Norwegian University
of Life Sciences

Master's Thesis 2019 60 ECTS
Faculty of Chemistry, Biotechnology and Food Science

Evaluation and optimization of procedures for HPV detection and genotyping in oral and urine samples in men

William Nicholas Tourniaire
Biotechnology with a focus on Microbiology

Foreword

This study was conducted at the Norwegian national HPV reference laboratory at Akerhus University Hospital (Ahus) between August 2018 and May 2019. The study is a master's thesis constituting a part of a master's degree in biotechnology at the faculty of chemistry, biotechnology and food science, at the Norwegian University of Life Sciences.
The study was conducted for and financed by Department of Microbiology and Infection Control, the Norwegian national HPV reference laboratory, Ahus and Olafiaklinikken. The work of this study was requested as a foundation for future research by the Norwegian national HPV reference laboraotory, Olafiaklinikken and the University of Tromsø

Primary supervisor: Christine M. Jonassen, NMBU.

External supervisor and project leader: Irene K. Christiansen, Ahus.

Secondary supervisors: Ole Herman Ambur, OsloMet and Trine Rounge, Kreftregisteret.

I would like to thank my supervisors, for counselling, advice and guidance throughout this research project. I would like to especially express gratitude to my external supervisor Irene K. Christiansen for arranging this study, for excellent counselling and guidance, and for allowing me to conduct my work at the HPV reference laboratory. I would also like to express gratitude to Mona Hansen, Hanne Kristiansen-Haugland and Sonja Lagström for my training in laboratory analysis, guidance in laboratory work and for the work they conducted as part of this study.
I want to thank Olafiaklinikken for sample collection and cooperation throughout the progression of the study. I want to especially express gratitude to Anne O. Olsen in charge of Olafiaklinikkens work contributed to the study, and to chief study nurse Marie Aasterød in charge of patient recruitment and sample collection, instruction and training of colleagues. I am also grateful to Patricia Merckoll and Sindre Ringvik for their contributions.

Oslo, Mai 2019

…….....

William Nicholas Tourniaire

Abstract

Human papillomavirus (HPV) has been well established as a carcinogen in human epithelial cancers; especially in the *uterin cervix*. Head and neck cancers (HNC) are increasing worldwide despite global reduction in smoking. The increase of HNCs has been linked to the increasing rate of HPV related oropharyngeal squamous cell carcinoma (OPSCC) and other HPV related cancers of the head and neck. To better understand cancer development and possible prevention of HPV related OPSCCs, as well as epidemiology of oral HPV and HPV related OPSCC, it is important to lay solid foundations for large scale studies.

This study aims to establish a protocol for sample collection, DNA extraction, detection and genotyping of HPV in oral and urine samples from men. To achieve this we seek to compare two types of oral sample methods, oral rinse and buccal brush; optimising an extraction protocol for oral and urine samples; compare the analytic capabilities of two HPV detection and genotyping methods, Luminex and NGS; and to obtain a brief insight in the HPV prevalence of the study population.

There were 138 men who have sex with men (MSM) included among patients seeking medical assistance at the sexual health clinic Olafiaklinikken in Oslo. Frome these we analysed 136 buccal brushes, 135 oral rinse and 98 urine samples. The HPV prevalence found with Luminex was: 2.94%, 4.45% and 17.34% for buccal brush, oral rinse and urine samples, respectively. With NGS the prevalence was: 8.09%, 6.66%, 17.34% for buccal brush, oral rinse and urine samples, respectively. Only 21% (4/19) of patients with a positive oral sample had the same genotype in both sample materials. The Cohen's Kappa values for concordance between the two detection methods in the buccal brush, oral rinse and urine samples were, 0.31, 0.60 and 0.40 respectively.

In all, the number of HPV positive samples in this study was too small to make any concise declarations. No sample material or detection method has been shown as clearly superior. However, there are indications of oral rinse as better suited sample material and NGS as a more sensitive detection method.

Sammendrag

Humant papillomavirus (HPV) kan forårsake ulike kreftformer og for livmorhalskreft er det funnet at viruset er nødvendig for utvikling av kreftformen. Kreft i hode-hals-regionen øker globalt, til tross for reduksjon av røykere. Denne økningen har blitt knyttet til økningen av HPV relatert plateepitelkarsinom i oropharynx og andre HPV relaterte kreftformer i hode-hals-regionen. For å få en bedre forståelse av kreftutvikling og mulige preventive tiltak for utvikling av HPV relatert kreft, samt epidemiologisk kjennskap til oral HPV og HPV relatert kreft, trengs det grunnleggende studier som kan tilrettelegge for større forskningsprosjekter. Dette studiet har som mål å etablere en protokoll for prøvetakning, DNA ekstraksjon, deteksjon og genotyping av HPV i oral og urin prøver fra menn. For å oppnå dette ønsker vi: å sammenligne to typer orale prøvetyper, munnskylleprøver og børsteprøve av kinnveggen; optimalisere en ekstraksjonsprotokoll for oral prøver; sammenligne de analytiske egenskapene til to HPV deteksjon og genotyping, Luminex og NGS; og å få et innblikk i HPV prevalensen i studiepopulasjonen.

138 menn som har sex med menn (MSM) ble rekruttert fra Olafiaklinikken i Oslo. Det ble totalt analysert 136 børsteprøver, 135 munnskylleprøver og 98 urinprøver. HPV-prevalensen med Luminex var 2.94%, 4.45% og 17,34% for hhv børsteprøver, munnskylleprøver og urinprøver. For NGS var tilsvarende prevalensen 8.09%, 6.66% og 17.34%. Kun 21% (4/19) av pasientene med positiv oralprøve hadde samme genotype i begge orale prøvematerialer. Cohen's Kappa verdier for å måle overenstemmelse mellom de to deteksjonsmetodene var for børsteprøver, munnskylleprøver og urinprøver hhv 0.31, 0.60 og 0.40.

Antallet HPV positive prøver i dette studiet var for få til å trekke sterke konklusjoner. Ingen av prøvematerialene eller deteksjonsmetodene er tydelig bedre enn den andre. Derimot ser vi indikasjoner til at munnskylleprøver er et bedre egnet prøvemateriale, og at NGS er en mer sensitive deteksjonsmetode.

Table of Contents

Words and abbreviations

HPV	human papillomavirus
PV	papillomavirus
DNA	deoxyribonucleic acid
Early Genes	E1, E2, E3, E4, E5, E6, E7; HPV genes used early in infectious cycle
Late Genes	L1, L2; HPV genes used late in the infectious cycle
Genera	plural for genus, a category in taxonomy between family and species.
Amniotes	animals whom produce offspring with an amniotic sack
HR-HPV	high-risk HPV
LR-HPV	low-risk HPV
WHO	World Health Organisation
OPSCC	oropharyngeal squamous cell carcinoma
STI	sexually transmitted infection
MSM	men who have sex with men
Anogenital	anatomical region, consisting of genitalia anus
Urogenital	anatomical region, consisting of urinary and genital organs
Viremia	the presence of viruses in the bloodstream
HNC	head and neck cancer
OSCC	oral squamous cell carcinoma
Olafiaklinikken	sexual health clinic in Oslo, operated by Oslo University Hospital
Beta globin	a major component of the haemoglobin molecule.
MGP	modified general primers, modified from GP6+/GP7+ primers
MFI	mean fluorescent intensity
NGS	next generation sequencing
T-test	statistical analysis; assesses the statistical significant difference between two sets of values
AIN	anal intraepithelial neoplasia
CIN	cervical intraepithelial neoplasia
AHUS	Akershus University Hospital
RPM	rotations per minute

1 Introduction

1.1 Human papillomavirus

1.1.1 HPV and cancer

Cervical cancer has been theorised as a sexually transmittable infection (STI) since first postulated by the Italian physician Rigoni-Stern in 1842. He established a connection between deaths due to cervical cancer and sexual activity. He made note of the unequal presence of cervical cancer in women who were or had been sexually active such as married women, widows and prostitutes, against assumed virgins such as unmarried women and nuns[1].

The plurality of human papillomavirus (HPV) was not established until advances made in the 1970s[1]. Beginning in the early 70's, Harald zur Hausen explored the hypothesis of HPV causing cervical cancer[1]. His work was based on reports of malignant conversion of genital warts to squamous cell carcinoma, and the failed attempts of linking herpes simplex type 2 with cervical cancer[1]. The importance of his work was recognised with the Nobel prise in medicine in 2008[2]. Since he first began his work, a lot of research on HPV has been conducted, mainly concerning its role in the development of cervical cancers, but also in other locations of infection plausibly leading to cancers.

Today, HPV has been established as a culprit to several types of cancer: cervix, penis, vulva, vagina, anus, oropharynx, oral cavity, lip and tongue[3]. The most significant of these is cervical cancer[3].

1.1.2 The virus

HPVs is a group of viruses belonging to the wider group of papillomaviruses (PV) in the family *Papillomaviridae*[4,5]. HPV is small, non-enveloped, with circular double stranded DNA[4,6]. All types of HPV share at least six early genes (*E1, E2, E4, E5, E6* and *E7*) and two late genes (*L1* and *L2*)[7]. HPV is divided into five genera, Alpha, Beta, Gamma, Mu and Nu papillomavirus[8]. The classification of HPVs is based on the Open Reading Frame of the *L1* gene[4], encoding the major capsid protein. The different genera are distinguished by less than 60% sequence similarity, genus between 60% and 70% similarity, and a type has less than 90% sequence identity with any other HPV type[4]. As of May 2019 there are 226 HPV genotypes registered with the International Human Papillomaviruses (HPV) Reference

Centre[9], four of these 226 types are withdrawn due to re-classification (HPV 46, 55, 64 and 79), but left in the registry to avoid confusion[8]. Papillomaviruses causing mucosal cancers in humans belong to the alpha papillomaviruses[10], of these there are currently 65 genotypes[9].

1.1.3 Tissue specificity

The PVs are found to infect birds, reptiles, marsupials and mammals, most of the aminotes[5,6,11]. Furthermore, the lack of cross-species transfer and ubiquitous presence among the amniotes suggests that the virus existed at the origin of the amniotes themselves[5]. This long co-evolution with little cross-species transfer has made them very adapted to life in their host and to specific epithelial niches[6,11]. Viruses such as PV having evolved with their host in this manner seem to have found a balance between viral replication and immune tolerance. Thus, PV can usually complete their entire life cycle whilst maintaining a presence in the

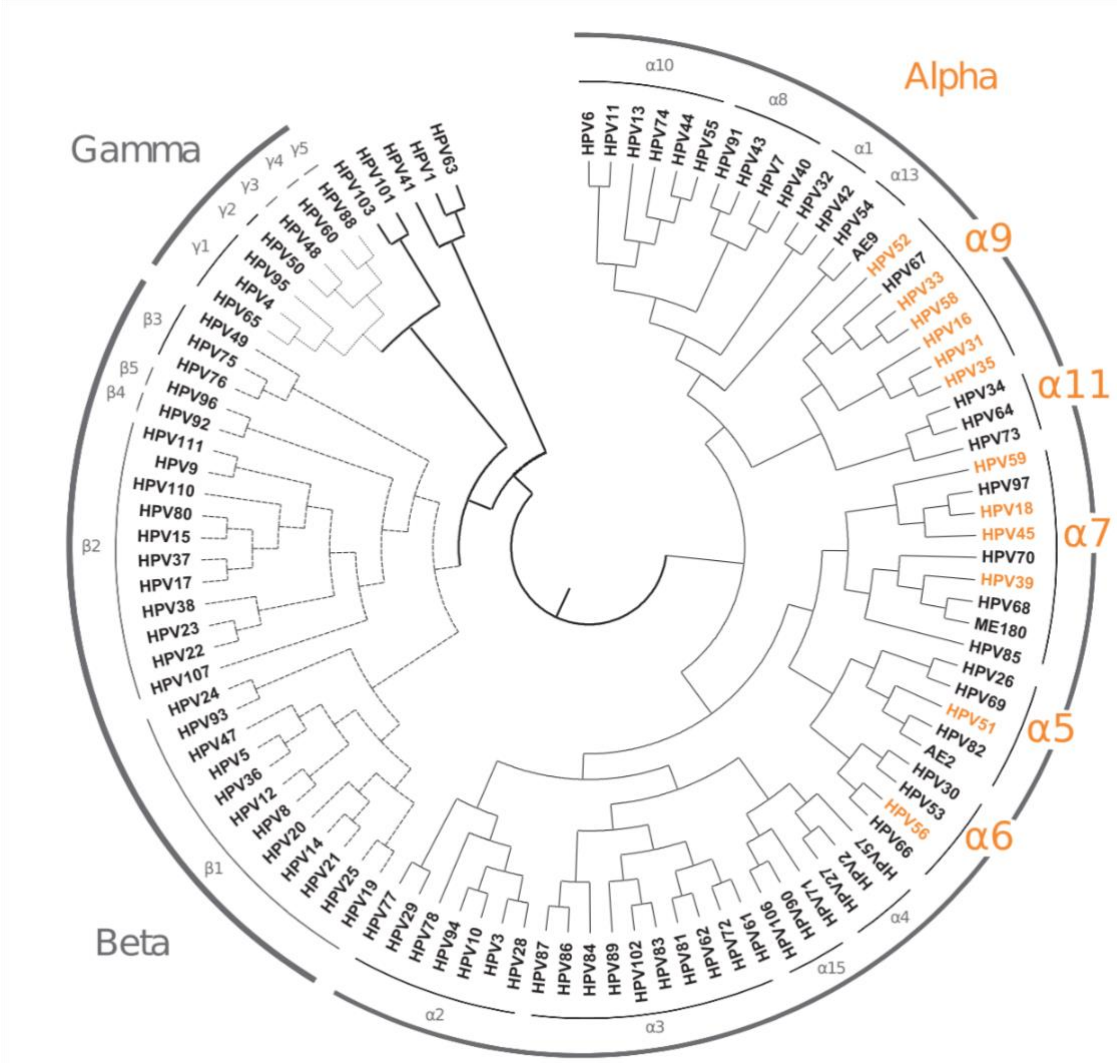


Figure 1: Phylogenetic tree of 100 HPV types. Highlighted are high-risk Alpha species as defined by WHO. Figure taken from IARC Monographs Human Pappilomaviruses[10].

