neck**. *International journal of cancer Journal international du cancer* 2003, **107**(3):401-406.
58. Van Tine BA, Dao LD, Wu SY, Sonbuchner TM, Lin BY, Zou N, Chiang CM, Broker TR, Chow LT: **Human papillomavirus (HPV) origin-binding protein associates with mitotic spindles to enable viral DNA partitioning**. *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**(12):4030-4035.
59. Psyrri A, Gouveris P, Vermorken JB: **Human papillomavirus-related head and neck tumors: clinical and research implication**. *Current opinion in oncology* 2009, **21**(3):201-205.
60. Boyle JO, Hakim J, Koch W, van der Riet P, Hruban RH, Roa RA, Correo R, Eby YJ, Ruppert JM, Sidransky D: **The incidence of p53 mutations increases**

with progression of head and neck cancer. *Cancer research* 1993, **53**(19):4477-4480.
61. Reed AL, Califano J, Cairns P, Westra WH, Jones RM, Koch W, Ahrendt S, Eby Y, Sewell D, Nawroz H *et al*: **High frequency of p16 (CDKN2/MTS-1/INK4A) inactivation in head and neck squamous cell carcinoma**. *Cancer research* 1996, **56**(16):3630-3633.
62. Westra WH: **Detection of human papillomavirus (HPV) in clinical samples: evolving methods and strategies for the accurate determination of HPV status of head and neck carcinomas**. *Oral oncology* 2014, **50**(9):771-779.
63. Garcia-Closas M, Egan KM, Abruzzo J, Newcomb PA, Titus-Ernstoff L, Franklin T, Bender PK, Beck JC, Le Marchand L, Lum A *et al*: **Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash**. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2001, **10**(6):687-696.
64. Holmes BJ, Westra WH: **The expanding role of cytopathology in the diagnosis of HPV-related squamous cell carcinoma of the head and neck**. *Diagnostic cytopathology* 2014, **42**(1):85-93.
65. Fakhry C, Rosenthal BT, Clark DP, Gillison ML: **Associations between oral HPV16 infection and cytopathology: evaluation of an oropharyngeal "pap-test equivalent" in high-risk populations**. *Cancer prevention research* 2011, **4**(9):1378-1384.
66. Lingen MW: **Brush-based cytology screening in the tonsils and cervix: there is a difference!** *Cancer prevention research* 2011, **4**(9):1350-1352.
67. Smeets SJ, Hesselink AT, Speel EJ, Haesevoets A, Snijders PJ, Pawlita M, Meijer CJ, Braakhuis BJ, Leemans CR, Brakenhoff RH: **A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen**. *International journal of cancer Journal international du cancer* 2007, **121**(11):2465-2472.
68. Fakhry C, Westra WH, Li S, Cmelak A, Ridge JA, Pinto H, Forastiere A, Gillison ML: **Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial**. *Journal of the National Cancer Institute* 2008, **100**(4):261-269.
69. Stephen JK, Divine G, Chen KM, Chitale D, Havard S, Worsham MJ: **Significance of p16 in Site-specific HPV Positive and HPV Negative Head and Neck Squamous Cell Carcinoma**. *Cancer and clinical oncology* 2013, **2**(1):51-61.
70. Salazar CR, Anayannis N, Smith RV, Wang Y, Haigentz M, Jr., Garg M, Schiff BA, Kawachi N, Elman J, Belbin TJ *et al*: **Combined P16 and human papillomavirus testing predicts head and neck cancer survival**. *International journal of cancer Journal international du cancer* 2014, **135**(10):2404-2412.
71. Ramshankar V, Soundara VT, Shyamsundar V, Ramani P, Krishnamurthy A: **Risk Stratification of Early Stage Oral Tongue Cancers Based on HPV**