population without causing apparent disease[5,11]. HPV in most cases cause chronic but asymptomatic infections[5,11], utilising sophisticated immune evasions or low level virion production to avoid immune detection in the infected host[6]. This unique adaptation to their host is a distinctive characteristic of the PVs[6]. HPV genotypes have a preference to distinct anatomical sites causing different pathologies in relation to genotype[6]. HPV viruses can typically be distinguished between cutaneous and mucosal, according to which type of tissue they infect[7]. Some HPV genotypes cause warts and some cause unapparent lesions that might progress to neoplasia and invasive malignant cancers[6]. HPV genotypes are classified into high risk HPV (HR-HPV) and low risk (LR-HPV), depending on their established relationship with malignant lesions and cancers. The World Health Organisation (WHO) classifies HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 59 as Class 1 Carcinogens for cervical cancer[10]. Often referred to as HR-HPV types, all of whom are clustered in the species groups, Alpha 5, 6, 7 and 9, belonging to the same clade of the Alphapapillomaviruses[12,13] (Figure 1). HPV manifesting as warts, especially at the anogenital region or the oral cavity, is often times the most unsettling and vexing to those afflicted[5]. The common causative HPV type for anogenital warts is HPV 6, and warts of the oral region is usually caused by HPV 11[6]. However, it is the group of HPV that may progress to neoplasia and cancer, that are of major medical importance, causing a wide range of epithelial cancers[1,5,11,14]. The genotypes causing cancers in the mucosal epithelium are all among the HR-HPV[5]. Of the HR-HPVs HPV 16 and 18 are known to cause most of the cancers of the cervix[15]. Of the HPV related oropharyngeal squamous cell carcinoma (OPSCC) it is believed that 70-90% are caused by HPV 16[14,16-19]. This gives HPV 16 a clear and important association to OPSCCs. HPV infections in the head and neck region are found to mainly develop in the oropharynx, the suggested theory for this is the nature of the tonsillar crypts which can act as a HPV reservoire[15]. The palatine and lingual tonsils are the most common site of oropharyngeal cancers caused by HPV infection[15].

1.1.4 Incidence of HPV at different anatomical sites

The global HPV prevalence in women with normal cervical cytology is 4% as provided by ICO/IARC Information Centre on HPV and Cancer[20]. However, there are regional variations; 12.1% in South America and 8.3% in Oceania, 4.5% in Northern Europe, 4.4% in North America and 3.4% in Asia[20].

Giuliano et al. found that when examining heterosexual men with a multitude of anogenital samples (Urethral, glans penis/coronal sulcus, penile shaft/prepuce, scrotal, perianal, anal canal, semen and urine) the HPV prevalence was 65.4%[21].
Urine samples have a HPV prevalence of 67% in women[22], and 29-37% in men[22,23]. Urine sampling for HPV detection in men is not clearly established as a suitable material. Giuliano et al. did not continue HPV analysis of 226 urine samples provided only 1 HPV positive, and so they ended the urine portion of their study[21]. In a review by Enerly et al., they observed a consistent lower prevalence of HPV in urine samples as opposed to other urogenital samples; coinciding with the low and inconsistent amount of detected beta globin they concluded that urine samples in men would not be optimal. Nonetheless, they reported the prevalence in studies reviewed of male urine as ranging between 5.8 – 36.7%, which was usually lower than other urogenital sample in their respective studies.

Prevalence of oral HPV in the normal healthy population ranges between, 4.0 - 11.5% in men[17,24-30], and 3.3 – 9.2% in women[17,24,27-30]. The prevalence varies from study to study, and even within studies: a Finnish study found the point prevalence of their women in a six year follow up to vary between 15 – 24%[31]. An Australian study examining different methods and sample material found the oral prevalence of: 10.4%, 11.5%, 3.1% and 16.7% in the same individuals[32]. Variation is also observed with distribution of genotypes: a study by Hearnden et al. found HPV 16 and 18 to be four of the total of 15 HR-HPV positive samples[33]. HPV 16 therefore constituted a small part of their identified HR-HPV. In contrast a study from Conway et. al found 12 of 17 HR-HPV infections to be HPV 16, making it the majority of HR-HPV's detected[28].

1.1.5 Men who have sex with men, HPV and incidence

Reports of oral HPV prevalence among healthy men who have sex with men (MSM) varies between 2-14%[26,34-36], and is higher in HIV-positive MSM 19%[36]. Prevalence of HPV in the anal canal is found between 30.8 - 53.3%[35,37]. Prevalence in the genitals is 9.5%[35]. King et al. reported 65% prevalence of any HPV type from anogenital samples (First void urine, intra anal swab and external genital swab)[34]. Interestingly according to Zou et al.[35] 10% of men who reported never receiving penetrative anal sex, were positive for HPV in their anal sample, possibly an indication of HPV infecting through other means than penetrative sex, e.g. contact between the genitals and anus. Furthermore, the risk of HPV 16 infection increases with increasing amounts of receptive anal sex partners[35].

A study of HPV in MSM at a sexual health clinic in London found that none of the participants with a valid HPV sample from anogenital site and oral rinse sample (n = 151) had the same HPV type detected at both sites[34]. Interestingly, this contradicts a Swedish study from a sexual health clinic in Stockholm looking at HPV in young females. They found that all participants having a valid oral and cervical samples (n = 22), showed the same genotypes detected from the oral sample in their cervical sample, but not necessarily the genotypes from cervical samples in their oral sample[27]. This could suggest a possible discrepancy between genders when it comes to concordance between oral and anogenital HPV infection.

1.2 Infection

1.2.1 HPV life cycle

Most research on HPV has been performed on the Alpha papillomaviruses, notably through the research of cervical cancers[11], but also anogenital- and oropharyngeal cancers[6]. However, as the same genotypes infect the aforementioned sites, which are all mucosal epithelium, we can expect that the mode of infection to be the same for cervical and oral HPV.

Most HPV genotypes are believed to require basal cells of the epithelium to successfully establish an infection, and is inextricably linked to the epithelial cell proliferation[6,11,38]. In the epithelium it is only the basal cells that undergo cell mitosis; superficial layers of cells simply undergo differentiation but no active DNA replication. For this reason to establish a successful infection epithelial trauma or micro-abrasions are necessary in many cases for the virus to gain access to the cells in active cell division (Figure 2)[5,6,11,39]. In addition, the viral genome needs to enter the cell nuclei of its host cell in order to utilise the replication

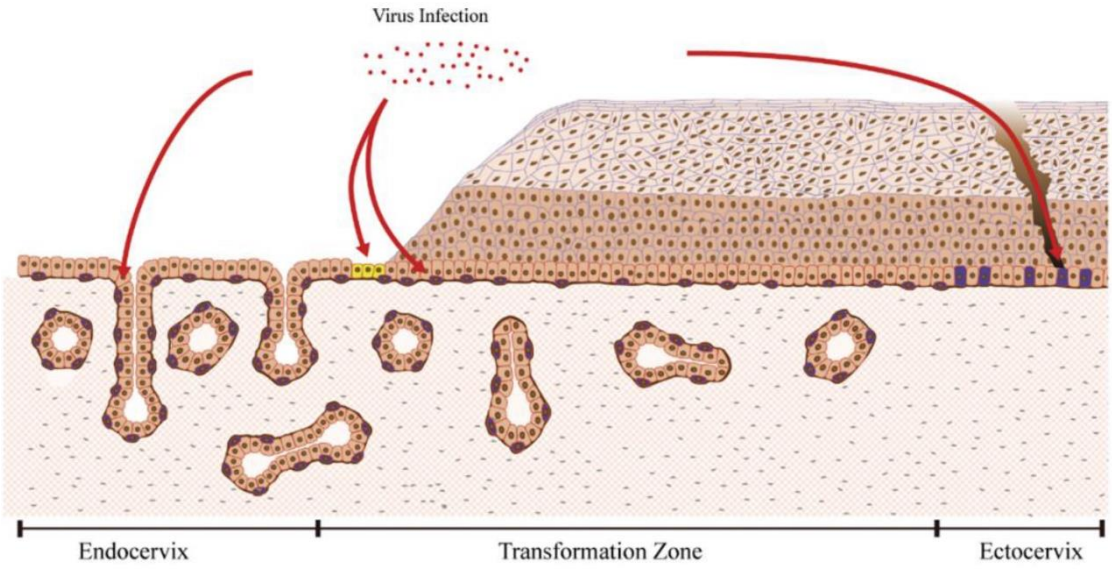


Figure 2: HPV infection of a. Hair follicle, b. between hair follicles and c. oropharynx in tonsillar tissue and entry into the basal cell layer. Figure taken from Egawa et al. 2015[6].

machinery. For HPV, access to the nucleus can only be achieved during cell division and HPV is therefore dependent on initial mitosis[11,38]. The cellular conditions during wound healing is thought to be critical in the establishment of LR-HPV infection. In addition, HPV seems to target vulnerable areas of the epithelial layers where access to basal cells is easier. Notably the transformation zone of the cervix, between columnar and stratified epithelium (Figure 2), as well as the anal transformation zone. But also; at the sites of specialized structures in epithelium such as: hair follicles, different sweat glands; and specialised epithelial structures such as salivary glands and tonsillar crypts (Figure 3)[5,6]. It is less clear if the requirements for a successful HR-HPV infection is dependent upon basal cells as they can force any cell into cell proliferation[11].

The life-cycle of HPV is tightly regulated by the virus itself. During the initial phase of infection, the E1 and E2 proteins are important for genome amplification and regulation of viral genome at a low copy number in the basal cells[6,11,39]. It is first when the daughter cells, from infected basal cells, differentiate and move through the epithelium that viral genome production reaches high levels and produce virions[39]. E2 remains an important regulator throughout later stages of the infection, notably to regulate the expression of viral oncogenes E6 and E7[6,11]. The E6/E7 proteins serve as important regulators for the progression of the cell cycle, and drive the S-phase re-entry in the upper epithelial layers[11]. The maturation of the virus is completed in the superficial layer where the conditions for producing stable infectious virions are met[11].

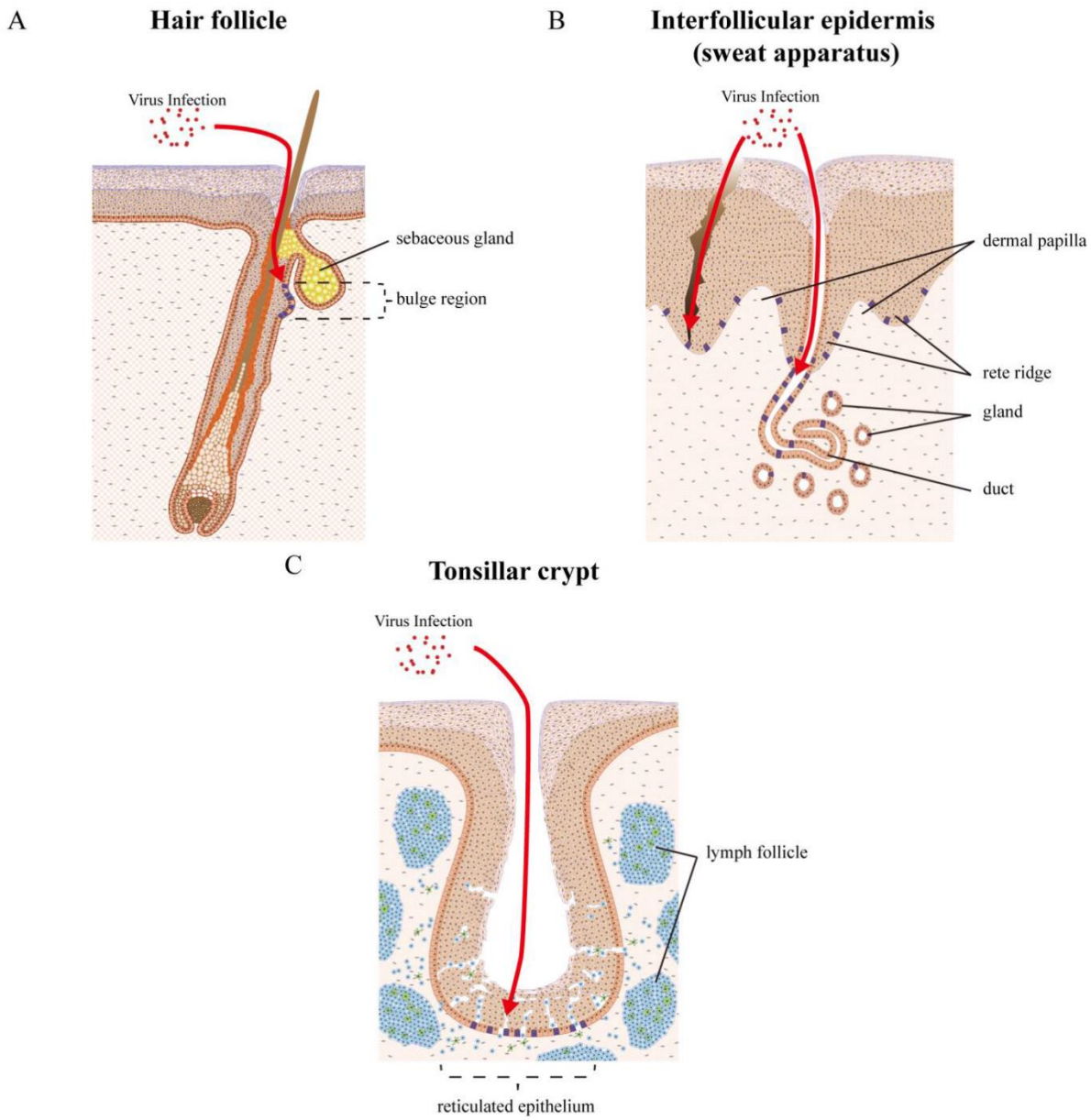


Figure 3: HPV virus infection of cervical epithelium and entry into the basal cell layer. Figure taken from Egawa et al. 2015[6].