Status and p16 Immunoexpression. *Asian Pacific journal of cancer prevention : APJCP* 2014, **15**(19):8351-8359.
72. Lavaf A, Genden EM, Cesaretti JA, Packer S, Kao J: **Adjuvant radiotherapy improves overall survival for patients with lymph node-positive head and neck squamous cell carcinoma**. *Cancer* 2008, **112**(3):535-543.
73. Rogers L, Siu SS, Luesley D, Bryant A, Dickinson HO: **Radiotherapy and chemoradiation after surgery for early cervical cancer**. *The Cochrane database of systematic reviews* 2012, **5**:CD007583.
74. Cooper JS, Pajak TF, Forastiere AA, Jacobs J, Campbell BH, Saxman SB, Kish JA, Kim HE, Cmelak AJ, Rotman M *et al*: **Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck**. *The New England journal of medicine* 2004, **350**(19):1937-1944.
75. Bernier J, Domenge C, Ozsahin M, Matuszewska K, Lefebvre JL, Greiner RH, Giralt J, Maingon P, Rolland F, Bolla M *et al*: **Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer**. *The New England journal of medicine* 2004, **350**(19):1945-1952.
76. Peters WA, 3rd, Liu PY, Barrett RJ, 2nd, Stock RJ, Monk BJ, Berek JS, Souhami L, Grigsby P, Gordon W, Jr., Alberts DS: **Concurrent chemotherapy and pelvic radiation therapy compared with pelvic radiation therapy alone as adjuvant therapy after radical surgery in high-risk early-stage cancer of the cervix**. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2000, **18**(8):1606-1613.
77. Bansal N, Herzog TJ, Shaw RE, Burke WM, Deutsch I, Wright JD: **Primary therapy for early-stage cervical cancer: radical hysterectomy vs radiation**. *American journal of obstetrics and gynecology* 2009, **201**(5):485 e481-489.
78. Herrero R, Quint W, Hildesheim A, Gonzalez P, Struijk L, Katki HA, Porras C, Schiffman M, Rodriguez AC, Solomon D *et al*: **Reduced prevalence of oral human papillomavirus (HPV) 4 years after bivalent HPV vaccination in a randomized clinical trial in Costa Rica**. *PloS one* 2013, **8**(7):e68329.
79. Heath EM, Morken NW, Campbell KA, Tkach D, Boyd EA, Strom DA: **Use of buccal cells collected in mouthwash as a source of DNA for clinical testing**. *Archives of pathology & laboratory medicine* 2001, **125**(1):127-133.
80. D'Souza G, Sugar E, Ruby W, Gravitt P, Gillison M: **Analysis of the effect of DNA purification on detection of human papillomavirus in oral rinse samples by PCR**. *Journal of clinical microbiology* 2005, **43**(11):5526-5535.
81. Frisch M, Biggar RJ, Goedert JJ: **Human papillomavirus-associated cancers in patients with human immunodeficiency virus infection and acquired immunodeficiency syndrome**. *Journal of the National Cancer Institute* 2000, **92**(18):1500-1510.
82. Kreimer AR, Alberg AJ, Daniel R, Gravitt PE, Viscidi R, Garrett ES, Shah KV, Gillison ML: **Oral human papillomavirus infection in adults is associated with sexual behavior and HIV serostatus**. *The Journal of infectious diseases* 2004, **189**(4):686-698.

83. Miller CS, White DK: **Human papillomavirus expression in oral mucosa, premalignant conditions, and squamous cell carcinoma: a retrospective review of the literature**. *Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics* 1996, **82**(1):57-68.
84. Kansy K, Thiele O, Freier K: **The role of human papillomavirus in oral squamous cell carcinoma: myth and reality**. *Oral and maxillofacial surgery* 2012.
85. Johne R, Muller H, Rector A, van Ranst M, Stevens H: **Rolling-circle amplification of viral DNA genomes using phi29 polymerase**. *Trends in microbiology* 2009, **17**(5):205-211.
86. Rector A, Tachezy R, Van Ranst M: **A sequence-independent strategy for detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle amplification**. *Journal of virology* 2004, **78**(10):4993-4998.
87. Dean FB, Nelson JR, Giesler TL, Lasken RS: **Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification**. *Genome research* 2001, **11**(6):1095-1099.
88. Esteban JA, Salas M, Blanco L: **Fidelity of phi 29 DNA polymerase. Comparison between protein-primed initiation and DNA polymerization**. *The Journal of biological chemistry* 1993, **268**(4):2719-2726.
89. Nelson JR, Cai YC, Giesler TL, Farchaus JW, Sundaram ST, Ortiz-Rivera M, Hosta LP, Hewitt PL, Mamone JA, Palaniappan C *et al*: **TempliPhi, phi29 DNA polymerase based rolling circle amplification of templates for DNA sequencing**. *BioTechniques* 2002, **Suppl**:44-47.
90. Stanley MA, Browne HM, Appleby M, Minson AC: **Properties of a non-tumorigenic human cervical keratinocyte cell line**. *International journal of cancer Journal international du cancer* 1989, **43**(4):672-676.
91. Forslund O, Antonsson A, Nordin P, Stenquist B, Hansson BG: **A broad range of human papillomavirus types detected with a general PCR method suitable for analysis of cutaneous tumours and normal skin**. *The Journal of general virology* 1999, **80 (Pt 9)**:2437-2443.
92. Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M: **Comparison of next-generation sequencing systems**. *Journal of biomedicine & biotechnology* 2012, **2012**:251364.
93. Collins FS, Morgan M, Patrinos A: **The Human Genome Project: lessons from large-scale biology**. *Science* 2003, **300**(5617):286-290.
94. Meiring TL, Salimo AT, Coetzee B, Maree HJ, Moodley J, Hitzeroth, II, Freeborough MJ, Rybicki EP, Williamson AL: **Next-generation sequencing of cervical DNA detects human papillomavirus types not detected by commercial kits**. *Virology journal* 2012, **9**:164.
95. Radford AD, Chapman D, Dixon L, Chantrey J, Darby AC, Hall N: **Application of next-generation sequencing technologies in virology**. *The Journal of general virology* 2012, **93**(Pt 9):1853-1868.