1.2.2 Immune response and evasion

The immune system is slow in fighting HPV; this is mostly due to the viruses tissue-specificity to the basal epithelial cells and intra-epithelial nature[11]. The infection is not in direct contact with the circulatory system and causes no viremia[11,40]. In addition, HPV infection is not lytic or destructive, and is not believed to cause inflammation[6,11,40]. HPV travels along with the differentiating cells until their natural destruction and shedding from the epithelium, the virions are released away from immune cells[11]. These factors together aid HPV in preventing detection by the innate imunsystem[11], and allow an infection to become persistent.

In addition to a natural circumvention of the immune system, HPV employs several methods for active evasion[5,11,40]. In infected cells, HPV can compromise signals to Langerhans cells, dendritic cells and macrophages; and interfere with major histocompatibility complex affecting the host cells ability to display viral peptides, resulting in immune tolerance[5]. In spite of effective evasion strategies, nearly all HPV infections are still cleared by the host[5,6,11], however in some incidences the infection persists, a prerequisite for the development of cancers. This is interesting as HPV 16 persists on average longer than most other HR-HPV types, and this could be a key component in its overrepresentation among HPV related cancers[5].

1.2.3 Chromosomal integration

As previously stated nearly all infections with HPV eventually regress, with only a small part progressing to cancer[13]. It is theorised that integration of the viral genome into the host is a key event in cancer progression[40]. The sites of integration are not random, and are focused to fragile sites, transcriptionally active regions and chromosomal instability regions[41]. The understanding of integration and the place of integration within the host genome has thus become of interest[15,41]. Examining whether integration plays a similar role in the different forms of HPV cancer would also be an interest. Integration itself is not a natural process in the lifecycle of HPV. It is a dead end, preventing the production of virions as it is unable to make new circular genomes to be packaged and transmitted[39].

Integration can results in dysregulation and increased expression of the oncogenes E6 and E7[39,40]. These are known to inactivate or degrade many components within the cell related to cell cycle, particularly the s-phase, and repair off cellular DNA damage[40], resulting in increased cellular proliferation and cancelation of cell cycle checkpoints[39]. E2 is the viruses' own negative regulator of E6/E7 expression, in some observed integration events this gene has been disrupted, this leads E6/E7 expression to go unrepressed[38-40].

1.3 Cancer

1.3.1 Types of HPV attributable cancers in men

Most HPV related cancers of the head and neck region (HNC) originate in the oropharynx as OPSCC[14,43]. The time from initial infection to cancer in the oropharynx is hard to ascertain but is believed to between 10 and 30 years[16,44]. HPV is however tied to other HNC besides OPSCC; oral squamous cell carcinoma (OSCC) HPV positivity has been found to be between 22.2 – 61.5%[45-47].

As HPV is a STI, HPV related OPSCC and other HNCs are strongly tied to sexual behaviour. Oral sex, number of sexual partners, vaginal sex partners, rimming partners and age of sexual debut[16]. Sexual behaviour is changing, oral sex is more common, and with it the increase in HPV related OPSCC and other HNC. The so-called "sexual revolution" in the 1960's in the United States is believed to be a contributing factor to the elevated incidence of HPV related HNCs seen today[16]. Men who have sex with both men and women have highest risk of HPV infection[48].

HPV related OPSCC make out a distinct epidemiologically and clinical form of cancer, when compared to other OPSCCs. They are tied to sexual behaviour and show favourable survival as compared to non-HPV related OPSCC, which have strong ties to tobacco and alcohol use[19]. The HPV E6/E7 genes inactivate p53 in the host cell, which is similar to HNC associated with tobacco or alcohol, which are often linked with mutation of p53[18]. Despite the epidemiological distinction between the OPSCC cancers, the different origins are not unaffected by each other, smoking has been linked to significantly higher HPV incidence. It is believed to both increase the risk of infection and to increase risk of persistence[16,49]. HPV infection has been found as nearly three-times higher in current smokers[48]. There are several reported co-factors which increase the chance of acquiring HPV or for existing infection to become persistent or cancerous. Marijuana use have been associated with increased HPV 16 positivity, explained by cannabinoids immunomodulatory effect which may suppress tumour immunity as well as immune response to viral pathogenes[18].

1.3.2 Incidence of HPV attributable cancer in men

HNC is caused by mainly two attributes: HPV and tobacco/alcohol use[16]. These are competing causes for HNC, but not mutually exclusive. The overall trend in smoking habits has seen a decline in tobacco related HNC, however the incidence of HPV related HNC is rising at such a pace that the overall incidence for HNC is increasing worldwide[16]. In the United States, HPV related HNC incidence increased 225% from 1988-1990 to 2003-2004[16]. At the same time, HPV negative HNC decreased with 50%[16]. HPV went from causing 18% of the oropharyngeal cancers in the 1980s to 80% of oropharyngeal cancers in the 2000s (reviews by Rettig and D'Souza)[16]. With these levels of increase HPV related HNC is portrayed to exceed cervical cancer in the United States by 2020[16,19]. The trends in the United States, is in line with the world trend of increase in HNC mostly due to HPV related OPSCC[19]. In Europe HPV went from causing 35% of HNC before 2000 to 73% after 2005[16]. This increase in HPV prevalence can be attributed to changes in sexual behaviour and unsafe sexual conduct. Indeed, genital infection with herpes simplex virus 1 & 2 is also increasing in the recent birth cohorts in the United States[19]. This trend has however not, to our knowledge, been documented in Norway. According to statistics provided by the Cancer Registry of Norway; cases of cancers in base of tongue and pharynx (ICD-10 codes: C01-02, C09-14) were 229 cases in 2008 and 353 cases in 2017 (153 and 250 of the cases were among men)[50]. Meanwhile the cases of cancer in *cervix uteri* (C53) was 293 cases in 2008 vs 316 cases in 2017[50], which is a lesser increase than OPSCC. A Norwegian study published in 2017, examined all available tumour tissues samples from OPSCC (C01, C09-10) patients admitted to Oslo University Hospital from January 2010 to December 2011[14]. As this hospital is the sole provider of treatment for OPSCC in south-eastern Norway (population: 2.8 million, 60% of the total Norwegian population) this study offers a unique insight into the prevalence of HPV in OPSCC for Norwegian population. Of the 166 usable samples, HR-HPV was found in 127 (77%)[14].

D'Souza et al. [51] found that among the partners of patients with HPV related OPSCC there was no higher HPV positivity than what is reported for the general population at the corresponding age. This suggests that once cancer has developed, the HPV genome is integrated and that productive viral infection has been cleared. In addition, their partners have likely been repeatedly exposed to oncogenic HPV by their partners but managed to clear their infections, which is expected to be common[51].

On an interesting note, a meta-analysis by Machalek et al.[37] states that 20-30% of any MSM tested had high-grade AIN (anal intraepithelial neoplasia), and 83% of anal cancers contained HPV DNA. The prevalence of anal HPV and anal dysplasia among the MSM greatly exceeds that of the cervix[37]. Furthermore, incidences of anal HPV cancers in HIV positive MSM are the same as cervical cancer in women before cervical screening programs were put in place. However, the progression form AIN to anal cancer in HIV negative men (1 in 4000) and HIV positive MSM (1 in 600) per year, are both considerably rarer but less established than the progression from CIN (cervical intraepithelial neoplasia) to cervical cancer, that is (1 in 80) per year[37].

1.3.3 Survival of HNC

Fossum et al.[14] found in their study, that survival was lower in patients with HR-HPV other than HPV 16. However, their sample pool was small, HPV 16 was detected in 108 samples and other HR-HPV in 19[14]. If this finding is to be confirmed, it would present challenges for further disease prevention as vaccination against all HR-HPV is currently not feasible. A different study by Rettig et al found that persistent infection with HR-HPV other than HPV 16 did not affect survival whereas persistent HPV 16 did[16]. Both studies had similar sample pools, Rettig et al. (n = 124) to Fossum et al (127). However, Rettig et al. sampled patients after treatment, and thus might have found lower incidences to Fossum et al. having sampled biopsies of the tumours themselves[14,16].

HPV related HNC show a favourable prognosis, even in those who experience disease regression after treatment[16]. Fossum et al found a survival rate of 71% for HPV related OPSCC[14]. They also found that among the HPV related OPSCC the non-smokers had a higher survival rate than smokers, but not significantly higher[14].

HPV related OPSCC show an advantage to survival over other OPSCC[15], this is mainly due to better response of treatment[18]. For this reason, HPV testing of OPSCC can be advantageous in treatment and estimates of survival chances. However, the survival advantage is not seen in OSCC cancers with HPV as the cause[15]. There is a discrepancy in studies showing no change in survival chances versus reduced survivability in HPV positive HNC in non-oropharyngeal sites. In a study by Retting (n = 124) a disease-free survival rate was observed at 92% and overall survival was 98%, two years after diagnosis. 14 participants experienced recurrence of disease after treatment, of these 6 died, constituting all the deaths

in their sample pool[43]. In the united states 5-year survival from HNC is found to be 65%, with a 50% increase in chance of survival if the tumour is HPV positive[16]. Tobacco however interferes with the favourable survival related to HPV positive tumour[16].

Of the different HPV types, persistent infection with HPV 16 is the most common type found to cause recurrence of disease, which in turn highly affects survival rate[43]. Detection of HPV 16 in oral rinses after treatment could thus be used as an indicator for recurrence of disases[43]. Risk of reoccurrence with oropharyngeal cancer is greater with HPV negative tumour (35.1%) versus HPV positive tumour (13.6%)[16].

1.4 Vaccination

According to WHO as of 2018, 91 countries have implemented HPV vaccination in their national vaccination program, with an additional six planning to do so in the future[52]. In Norway, HPV vaccination has been a part of the Norwegian childhood immunization program since the schoolyear of 2009[53,54]. All girls born in 1997 and later have been offered vaccination against HPV in the seventh grade (11-12 years)[53,54]. The quadrivalent vaccine (with antigens against HPV 6, 11, 16 and 18) was used from the start of the vaccination program to 2017[54]. Following the fall of 2017, the offered vaccine was changed to the bivalent Cervarix (with antigens against HPV 16 and 18)[55]. All girls born 1991 and later were, as of November 2016, offered vaccination with Cervarix as part of a limited two year catch-up program.[55] As of the fall of 2018 boys in seventh grade have been offered vaccine against HPV[55]. No catch-up program has been offered to boys. However due to the switch from Gardisil to Cervarix 18 000 doses of Gardasil were in excess, sufficient for vaccinating 6000 people. Starting in the fall of 2018 it was decided to offer these vaccine doses to drug-users, sex workers and MSM born after 1991.[56]

However, when only vaccinating girls and relying on herd immunity for the protection of men in the population, the protection does not extend to the MSM population[37]. Boys in the seventh grade in Norway are now offered vaccination, and this will have a favourable impact from an MSM standpoint. Although this was a part of the listed benefits in vaccinating boys, it was not the principle argument for their inclusion in the vaccination program. When tasked with evaluating vaccination of boys, the Norwegian Institute of Public Health placed emphasis on the added reduction in risk of cancer in men; the additional coverage provided to

men who will not need to rely solely on herd immunity from vaccinated girls; and the awareness of HPV as a unisexual hazard allowing men to take responsibility for their own HPV status[57].

1.4.1 Vaccine surveillance

HPV infections are quite common, the probability of women and men with at least one sexual partner acquiring an HPV infection during their lifetime is: 84.6% for women and 91.3% for men[58]. In contrast to its commonality, the cancers it induces is not something most adults experience. HPV causes cancer in only a few rare instances compared to the amount of infections[12,13]. In addition, as cancer progression is slow taking up to a decade. These factors make any assessment of vaccine effectiveness a challenge. The decline in cervical cancer incidence is therefore not immediate after vaccine implementation; a key aspect in the vaccination of HPV is therefore a good vaccination surveillance program. It must measure the effect of the vaccine in the vaccinated population, and unvaccinated population to assess herd immunity; and the vaccines effect on genotype distribution. It is important to observe any possible changes in distribution of genotypes when vaccination is widespread. Will the vaccination cause a type replacement with other genotypes increasing in prevalence? Will vaccination cause an overall reduction in HPV infections, by protection against vaccine-types and possible cross-protection against other genotypes? To observe this possible change, it is essential to have a picture of the genotype distribution prior to widespread vaccination. It is important to do larger studies establishing a baseline of the genotype distribution and prevalence of HPV in the population, so that after widespread vaccination it is possible to compare the new distributions and prevalence.

Large-scale studies can assess the current distribution in the population, and how this change is in a vaccinated population. In the surveillance of the Norwegian HPV Immunization Program, a cross-sectional study of 17740 urine samples found that vaccine-HPV types were reduced by 77 percent from an unvaccinated birth-cohort to the vaccinated cohort, when measured in girls aged 17[54]. Feiring et al. concluded that the vaccination program in Norway had reduced vaccine-types by 90% in vaccinated girls, and 54% in unvaccinated girls, pointing at the heard effect as well as direct effect of large-scale HPV vaccination[54].

1.5 Importance of research into oropharyngeal squamous cell carcinoma in men

HPV has been firmly established as a cause for various epithelial cancers. Its importance in other body sites than the cervix is becoming clear, especially the increase in OPSCC among men. The increase in oral sexual behaviour is believed to be a driving force for the increase in OPSCC[44]. For cervical cancer, the time between infection and development of cancer has been estimated to at least 10 years. For OPSCC this has been estimated to somewhere between 10 and 30 years[16,44]. Unlike cervical cancer, there is no screening program for HPV related cancers at other body sites, affecting mostly men, and no system to successfully prevent cancer development if abnormalities were discovered[44]. All this considered, men are becoming an increasingly at-risk group that are dependent on future research and clinical intervention strategies. The incidence of HPV positive OPSCC is not expected to decrease before 2060, even with high vaccine coverage[44]. Indeed, projections indicate HPV positive OPSCC to surpass HPV positive cervical cancer in the United States by 2020[16,19]. This is due to both effective reduction in cervical cancers by effective screening programs and vaccination, and to the increase of OPSCC[19].