96. Eklund C, Zhou T, Dillner J, Network WHOHPL: **Global proficiency study of human papillomavirus genotyping**. *Journal of clinical microbiology* 2010, **48**(11):4147-4155.
97. Mori S, Nakao S, Kukimoto I, Kusumoto-Matsuo R, Kondo K, Kanda T: **Biased amplification of human papillomavirus DNA in specimens containing multiple human papillomavirus types by PCR with consensus primers**. *Cancer science* 2011, **102**(6):1223-1227.
98. Schmitt M, Dondog B, Waterboer T, Pawlita M, Tommasino M, Gheit T: **Abundance of multiple high-risk human papillomavirus (HPV) infections found in cervical cells analyzed by use of an ultrasensitive HPV genotyping assay**. *Journal of clinical microbiology* 2010, **48**(1):143-149.
99. Martin E, Dang J, Bzhalava D, Stern J, Edelstein ZR, Koutsky LA, Kiviat NB, Feng Q: **Characterization of three novel human papillomavirus types isolated from oral rinse samples of healthy individuals**. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology* 2014, **59**(1):30-37.
100. Edelstein ZR, Schwartz SM, Hawes S, Hughes JP, Feng Q, Stern ME, O'Reilly S, Lee SK, Fu Xi L, Koutsky LA: **Rates and determinants of oral human papillomavirus infection in young men**. *Sexually transmitted diseases* 2012, **39**(11):860-867.
101. de Leon DC, Montiel DP, Nemcova J, Mykyskova I, Turcios E, Villavicencio V, Cetina L, Coronel A, Hes O: **Human papillomavirus (HPV) in breast tumors: prevalence in a group of Mexican patients**. *BMC cancer* 2009, **9**:26.
102. Berkhout RJ, Tieben LM, Smits HL, Bavinck JN, Vermeer BJ, ter Schegget J: **Nested PCR approach for detection and typing of epidermodysplasia verruciformis-associated human papillomavirus types in cutaneous cancers from renal transplant recipients**. *Journal of clinical microbiology* 1995, **33**(3):690-695.
103. de Koning M, Quint W, Struijk L, Kleter B, Wanningen P, van Doorn LJ, Weissenborn SJ, Feltkamp M, ter Schegget J: **Evaluation of a novel highly sensitive, broad-spectrum PCR-reverse hybridization assay for detection and identification of beta-papillomavirus DNA**. *Journal of clinical microbiology* 2006, **44**(5):1792-1800.
104. Drummond AJ, Rambaut A: **BEAST: Bayesian evolutionary analysis by sampling trees**. *BMC evolutionary biology* 2007, **7**:214.
105. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H: **Classification of papillomaviruses**. *Virology* 2004, **324**(1):17-27.
106. Chen Z, Freitas LB, Burk RD: **Evolution and classification of oncogenic human papillomavirus types and variants associated with cervical cancer**. *Methods in molecular biology* 2015, **1249**:3-26.
107. Karagas MR, Nelson HH, Sehr P, Waterboer T, Stukel TA, Andrew A, Green AC, Bavinck JN, Perry A, Spencer S *et al*: **Human papillomavirus infection and incidence of squamous cell and basal cell carcinomas of the skin**. *Journal of the National Cancer Institute* 2006, **98**(6):389-395.