HPV is present in the oral cavity and oropharynx, and the global trend is an increasing rate of HPV positive OPSCC. More research is needed, and investigation into possible screening programs and better HPV detection methods for men is important. To conduct these necessary largescale studies, we need to know how to best detect HPV in oral samples, including sample material, DNA extraction, HPV detection and genotyping methods Establishing this is the first step in approaching larger studies of HPV infections in men.

2 Aim of this study

Considering the current state of HPV attributable OPSCCs, and the still substantial lack in knowledge and understanding, further studies are needed. It is of interest to the Norwegian national HPV reference laboratory to form the foundation for future study of oral HPV. A research cooperation have been made between the Norwegian national HPV reference laboratory, Olafiaklinikken, and Tromsø University. The cooperation with Tromsø University is on the HPV research as a part of the larger "Tromsøundersøkelsen" (Tromsø 7). They wish to increase knowledge and understanding of incidence and risk factors of oral HPV infection in the general population and in patients with HNCs. Increased knowledge of HPV will serve as the foundation to improve preventative treatment, and increase research of oral HPV and HNCs. Well established protocols and procedures are important for future studies. This study seek to explore different approaches to oral HPV research, and to lay the foundation for future study.

In order to increase HPV positive samples for a better comparison of the HPV detection and genotyping methods, urine samples from the same participants were also included. This also served as an insight to the possible concordance between HPV in oral and urine samples.

The primary aim of this study is to:

Establish a protocol for sample collection, DNA extraction, detection and genotyping of HPV in oral and urine samples in men.

In order to reach this aim, we will:

1. Compare two sample collection methods for the detection of HPV in the oral cavity/oropharynx (oral rinse and buccal brush).
2. Optimising DNA extraction for oral samples.
3. Compare the analytical performance of two HPV methods for the detection and genotyping of HPV in paired oral and urine samples.

Secondary aims which was also of interest:

1. To obtain a brief insight of HPV genotype distribution in a sexually active MSM population.
2. To examine the willingness to provide samples for research purposes, in an MSM population recruited through a sexual health clinic.

3 Material & Methods

3.1 Study population and samples collected

The population of this study are MSM who seek medical assistance at the sexual health clinic Olafiaklinikken in Oslo at their evening drop-in time. Olafiaklinikken is a government run medical centre administered under Oslo University Hospital, focusing on sexual health. On Tuesday evenings, Olafiaklinikken has a special drop-in offer for MSM who wish to be tested for STI. As this study seek to explore approaches to HPV analyses and is not designed to collect large scale epidemiological data, no accompanying information from the patients was recorded. The participants are in this study referred to as patients as they are recruited in a consultation and examination for STIs.

Of these patients, all were approached with the option to participate during their consultation, with no further inclusion criteria. Patients could participate only once; to avoid double enrolment all patients were asked if they had already participated.

The initial goal was to collect samples from at least 100 patients, with the possibility of including more. The number of HPV positives were low among the first 100 patients, the decision to keep recruiting for 3 extra Tuesdays was made. However, these extra patients did not provide urine samples as we had initially only ordered 100 sample tubes for urine sampling, and delivery time for more sample tubes was too long.

During the study period, 139 patients were recruited to the study. One patient had to be withdrawn as the waver had not been signed, and the samples provided were discarded; leaving us with 138 included. The average number of patients included per evening was 11.5 with the least being eight and the most being 16.

Of the 138 patients included, we received 136 buccal brushes, 138 oral rinse samples, and 98 urine samples. During DNA extraction, three of the oral rinse samples were not successfully extracted due to laboratory error, reducing the number of oral rinse samples to 135, leaving five of the patients without both buccal brush and oral rinse sample (Figure 4).

One urine sample never arrived at Ahus, with no recorded reason for its absence, or leftover urine sample tube at Olafiaklinikken. Patients would fill their urine samples themselves; most likely this was not done.

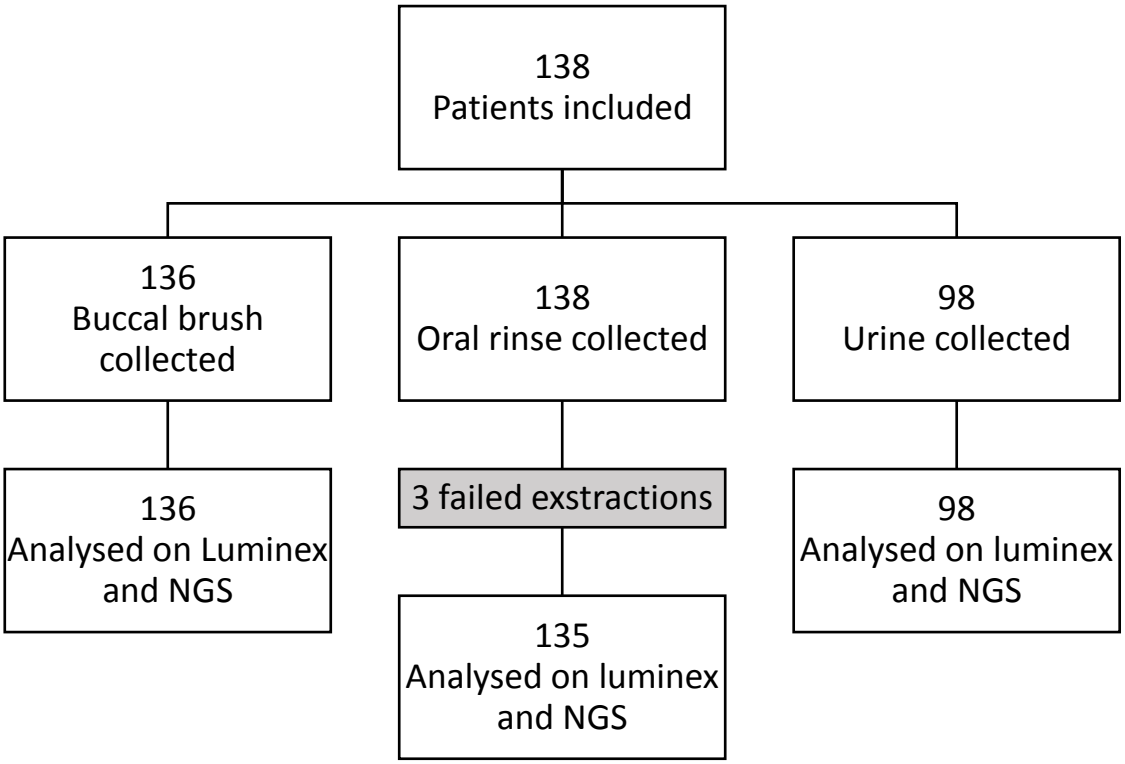


Figure 4: Flowchart of recruitment, samples collected and analysed for buccal brush, oral rinse and urine samples. Samples excluded and reason for exclusion is noted in shaded box

3.2 Ethics and information privacy

This study has been evaluated and approved by the regional committee for medical and health research ethics (REK sør-øst), reference 2018/1244.

Evaluation of information privacy has also been evaluated by the Ahus Data Protection Officer. There has been a Data Protection Impact Assessment (DPIA) performed for this study to protect the information and privacy of the patients.

Informed consent was provided from each patient, by signing an information letter which was archived as per instructions from REK.

3.3 Clinical material and collection

3.3.1 Recruitment and sampling

The sample procedure was designed in such a way that participation in the study would not significantly increase the time spent on each patient. Sampling should not be unnecessarily intrusive and not painful; the testing not add needlessly much work on the nurses who see the patients; and most importantly not interfere with the diagnostic tests of which the patients initially seek the clinic for. Excess time spent per patient would reduce the number of patients they could treat in an evening and possibly prevent someone from getting treatment. Tuesday evenings is a drop-in offer and if their capacity is exceeded, they have to close the doors. With this in mind, we arranged our samples to come after the diagnostic samples, with clear and precise instructions for the nurses and leeway for inclusion of participants; meaning that if a situation should arise where a patient was distraught, e.g. HIV-positive rapid test. Then the nurse would not collect a sample. Additionally, if challenges with collection arose, like spilling a considerable amount of the oral rinse solution, the nurse would not retake a sample but send the material received.

During the consultation, patients were asked if they would participate. If yes, the patient would sign a waiver; acknowledging that they understood the information provided by the waver and by the consulting nurse, and thereby consented to providing de-identified samples. No information from the patients were collected. Patients were permitted to withdraw – without providing reason, their sample from the study at any given time. All diagnostic tests were taken prior to study samples, as to not interfere with the diagnostic tests of significance to the patients. As such, the possibility of the diagnostic tests interfering with this study cannot be overruled.

Samples were labelled with de-identified numbers, with the link to the individual patients only available by the chief study nurse at Olafiaklinikken, in a locked cabinet. The only reason for de-identifying samples rather than anonymising is so the participant would be able to withdraw the sample after it had been collected, if they so wished.

3.3.2 Buccal brush collection protocol

Buccal brush samples were collected with: Cytobrush® Plus cell collector (Medscand® Medical). The Cytobrush was chosen based on pre-existing experience from the Norwegian National HPV Reference Laboratory, in which the alternative cotton swab had showed low DNA yield. Although not the most used approach to sampling oral HPV, it is for instance used when targeting visible malignant lesions in the oral cavity[47]. The sample was collected from the inside of both cheeks – the buccal lining, as described below. This was decided, as we did not want to reach down to the oropharynx. The routine samples at Olafiaklinikken already include swabs to the oropharynx for gonorrhoea testing; it was decided against adding further brushing to the oropharynx. This could be uncomfortable or painful for the participants, and possibly affect recruitment, as the patients would reap no benefits from these samples. Furthermore, we were curious as to whether buccal brush was successful in acquiring HPV and if the genotypes in the buccal lining differed from the oropharynx.
Two different brushes, one for each cheek, were used with the aim of providing sufficient material for DNA analysis. The nurse would rub the Cytobrush over the buccal lining inside the cheek with rotating movements up and down six times. To ensure that the brush collected enough cells, pressure would be applied, while taking care not to create sores or tears that would be uncomfortable. The brushes were suspended in a sample tube prefilled with 4 ml of ThinPrep® PreservCyt® (Hologic Inc.), a transport medium designed to preserve the epithelial cells collected in cervical samples. Excess shaft of the brush was clipped with plyers. Both brushes from the same participant were stored in the same tube.
The buccal brush samples were stored in refrigerator at Olafiaklinikken until the next day when samples were transported to Ahus and extracted on the day of arrival.

3.3.2 Oral rinse collection protocol

The patients were provided with a 50ml sample tube containing 10ml saline solution (sodium chloride (NaCl) 9 mg/ml). Oral rinse is a common sample method for examining oral HPV [17,24,26-29,32,34,51,59,60]. Saline was preferred over mouthwash; the latter however have been most common in similar studies. We however made the decisions as mouthwash might possibly interfere with the DNA extraction process. The saline solution was transferred into the sample tubes at Olafiaklinikken.

The 10 ml of saline were used for oral rinsing for 30 seconds, preferably as far back in the throat as possible. Patients were permitted to take brakes and instructed to try not to swallow any of the saline. The solution was then redeposited in the 50 ml tube and labelled by the consulting nurse. If some of the solution was swallowed or spilled, the test was not redone, but the remaining sample was sent to the laboratory, the nurse would record any such inconsistencies during sample collection.

Like the buccal swabs, the oral rinse samples were stored in refrigerator at Olafiaklinikken until the next day when samples were transported to Ahus and extracted on the day of arrival.

3.3.3 Urine collection protocol

In addition to the oral samples, urine samples were collected from the patients at Olafiaklinikken. Urine was collected by a beaker provided by the study nurse. This was filled with first flush urine in the clinic's restrooms after consultation. The clinic utilizes first flush urine as their primary analysis is the detection of bacteria and viruses, which is ideal for analysis of HPV. However, the sample time in this study was between 3PM and 8PM, and so none were morning urine. Morning urine may be richer in microbial and cellular material than late afternoon urine. The patient would fill a provided urine sample beaker and attach the lid. The sample tubes contain low pressure and when inserted through the lid are filled by suctioning urine. Patients were instructed to fill the sample tubes needed by Olafiaklinikken first, and then to fill the sample tube for use in this study. Once filled the samples were delivered to a deposit box outside the Olafiaklinikken laboratory door. Urine was collected in 10 ml sample tubes: C&S Boric Acid (BD Vacutainer®) REF 364955, containing boric acid to prevent microbial growth. As our sample was collected last the amount of urine could vary, but every sample was of enough quantity.

The samples were stored in refrigerator at Olafiaklinikken until the next day when samples were transported to Ahus where they were frozen at -20ºC.

3.4 Sample preparation

3.4.1 Buccal brush samples; processing

Upon arrival at the laboratory, tubes with the buccal brush samples (1 tube with two brushes per patient) were thoroughly vortexed for 20 seconds, three times, with 10 minutes break between each round. The break in vortexing is there to prevent foaming, and allow cells stuck to the brushes to soak and loosen for the next round. After vortexing, the brushes were removed with disposable plyers after which the tubes were centrifuged at for 15 minutes. The supernatant was discarded by inversion of the tube. The pellet would not be resuspended in the remaining liquid (approximately 200 µl). A slight but reasonable deviation from the intended process as sample prep and extraction could be done consecutively. DNA extraction followed immediately after centrifugation. The pellet was re-suspended directly in the NucliSENS® easyMAG® Lysis Buffer *(BioMérieux).*

The speed of centrifugation was changed two times during the study; initially 450 × g (n = 57), then changed to 1000 × g (n = 32) and later changed to 1500 × g (n = 47). The basis for altering speed was several observations of loose pellets after centrifugation. This led to a doubt on whether or not the centrifugation was sufficient, and the possibility of losing sample material. According to the procedure from the Tromsø University Hospital, our samples were spun at 450 × g for 15 minutes. When we considered altering this, we sought advice from Tromsø, who had been experienced the same uncertainties and recommended increase in centrifugation speed. The most commonly used strength in the articles reviewed was 3000 × g for either 15[48,61] or 10 minutes[26,60], while some go as high as 5000 × g for 10 minutes[35] or as low as 250 × g for 10 miniutes[33]. The centrifugation speed was increased from 450 × g to 1000 × g and later for additional safety increased to 1500 × g.