108. Asgari MM, Kiviat NB, Critchlow CW, Stern JE, Argenyi ZB, Raugi GJ, Berg D, Odland PB, Hawes SE, de Villiers EM: **Detection of human papillomavirus DNA in cutaneous squamous cell carcinoma among immunocompetent individuals**. *J Invest Dermatol* 2008, **128**(6):1409-1417.
109. Kullander J, Handisurya A, Forslund O, Geusau A, Kirnbauer R, Dillner J: **Cutaneous human papillomavirus 88: remarkable differences in viral load**. *International journal of cancer Journal international du cancer* 2008, **122**(2):477-480.
110. Muller M, Kelly G, Fiedler M, Gissmann L: **Human papillomavirus type 48**. *Journal of virology* 1989, **63**(11):4907-4908.
111. Mellin H, Dahlgren L, Munck-Wikland E, Lindholm J, Rabbani H, Kalantari M, Dalianis T: **Human papillomavirus type 16 is episomal and a high viral load may be correlated to better prognosis in tonsillar cancer**. *International journal of cancer Journal international du cancer* 2002, **102**(2):152-158.
112. Gillison ML, Broutian T, Pickard RK, Tong ZY, Xiao W, Kahle L, Graubard BI, Chaturvedi AK: **Prevalence of oral HPV infection in the United States, 2009-2010**. *Jama* 2012, **307**(7):693-703.

Juliet Dang Curriculum Vitae
jhtdang@uw.edu

SUMMARY OF QUALIFICATIONS

- Translational research experience through 5 years of PhD graduate studies
- 9 years of experience as a Registered Clinical Dental Hygienist
- Recognized for outstanding communication with both patients and management
- Exceptional leadership and mentoring skills
- Consistently involved in detail-oriented projects, which includes deep organization, planning, writing, and problem solving
- Strong oral presenter demonstrated by over 12 years of experience in the performing arts

EDUCATION

PhD in Oral Biology — June 2015
Thesis: Identification and characterization of novel HPVs in oral rinse samples.
School of Dentistry
University of Washington, Seattle, WA

Master of Oral Biology — June 2010
School of Dentistry
University of Washington, Seattle, WA

Diploma of Dental Hygiene — April 2006
School of Dentistry
University of Manitoba, Winnipeg, MB

Bachelor of Science — April 2002
Major: Biology, Minor: Chemistry
University of Winnipeg, Winnipeg, MB

ACADEMIC HONORS

- American Dental Hygienists Association, Dr. Esther Wilkins Scholarship
  July 2014
- Institute of Translational Health Sciences Scholar, University of Washington, Seattle, WA
  June 2013 – June 2014
- Magnuson Scholars Award, University of Washington, Seattle, WA
  March 2012
- Lifelong Learning Award, Washington State Dental Hygienists' Association, Seattle, WA
  October 2011

- Top Scholar Award for Graduates, University of Washington, Seattle, WA
  Fall Quarter 2010
- Visionary Scholarship, The Canadian Dental Hygienists Association, Seattle, WA

  Winter 2008
- School of Dentistry Scholarship, University of Manitoba, Winnipeg, MB
  January 2005
- Millennium Award, Government of Canada, Winnipeg, MB
  January 2004
- Top Students Award, Cambrian Credit Union, Winnipeg, MB
  June 2009
- Entrance Scholarship, University of Winnipeg, Winnipeg, MB
  September 1999

PUBLICATIONS

Academic Papers:

1. Juliet Dang, Nancy B. Kiviat, Qinghua Feng, Keith D. Eaton, Hona Jang. **Detection of HPVs type 16 and 18 in oral rinse samples from OPSCC and non-OPSCC patients**. *In submission to BMC Oral Health.*
2. Juliet Dang, Nancy B. Kiviat, Qinghua Feng. **Prevalence of HPVs type 16 and 18 within a dental student clinic setting.** *Canadian Journal of Dental Hygiene June 2015.*
3. Juliet Dang & Qinghua Feng**. Oral cavity cancer versus oropharyngeal cancer: A comparison of patients infected with HPV16**. *Submitted to International Journal of Dental Hygiene.*
4. Erin Martin, Juliet Dang, Davit Bzhalava, Joshua Stern, Zoe R Edelstein, Laura A. Koutsky, Nancy B. Kiviat, Qinghua Feng. **Characterization of four novel human papillomavirus types isolated from oral lavage samples of healthy individuals.** *J Clin Virology.* 2014 Jan; 59(1):30-7.