3.4.2 Oral rinse samples; processing

Upon arrival to Ahus the 50 ml tubes with oral rinse were centrifuged at 450 × g/1000 × g/1500 × g (n = 56, 32 and 47, respectively) for 15 minutes. The changes to centrifugation, as mentioned, was done to the oral rinse as well. The supernatant was discarded by inverting the tubes. The pellet was then re-suspended in lysis buffer and treated in the same manner as the buccal brush samples for the remainder of the DNA extraction.

3.4.3 Urine samples; processing

Upon arrival the tubes were frozen at -20ºC until further processing. Once they were to be extracted the samples were thawed overnight in a refrigerator unit in the laboratory. Once thawed the urine sample was thoroughly mixed, pipetting up and down ten times, or more if needed. 1 ml of sample material was transferred to easyMAG extraction cartridges, to which 1 ml of lysis buffer was added. These were thoroughly mixed.

3.5 DNA extraction and validation

3.5.1 NucliSens easyMAG

DNA extraction was performed with the automated extraction platform NucliSens® easyMAG® (BioMérieux), hereafter referred to as easyMAG. In brief: This method utilizes magnetic silica particles which binds to nucleic acids in the presence of chaotropic salts. Sample material were lysed with lysis solution of chaotropic salts and ethanol, denaturing proteins and releasing the genetic material from viruses and bacteria. The nucleic acids bind to the silica particles. Unbound material is removed through several washing steps. After completion of the washes, the nucleic acids are eluted into a buffer solution through the means of heat treatment, separating them from the silica particles. For list of reagents see Appendix 1.

3.5.1.1 Establishing extraction protocol

Prior to the extraction of our samples collected at Olafiaklinikken, we wanted to examine a few different approaches to DNA extraction and settle for one approach. For this purpose, samples collected from Tromsø University Hospital were used. These samples were pellets resuspended in remaining supernatant and stored at -80°C. For each approach 10 samples were used. DNA extraction was eluted in either 60 µl or 100 µl, performed with either all the sample volume or half the sample volume and extracted with either easyMAG or QIAamp (Tabel 1). To find the volume of a sample a pipette was used. The volume was defined as when the pipette contained the whole sample with no air in the tip, if this was not the case the sample would be redeposited to its tube and the pipette setting changed accordingly; this was repeated until volume was found. The samples compared with the two extraction methods used half of the sample material with one method, and half the sample volume with the other to allow a more direct comparison.

Table 1: The different extraction runs and the variation in the method between them, with the Tromsø samples to establish an extraction protocol.

Extraction run	**1th**	**2nd**	**3rd**	**4th**
Extraction method	easyMAG	easyMAG	easyMAG/QiaAmp	easyMAG
Volume of original sample	All	Half	Half/Half	All
Eluate volume	60µl	60µl	60µl	100µl
Number of samples used	10	10	10	10

The variables examined in the different approaches were sample volume and eluate volume, besides these variables the extraction with easyMAG were the same; adding lysis buffer manually, using 50µl of magnetic silica particles, magSIL and extracting on easyMAG.

QIAamp was compared to easyMAG, to assess the potential of a manual extraction method. Extraction was conducted as described in the protocol: DNA Purification from Blood or Body Fluids (Spin Protocol) a section of the QIAamp® DNA Mini and Blood Mini Handbook[62], which accompanied the QIAamp kit. Supplies and reagents were provided by corresponding kit: QIAamp DNA Blood Mini Kit (50), *Qiagen* (Catalogue number: 51104). In brief: ProteinaseK is added to the samples to lyse the cells and free the genetic material. The sample is then added to a spin column which when centrifuged, traps the nucleic acids in a filter. The samples were then washed through several steps of centrifugation with different wash-solutions. After washing, eluation buffer was added, releasing the nucleic acids from the filter. The final centrifugation yields nucleic acids in Buffer AE provided by QIAamp manual extraction kit.

3.5.1.2 Automated DNA extraction procedure for buccal brush, oral rinse and urine samples

The initial evaluation of extraction protocols performed in 3.4.1.1 was conducted to determine the desired extraction protocol. For the buccal brush and oral rinse samples it was decided to input all sample material and elute in 100 µl.

DNA extraction from the oral samples started with the resuspension of the pellets obtained after centrifugation. With 2 ml of the easyMAG lysis buffer, added straight to the original sample tube. The pellets were carefully re-suspended and thoroughly mixed. While mixing the sample, crumbs or other obstructions would be caught in the tip. These obstructions would then be carefully broken apart or moved to the wall of the sample tube. The removal of possible obstructions was essential, as they would wedge in the easyMAG pipettes during the automated extraction, resulting in a failed extraction. Any samples which failed at extraction would be attempted extracted again, starting with resuspending the magSIL in lysis buffer and initiating a new run with easyMAG. This was possible as easyMAG halts the progress of the samples were pipettes were blocked, the bonds between the nucleic acids and magSIL have not been broken. Highly viscous samples would be mixed for longer than other samples until the mucus was thoroughly dissolved. The re-suspended sample was transferred to the easyMAG cartridges, labelled with sample id. To the cartridge 50µl of magnetic silica particles magSIL was added followed by extraction on easyMAG

The urine samples were extracted as per the in-house approach. This approach is the same as for the oral samples with two exceptions: 1ml urine sample was added to the easyMAG cartridge with 1 ml lysis buffer; and the elution volume was 60 µl.

3.5.3 Nanodrop 2000 Spectrophotometer

Nucleic acid concentration was measured with Nanodrop 2000 Spectrophotometer (Thermo Scientific) which measures optical density of liquids in reference to the medium the sample is solved in (the Blank). The principle is based on spectrophotometry, and nucleic acids absorption of UV light at 260 nm. The absorption in the samples are used to calculate the nucleic acid concentration in the sample. From here on nucleic acid concentration is referred to a as Nanodrop values.

Nanodrop was preformed after the extraction, measuring every sample. After the instrument had been properly prepared and the blank calibrated with the easyMAG extraction buffer, 2 µl sample was deposited on the detection node and analysed.

Nanodrop was performed in the same manner for both the Tromsø samples during the establishing of the extraction protocol and the HPV study samples from Olafiaklinikken.

3.5.4 Beta globin PCR

The detection of the human gene HBB (haemoglobin beta) – hereafter referred to as beta globin – is an assurance of adequate DNA in the sample by measuring the presence of human DNA. Beta globin is a major subunit of haemoglobin, which is vital for oxygen transportation in the bloodstream. Real-Time PCR quantifies the amount of Human DNA in the sample by the detection of the HBB gene. It serves as a positive control for human DNA in a sample before conducting further analysis and serves as evidence of successful extraction. A sample with a positive beta globin result and a negative HPV detection result ensures that the result is not a false HPV negative due to inadequate material.

The Real-Time PCR method utilized in this study, and in routine analysis at the Norwegian National HPV Reference Laboratory uses Taqman Universal Mastermix, *Applied Biosystems* (Catalog number: 4364340). Analysis was done on the Real-Time PCR instrument: Stratagene Mx3005+ QPCR System (Agilent Technologies). For list of reagents, see Appendix 2.

The master mix was prepared in a template-free-laboratory to prevent contamination. Mastermix consisted of, per sample: 12.5 µl TaqMan Universal Mastermix and 7.5 µl primer/probemix. Mastermix was prepared in excess to compensate for inaccuracies using a multichannel pipette via a tray. The master mix was aliquoted to a 96'well PCR plate, dispensing 20 µl in each well. In the first two columns 5 µl of standards and controls (positive and negative) were added in parallels. In the remaining wells 5 µl of sample was added. Total volume in each well was 25 µl. The standards were Human Genomic DNA, a 10-fold dilution series with 6 tubes starting from 20'000 copies/µl down to 0,2 copies µl. The standard curve is used to translate the signal strength of the samples to an amount of copies – the copy number. The plate was vortexed and spun before it was placed in the Real-Time PCR machine. PCR cycles were as follows: 2 min at 50ºc, 10 min at 95ºc; then 50 cycles of: 15 seconds at 95ºc, 1 minute at 60ºc

3.6 HPV detection and genotyping

Two methods for the detection and genotyping of HPV were used in this study. Genotyping with Luminex that is based on DNA hybridization, and NGS that is based on DNA sequencing. The methods are described in further detail below. Both methods include PCR using the HPV MGP primers (modified GP5+/6+ general primers). These primers have been developed by Söderlund-Strand et al.[63]. The GP5+/6+ primers target the L1 gene of the HPV genome; and the MGP primers have been specifically designed for increased sensitivity to HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66 and 68. The basis for amplicon generation is thus the same for both methods, but the PCR cycles were not the same; direct comparison is therefore not appropriate. The PCR cycles for Luminex and NGS are shown in figure 5 and 6 bellow.

95 ºC, 10 min
5 cycles:
    95 ºC 30 sec
    42 ºC 30 sec
    72 ºC 45 sec
45 cycles:
    95 ºC 30 sec
    64 ºC 30 sec
    72 ºC 45 sec
4 ºC, ∞

Figure 5. PCR cycles for MGP amplification for Luminex detection and genotyping.

98 ºC, 30 sec
5 cycles:
    98 ºC 10 sec
    42 ºC 30 sec
    72 ºC 15 sec
45 cycles:
    98 ºC 10 sec
    64 ºC 30 sec
    72 ºC 15 sec
72 ºC, 10 min
4 ºC, ∞

Figure 6. PCR cycles for MGP amplification for NGS detection and genotyping

3.6.1 Luminex

Detection and genotyping of HPV was done on a Luminex® 100/200™ system (Luminex Corporation). The process of detection and genotyping with this approach is in this study referred to simply as Luminex. The method is in-house and based on hybridization of genotype-specific oligonucleotide probes coupled to fluorescence labelled polystyrene beads for detection with the Luminex suspension array technology as developed by Schmitt et al.[64]. Our in-house Luminex methods contains probes for detection and genotyping of 37 HPV types (HPV 6, 11, 16, 18, 26, 30, 31, 32, 33, 35, 39, 40, 42, 43, 45, 51, 52, 53, 54, 56, 58, 59, 61, 66, 67, 68, 69, 70, 73, 74, 81, 82, 83, 86, 87, 89, 90 and 91), in addition to universal probes made to detect other variants.

Luminex principle in brief: HPV amplicons generated with MGP is hybridised with the genotype-specific probes coupled to small magnetic beads. These beads, which are detected by a red laser, contain different concentrations of two fluorophores, opening the possibility for the differentiating of up to 100 different hues. A second, green, laser detects whether the beads in question is hybridised with biotin-streptavidin-R-phycoerythrin-labelled PCR-product. This compound absorbs the green-blue light from the green laser, returning only yellow-orange. The Mean Fluorescence Intensity (MFI) from the R-phycoerythrin-labelled PCR-product determines the presence of HPV DNA. The MFI can vary depending on the quality of the specific bead product binding; as such this method is not suited for direct quantification (Figure 7).

One run with a 96-well plate for Luminex uses 80 patient samples, the other wells are used for; positive controls: HPV 16, 18, 31 and 33 as well as MGP-1 (HPV 6, 16, 18 and 51), MGP-2 (HPV 11, 31, 33 and 58), MGP-3 (HPV 39, 45, 52 and 56) and MGP-4 (HPV 35, 35-6624A, 59, 66 and 68orth); Water was used as a negative control. As the quality of the mastermix deteriorates within a few hours it is essential to work efficiently until the samples are placed in the PCR instrument. To each well 20 µl master mix was dispensed followed by 5 µl of sample material or controls. Total reaction volume in each well was 25 µl. PCR was performed with the cycles shown in figure 5. After amplification, the samples were either frozen at - 20ºC until such time as the process could be continued or left in the PCR instrument at 4 ºC overnight and then frozen or analysed.

As proper training for Luminex genotyping is time-consuming, in addition to the present study's limited sample number which would not be enough for a completing training, it was decided that all Luminex analysis, after PCR amplification, should be undertaken by the laboratory personnel at the HPV reference laboratory.

Performed by others were: Hybridization of the amplified product to the Luminex beads, by using, per sample: 34.1 µl hybridisation buffer and 0.1375 µl of each genotype-specific bead type. The nature of the bead's fluorescence capabilities makes them highly photosensitive, and so the work must be conducted in low lighting. Hybridization takes place at 95 ºC, then cooled in an ice bath. After hybridisation the products are washed. To each well 70 µl of biotin-streptavidin-R-phycoerythrin is added and incubated, followed by a second wash to remove unbound material. After hybridisation and successive washing, the samples were ready for HPV detection through Luminex analysis.

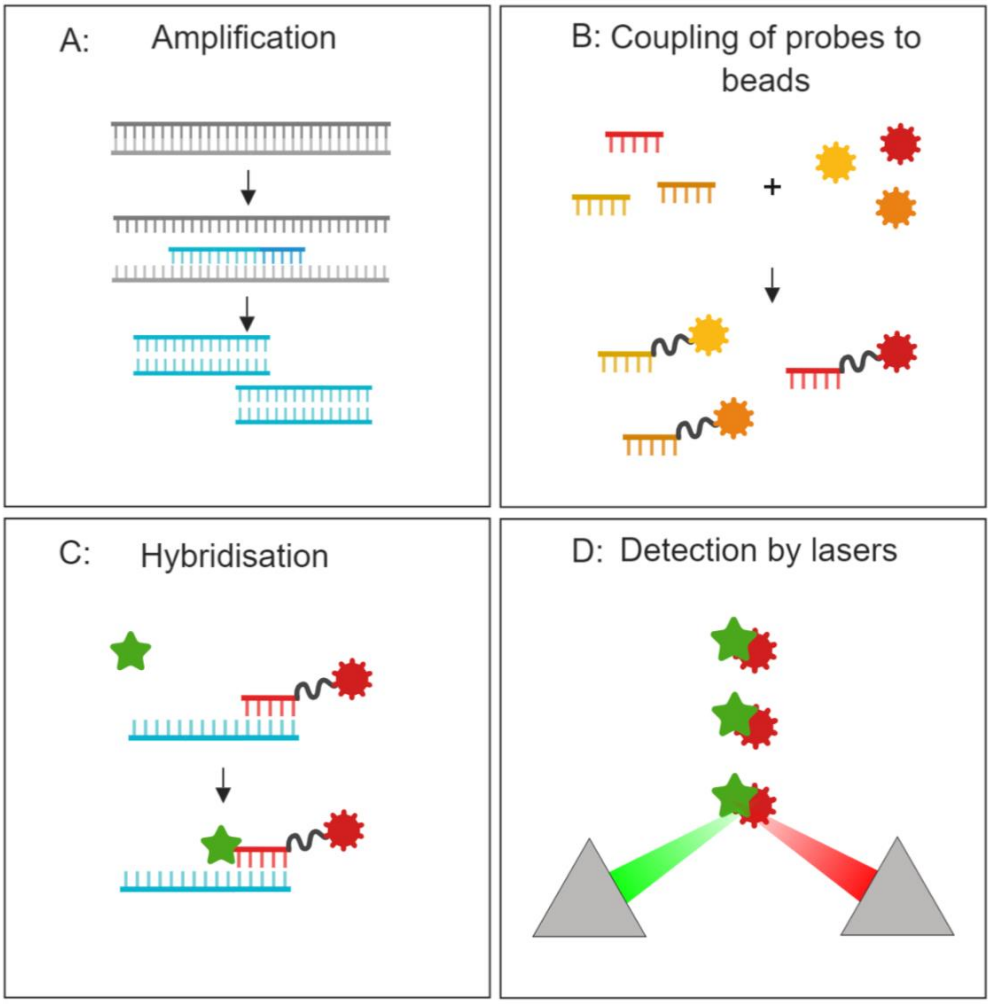


Figure 7: Detection and genotyping with Luminex. A: amplicon generation through PCR with MGP primers. B: Beads of different hues are attached to genotype specific probes. C: a probe is matched and attach to the sequence in our amplicon allowing biotin-streptavidin-R-phycoerythrin to attach to the probe. D: red laser detects the hue of the beads; green laser detects the streptavidin coupled PCR- product.

It has been recommended by the supplier to avoid high DNA concentrations due to potential sequestration of magnesium ions that are required for the polymerase to function. Our first run with Luminex gave no positives, which we did not expect; to exclude the possibility of false negativity due to too high DNA concentrations a re-run of these samples was performed. For the re-run, aliquots of the samples diluted to 10 µl were used. The samples were however, confirmed negative also after dilution.

For list of reagents, supplies and equipment see Appendix 3 - 5.

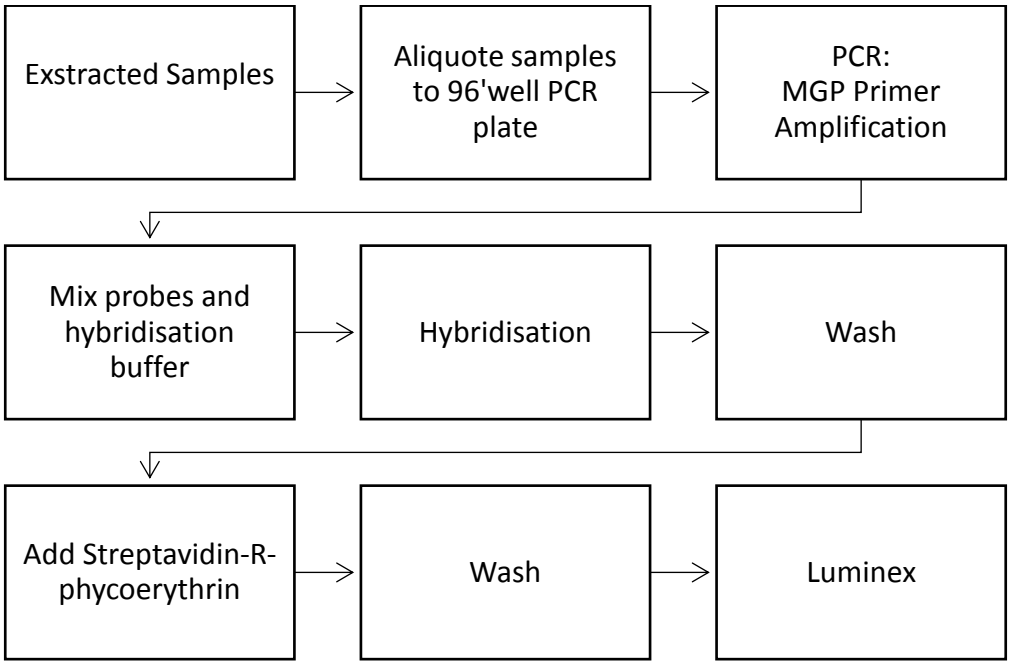


Figure 8. Flowchart depicting the flow of samples from finished extracted product until Luminex detection.

3.6.2 NGS – Next generation sequencing

3.6.2.1 Pre-Sequencing

HPV was detected and genotyped with NGS using MGP primers, and sequencing on the Illumina® Miseq System. In brief: Amplicons of our sample material are produced with MGP primers through PCR. The products are indexed and pooled in a library of all samples to be sequenced. The library is deposited to a flow cell covered in oligonucleotides complementary to the adaptor at the end of our indexed strands; our products attach to the flow cell. Through the means of bridge amplification our products are amplified to larger clonal clusters, this allow for better detection than single strands. The strands are then sequenced by synthesis, polymerases build the parallel strand attaching modified nucleotides which emits lighting when synthesised. The different nucleotide bases emit different coloured

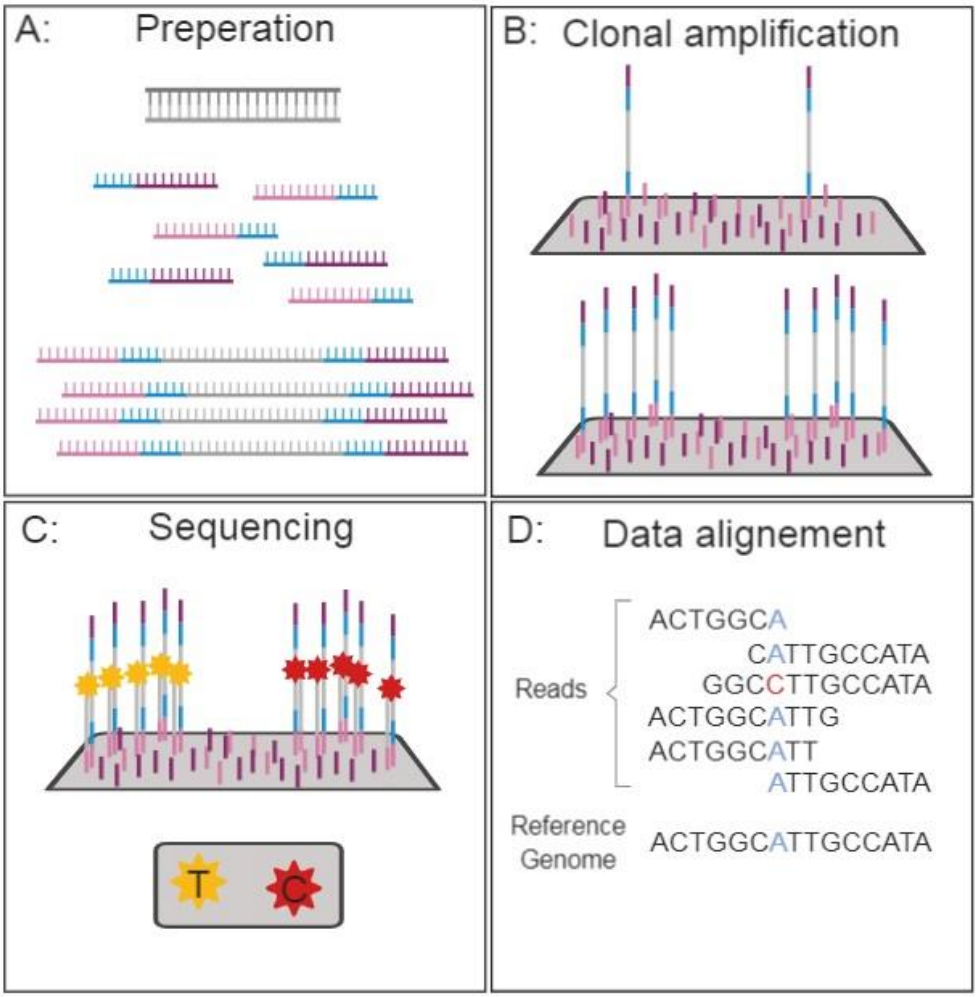


Figure 9: A: target gene is amplified, and index and adapters are attached through PCR. B: sequences attach to oligonucleotides in the lanes of the flow cell, oligonucleotides match to the adapters attached to our sequences. Clonal amplification through bridge amplification generates large clusters for improved detection. C: sequencing by synthesis, nucleotides attached fluoresces with a colour specific to the base of the nucleotide. D: sequenced data is aligned against a reference genome to assign origin. Figure adapted from documents by Illumina[65].

light. The nucleotides are also reversibly terminator-based, which prevents the polymerase from adding more than one nucleotide. When the signal is detected, the last nucleotide added is modified to a state allowing a successive nucleotide to be added. The lights emitted during synthesis are photographed and translated by the system into a sequence of nucleotides. Once processed this sequence can be mapped to reference genomes to assign origin, and by reading the index we can trace the sequence to its sample[65] (Figure 9).

The process from extracted DNA to HPV detection and genotyping with this approach, as described in further detail below, is in this study referred to as NGS for simplicity.

Four 96'well plates with 5 µl from each of the collected 369 samples were dispensed in their respective wells, as well as two HPV positive cell line controls: HeLa (HPV 18) (N40062, New England Biolabs) and SiHa (HPV 16) (ATCC® HTB-35™). Once this step was complete, laboratory personnel currently finalising the protocols for this method, preformed the successive steps.

Performed by others: Mastermix with MGP primers were dispensed to each well and amplified with PCR reaction on an Applied Biosystems™ ABI5700 PCR instrument. Amplified product can be stored at 4 ºC for 2 – 4 weeks. As all samples in this study were analysed in the same run, their treatment was the same.

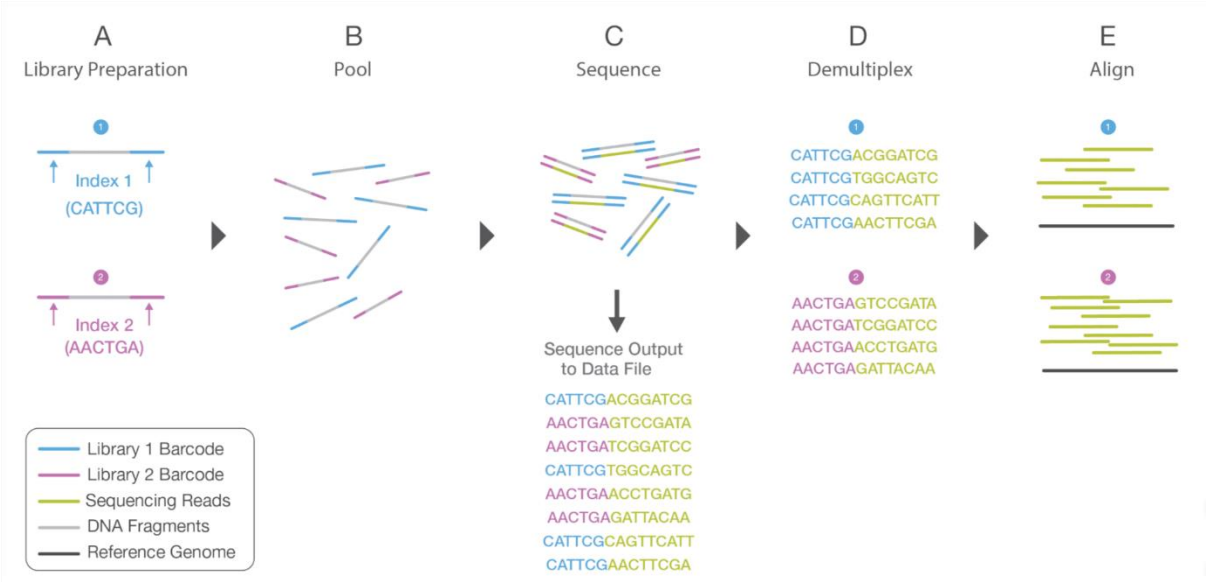


Figure 10: Library preparation and multiplexing through indexation. A. Different samples are indexed with unique indexes. B these are pooled together into a single library. C. Indexes are sequenced along with the sample DNA. D. Based on the sequenced indexes the reads can be sorted based on their origin. E. Reads from a the same sample can be aligned to a reference genome. Figure adapted from documents by Illumina[65].

To assess quality after amplification, a random selection of eight samples, two from each PCR plate with HeLa and SiHa as control were run on E-gel™ Invitrogen™ after MGP amplification. E-gel was preformed again after index PCR, on 10 different samples. To assess fragment the sample library was run with Agilent 2100 Bioanalyser.

After amplification, samples require indexing in order to assign sequencing results to individual samples. NGS with Miseq Illumina is performed on a pooled library of all our samples. A unique oligonucleotide, an index, is attached to the ends of the amplified PCR product. To achieve a maximum number of unique labels with the minimal number of specifically tailored indexes, we use a combination of two index primers in each sample. In a 96'well PCR plate each well in the same row receives 1.5µl of the same index (SX501-SX508, 5 µM), each well in the same column receives 1.5µl of the same index (SX701-SX712, 5 µM). The indexes are hybridised to amplified product by means of PCR, the indexes are attached at the end of primers specific for the products achieved with the MGP primers. Indexing also serves to attach the oligonucleotide sequence matching the lanes on the sequencing plate in the Miseq Illumina instrument, which are needed for the attachment and bridge amplification. After indexing the samples are pooled to a single library in a 1,5 ml sample tube before it is sequenced with the Illumina® Miseq System. Appendix 8.