Conference and Presentations:

Oral Presenter
Title of presentation: Identification and characterization of novel HPVs in oropharyngeal squamous cell carcinoma (OPSCC)

- ❖ Lake Washington Dental Hygiene Society Meeting
  January 15th, 2015
  Bellevue, WA
- ❖ 3rd North American/Global Dental Hygiene Research Conference
  October 18th, 2014
  Bethesda, MD
- ❖ Washington State Dental Hygienists Symposium
  April 25th, 2014

Seattle, WA

- Canadian Dental Hygienists Association Conference
  October 4th, 2013
  Toronto, Ontario

  Title of presentation: Pilot study of oncogenic HPVs in oral lavage samples from HIV positive Senegal women
- American Association for the Advancement of Science, Pacific Division Annual Meeting
  June 26th, 2012
  Boise, Idaho
- Greater Seattle Dental Hygienists' Society General Meeting
  February 21st, 2012
  Seattle, WA
- Second North American/Global Dental Hygiene Research Conference
  October 21st, 2011
  Bethesda, MD

RESEARCH EXPERIENCE

Human Papillomavirus (HPV) Laboratory — September 2010 - Present
Dr. Nancy Kiviat, University of Washington
Projects:

- Identification and characterization of novel HPVs in human oropharyngeal squamous cell carcinoma (OPSCC)
- Pilot study of oncogenic HPVs in oral lavage samples from HIV positive Senegal women

Periodontics Research — September 2009 – June 2010
Dr. Frank Roberts, University of Washington
Project:

- Severe chronic neutropenia and oral health status

TEACHING EXPERIENCE

Teacher Assistant for the First Year Dental Students Clinic — Winter 2009
School of Dentistry, University of Washington, Seattle, WA

Teacher Assistant for Undergraduate Biology Lab — Fall 2002
University of Winnipeg, Winnipeg, MB

WORK EXPERIENCE

Clinical Dental Hygienist — June 2006 – Present
Seattle, WA
Winnipeg, MB

LEADERSHIP EXPERIENCE

- Editor and Writer, Xin Chao Seattle Magazine — June 2013-Present
- Marketing Director, Alar Productions — March 2012-March 2013
- Marketing Director, Miss Vietnam Washington Association — March 2012-Present
- Vice President of The Dental Hygiene Class of 2006 — August 2005-April 2006
- President of The Dental Hygiene Class of 2006 — August 2004-April 2005

COMMUNITY SERVICE EXPERIENCE

- Mentor for The Miss Vietnam WA Association — March 2012-Present
- Tutor/Mentor for University of Washington's Dream Project — October 2011
  Seattle, WA
- Volunteer for Vietnamese Community Fundraisers — January 2009-Present
  Seattle, WA
- Volunteer at The Children's Hospital — May 2001-October 2003
  Winnipeg, MB

PROFESSIONAL AFFILIATIONS & MEMBERSHIPS

- Washington State Dental Hygienists' Association
- American Dental Hygienists Association
- College of Dental Hygienists of Manitoba
- American Dental Education Association
- American Association for the Advancement of Science

CONFERENCES ATTENDED

- HPV 2014 Conference
  August 2014
  Seattle, WA
- Translational Science 2014 Meeting
  April 2014
  Washington D.C.
- Canadian Dental Hygienists Association Annual Conference
  October 2013
  Toronto, Ontario

- American Association for the Advancement of Science, Pacific Division Annual Meeting
  June 2012
  Boise, Idaho
- North American/Global Dental Hygiene Research Conference
  October 2011
  June 2009
  NIH Campus, Bethesda, MD
- Ethical Considerations in Research Collaborations
  September 2011
  University of Washington, Seattle, WA
- Manitoba Dental Association Conference
  2005 and 2006
  Winnipeg, MB

This doctoral dissertation is dedicated to my mother Nga (Anna) Thi Thai who always made education my number one priority, which was hers as well

Je t'aime.