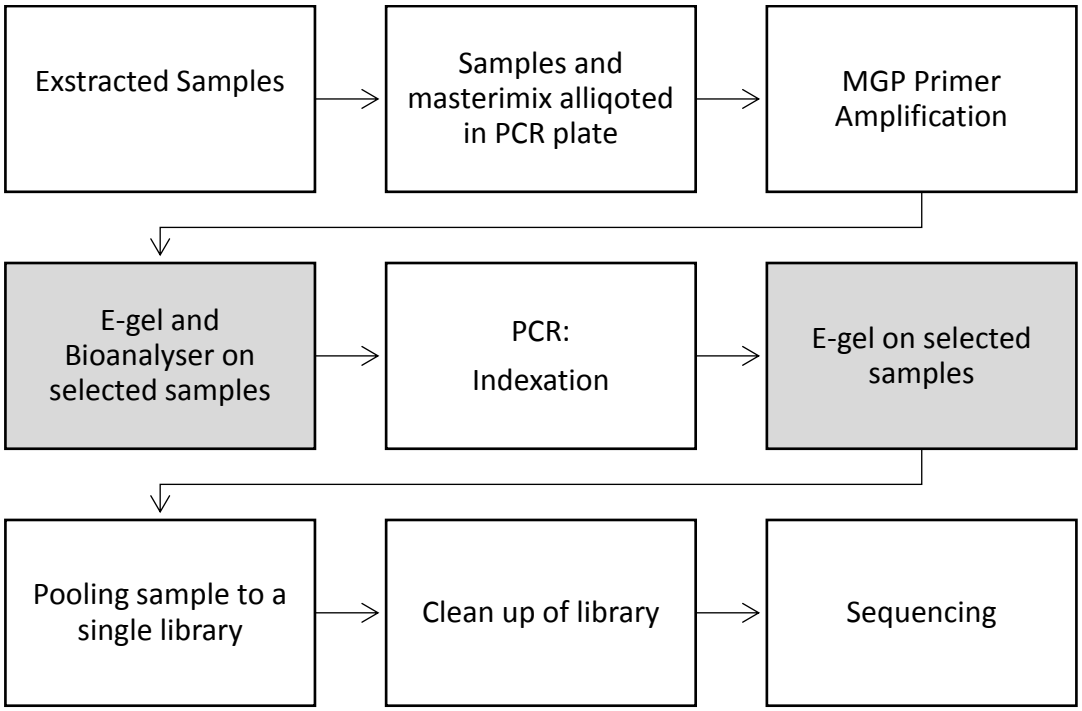


Figure 11. Flowchart depicting the flow of samples from finished extracted product until NGS detection. Steps shaded in grey were conducted with only a few arbitrary samples.

3.6.2.2 Post-sequencing

The treatment of the output data form the Illumina sequencing, were performed by others at FoU, MIKS, Ahus and at Cancer Registry of Norway.

The file format of the raw data from sequencing were FastQ, providing the sequence with accompanying quality score. Each sample provides forward and reverse reads in separate FastQ files. As 369 samples were sequenced, this amounts to 738 files, each of which must be evaluated for quality. Quality evaluation was performed by the creation of a fastQC box plot which translates the quality score of the FastQ files to an interpretable format. Since reviewing each plot is inefficient, a second program multiQC is used. MultiQC gathers all the fastQC plots and creates one single plot where a general overview is given, allowing for a closer review of the quality and assessment of samples with quality that stands out.
Based on the quality control, the sequenced strands are trimmed. It is common for sequenced strands to have poor quality at the start and end of the read, as well as the possibility of poorer quality elsewhere on the strand. A program, Nesoni.clip, cuts out poorly called nucleotide bases below a certain quality, as well as completely removing short sequences with less than 50 base pairs remaining after trimming.
Once the FastQ files have been trimmed a program, bowtie2, maps the sequenced strands to a

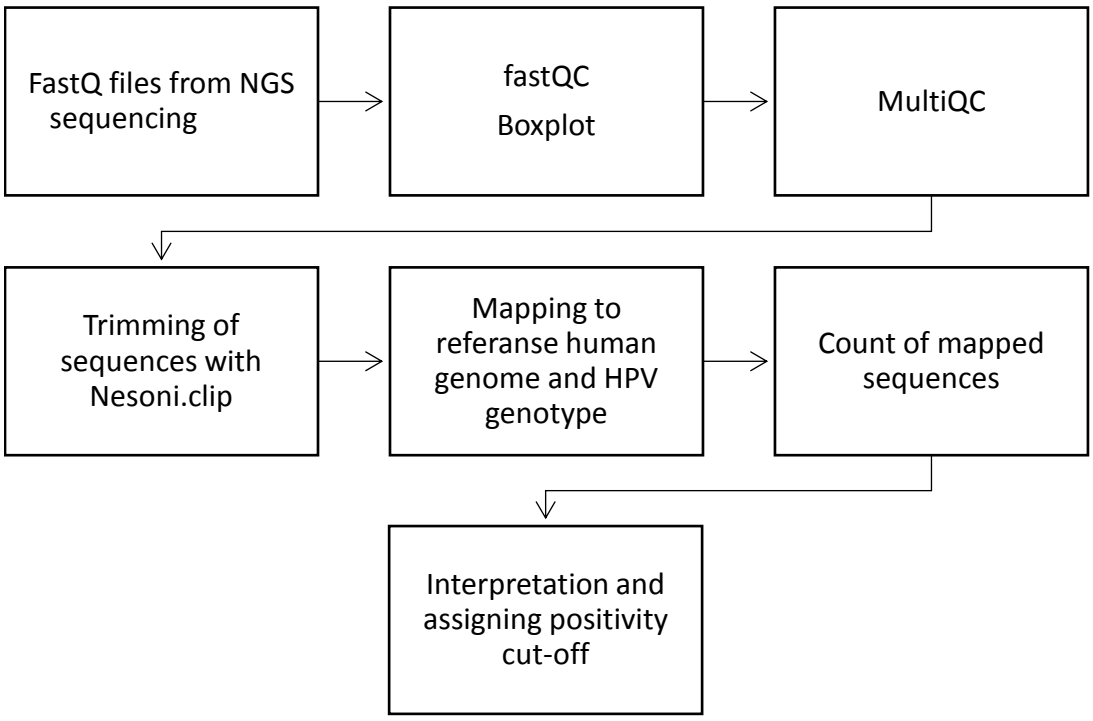


Figure 12. Flowchart depicting the flow of samples from sequence data to genotyping results.

given reference. In our case a reference human genome (hg19/GRCh37) and 183 genomes from reference HPV genotypes found in PaVE[66]. The files for the forward and reverse sequence are individually mapped, but are removed if they map to different areas, for instance two different HPV genotypes. After mapping the sequences are stored in a new file format a BAM/SAM file, which includes description of the mapping. The mappings are then scored and assigned a CIGAR string accordingly, that will be used to assess the mapping. The fourth step in this pipeline is counting mapped sequences to each HPV type creating a table of counts per sample per HPV genotype.
This whole pipeline is automated through a Bash script. The only manual input is the review of the MultiQC and any FastQC of interest.

After the bioinformatics analysis, excel sheets were sent back to Ahus. These contained counts of sequences for each genotype per sample. The final step is a review of these findings as well as establishing a cut-off, how many counts of a mapped sequence is necessary to confirm a positive. As this study seek to explore different approaches and serve no clinical importance, cut-off is less consequential. The NGS protocols used in this study are not finalised, and so contained no accompanying recommendations. Cut-off was assigned after a review of the findings. Cut-off was assigned at 50 reads, samples with less counts were thus regarded as negative.

3.7 Statistics

All statistical calculations were performed in Microsoft Excel 2013. Alpha level for statistical significance was set at $\alpha = 0.05$, as such p-values were considered of statistical significance when $p < 0.05$.
Student's t-test was used to determine statistical significance on Nanodrop and beta globin values of the samples used in the establishment of extraction protocol and on the buccal brush, oral rinse and urine samples. The T.TEST function in Excel was used to calculate the p-values with t-test. Three variations of the t-test were needed; t-test on values with equal variance, unequal variance and paired data. To determine if the data were of equal or unequal variance, f-test was used with the F:TEST function in Excel.

To compare the analytical performance between Luminex and NGS, Cohen's Kappa values were calculated. Cohen's Kappa was also calculated for the comparison between buccal brush and oral rinse in their ability to make the same HPV genotype identifications.

4 Results

4.1 Patient populations

The progression of the study was well ordered without logistical or practical difficulty; the collection of material and sample analysis was not hindered nor delayed by unexpected challenges. Study participation was satisfactory and patients at Olafiaklinikken were very willing to provide sample material. The nurses who conducted the sampling reported no distrust or scepticism from their patients. The collaboration on this project between Ahus and Olafiaklinikken has laid a positive foundation for future cooperation.

4.2 DNA extraction

4.2.1 Establishing extraction protocol

To establish an extraction protocol, four extractions with alterations were preformed and compared to each other, (Table 1). The results of nucleic acid concentration measured on Nanodrop, and beta globin gene copies and the 95% Confidence Interval (CI) are presented in Table 2. The difference between the extractions based on Nanodrop and beta globin are illustrated with boxplots (Figure 13 & 14).

Following the approach described in 3.4.1.1 and assuming normal distribution, two tailed t-test was performed; with $\alpha = 0.05$.

There was an equal variance between extraction of full sample versus half sample volume, determined by f-test, p-value 0.82 and 0.53 for Nanodrop and beta globin respectively. Therefore, a t-test with equal variance was used, finding no statistical significance, p-value: 0.94 and 0.87 for Nanodrop and beta globin, respectively.

When comparing QIAamp and easyMAG the same sample was split in half, extracted with each method. The t-test of paired samples was conducted. According to Nanodrop values there is no statistical significance, p-value: 0.36; however, the beta globin values found statistical significance, p-value: 0.02.

Lastly comparing the different elution volumes, 60 µl versus 100 µl. Conducting f-test found their variance to be unequal, p-value 0.02 and 0.02 for Nanodrop and beta globin, respectively. Conducting a t-test of unequal variance found no statistical significance, p-value 0.37 and 0.13 for Nanodrop and beta globin, respectively.

Table: 2: Average, mean and 95% confidence interval (CI), of the 4 different extractions with their distinct feature (1 with all sample volume, 2 with half sample volume, 3 samples split in half and extracted with both EasyMag and QIAamp, and 4 extracted with an eluate of 100 µl). All samples for run 1, 2 and 3 were eluated in 60 µl.

Extraction run with alteration		Nanodropp (ng/µl)			Beta globin (Copies)		
		Mean	Average	95% CI	Mean	Average	95% CI
1	All	50.7	72.5	± 33.3	11800	12900	± 5600
2	Half	65.6	70.7	± 32.2	7490	12200	± 7490
3	EasyMag	55.1	49.8	± 10.6	6940	7340	± 2750
	QIAamp	51.8	48.5	± 11.4	8490	9900	± 4210
4	100µl	55.5	55.3	± 14.1	8250	7840	± 2430

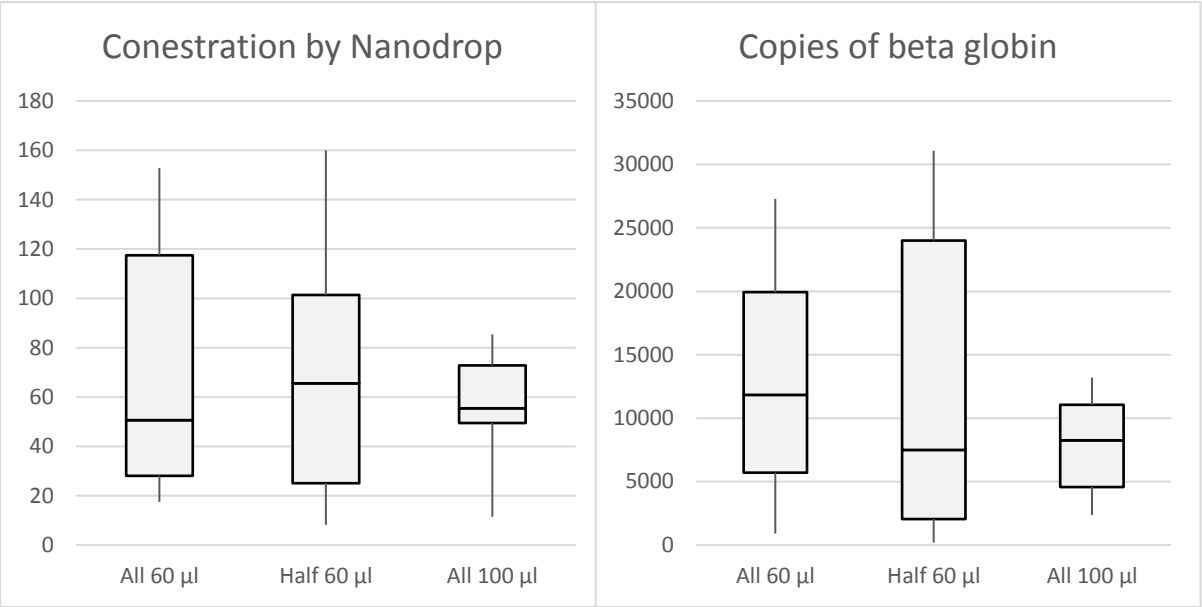


Figure 13: Boxplot of DNA concentrations as measured in nucleic acid concentration with Nanodrop and copies of beta globin gene with Real-Rime PCR. In the three variations of extraction on easyMAG: All of sample material with 60 µl eluate, half of sample material with 60 µl eluate and all sample material with 100 µl eluate.

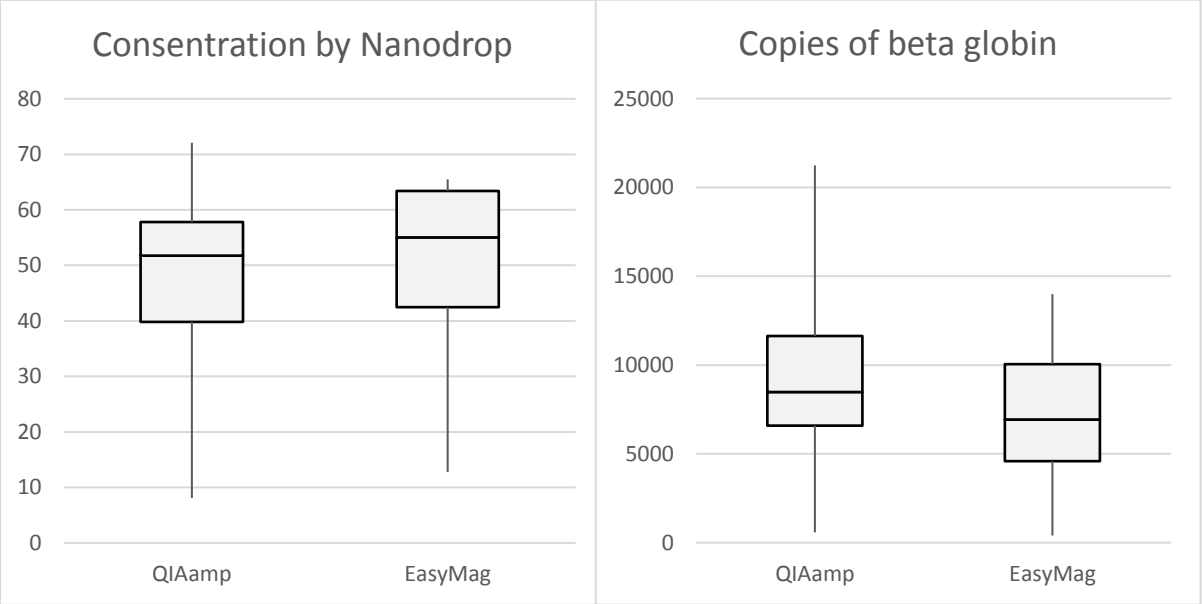


Figure 14: Boxplot of DNA concentrations as measured in nucleic acid concentration with Nanodrop and copies of beta globin gene with Real-Rime PCR. From the 10 samples analysed on both easyMAG and QIAamp

4.2.2 DNA concentration: buccal brush, oral rinse and urine samples.

The nucleic acid concentration and beta globin values for buccal brushes, oral rinse and urine are presented in (Table 3), none of the samples was negative with beta globin. There were nine HPV identifications in samples with less than 20 copies of beta globin, and one were in a sample with less than 10 copies. Mean, average and 95% confidence interval are presented in Tabel 3, distribution of values is illustrated with boxplots in Figure 15.

The difference is evident and confirmed with statistical testing. T-test found statistical significance between buccal brush and oral rinse, p-values 8.98E-23 and 2.49E-21 for Nanodrop and beta globin values, respectively. Similarly, there is statistical significance between oral rinse and urine, p-values 1.96E-33 and 1.65E-19 for Nanodrop and beta globin, respectively; and between buccal brush and urine, p-values 1.24E-21 and 6.08E-02 for Nanodrop and beta globin, respectively.

Table 3: Mean, average and confidence interval of the buccal brush, oral rinse and urine samples of the samples collected at Olafiaklinikken.

Sample Material	Nanodropp (ng/µl)			Beta globin (Copies)		
	Mean	Average	95% CI	Mean	Average	95% CI
Buccal brush	12.8	15.0	1.5	595	727	96
Oral rinse	43.8	46.7	5.2	2620	3440	465
Urine samples	2.8	4.3	1.3	60	526	332

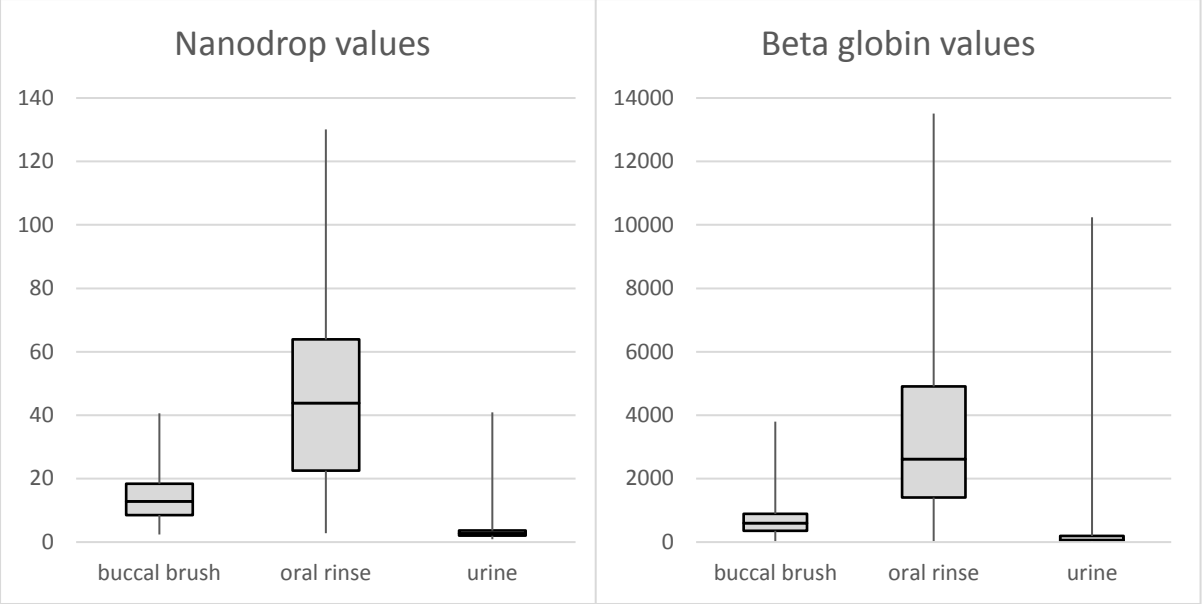


Figure 15: Boxplot of DNA concentrations as measured in nucleic acid concentration with Nanodrop and copies of beta globin gene with Real-Rime PCR. From the buccal brush; oral rinse and urine samples.

4.3 HPV prevalence for sample material and detection method

The prevalence of HPV detected by the two methods varied slightly for the oral samples. Interestingly the prevalence found in the urine samples was the same with Luminex and NGS, with 17 samples detected by both methods even though only eight of these positive were correlating. NGS data for all positives and for all positives outside the Luminex repertoire are provided. The latter of which are shaded yellow. Buccal brush and oral rinse sample both show higher prevalence with NGS. Urine samples interestingly, found less of the Luminex detectable HPV types in NGS than Luminex, although NGS found two samples positive for HPV types outside the Luminex repertoire (Table 4).

Table: 4: Overview of the total number of samples, number of HPV positive samples and the prevalence of HPV in buccal brush, oral rinse and urine samples.

Sample Material	Total number of samples	Luminex		NGS	
		Positive samples	Prevalence	Positive samples	Prevalence
Buccal brush	136	4	2.9 %	11	8.1 %
				8	5.9 %
Oral Rinse	135	6	4.5 %	9	6.7 %
				7	5.2 %
Urine	98	17	17.3 %	17	17.3 %
				15	15.3 %

For prevalence by NGS there are double values, in grey squares are samples with all HPV-types detected and in yellow are HPV-types detected that are also detectable by Luminex.

4.4 HPV genotyping

In Figure 16 the genotyping results for the Olafiaklinikken samples for both Luminex and NGS are summarized. Here the genotypes detected by all three sample materials are presented together. There was 24 patients positive in at least one sample for one or more HPV type(s) with Luminex, and 32 identifications were made in samples from these patients. For the NGS results, 33 patients were positive, and 44 identifications were made. There were 96 patients who were negative in all three sample forms. In total, 29 different HPV genotypes were detected (HPV 3, 6, 10, 13, 16, 18, 31, 32, 33, 39, 40, 43, 45, 51, 58, 59, 66, 67,68, 69, 74, 81, 82, 83, 86, 87, 89, 91 and 114).

Four genotypes were found only by Luminex (HPV 40, 43, 74 and 82). Nine genotypes were found only with NGS; four of these are detectable with Luminex (HPV 45, 58, 68 and 89) and the other five are not part of the 37 HPV type repertoire of Luminex (HPV 3, 10, 13, 32 and 114). No detections of the universal probe with Luminex was made.

Of the 42 patients with a positive HPV sample, only five had HPV identified in more than one sample type (Table 6). Seven patients had a multiple infection (Table 5). Of these, only one was not a urine sample. Luminex and NGS only agreed on multiple infection in two of the seven samples, only one of which had the same multiple genotypes with both methods.

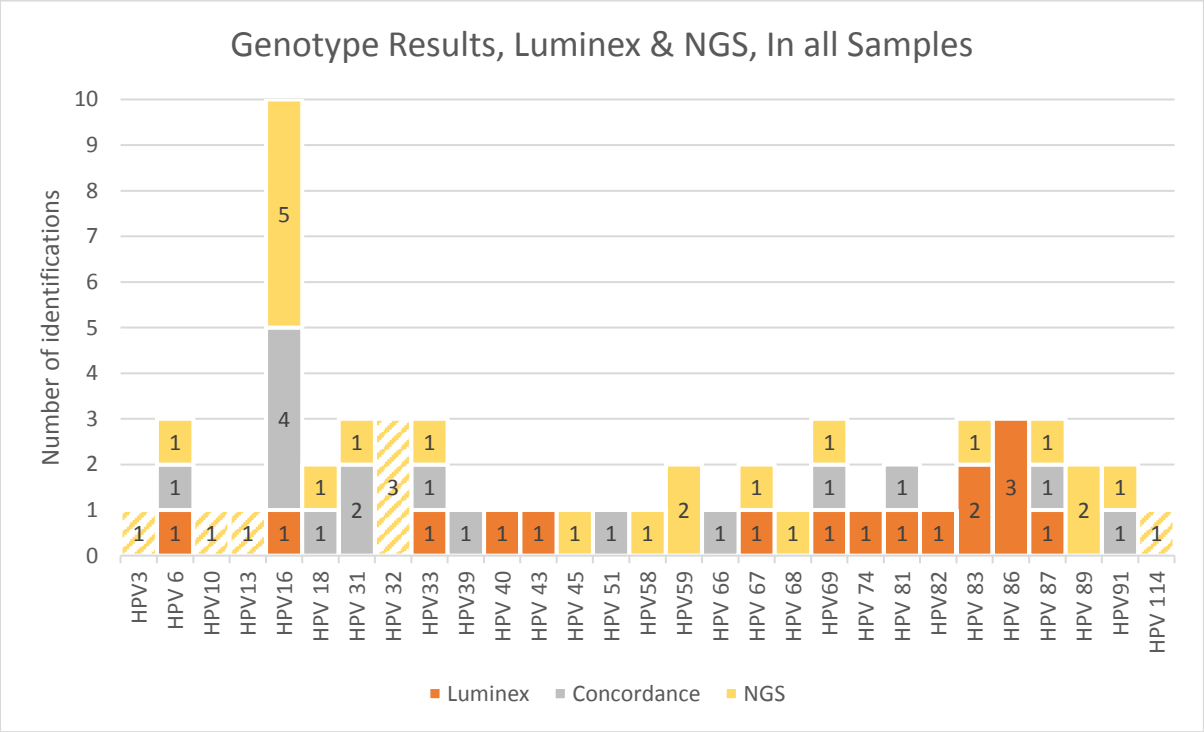


Figure 16. Genotype results for Luminex and NGS, all sample material types. Concordance between the two methods in grey, Luminex in red, NGS in yellow. White and yellow stripes represent NGS detections of HPV genotypes that are not detectable by Luminex.

Table 6: Patients with more than one HPV positive sample type. HPV status of both Luminex and NGS is provided to show correlation. Samples for these patients in which both Luminex and NGS were negative is not included.

Patient	Sample Type	Luminex	NGS
6	Buccal brush	-	HPV32
	Oral Rinse	-	HPV32
18	Buccal brush	-	HPV3
	Urine	HPV83	-
25	Buccal brush	**HPV16**	HPV67
	Oral Rinse	**HPV16**	**HPV16**
37	Buccal brush	-	HPV69
	Oral Rinse	HPV69	**HPV59**, HPV69
	Urine	HPV69	-
39	Buccal brush	**HPV16**	**HPV16**
	Oral Rinse	**HPV16**	**HPV16**

Highlighted in bold are the high-risk HPV-types (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 66 as)

Table: 5: Patients with multiple HPV infections shown by either Luminex or NGS.

Patient	Sample Type	Luminex	NGS
7	Urine	-	**HVP16**, **HPV18**
37	Oral Rinse	HPV69	**HPV59**, HPV69
44	Urine	HPV81, HPV87	HPV87
58	Urine	**HPV31**	HPV6, **HPV31**
66	Urine	HPV74, HPV81	**HPV45**, HPV81
84	Urine	**HPV16**, **HPV31**, **HPV51**	**HPV16**, **HPV31**, **HPV51**, HPV68
95	Urine	HPV86, HPV87	-

Highlighted in bold are the high-risk HPV-types (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 66 as)

4.5 Sample material specific results and concordance

When considering the genotype results for specific sample material (Figure 17), we find that among the urine samples there are more positive results, with a higher diversity of genotypes, than for the oral samples. The urine samples identified 23 genotypes. This opposed to the two oral site sample materials, buccal brush and oral rinse, which detected 10 and eight genotypes, respectively.

In some instances, Luminex make detections in a grey area (weak inconclusive result). These should be confirmed with another method or a second run with Luminex. NGS did not identify any of the grey area detections for any of the samples in question. However, due to this study exploring both Luminex and NGS, these detections were not considered confirmed negative by NGS and have been classified as neither positive nor negative; excluding them from calculations. The samples in question were one buccal brush, two oral rinse and one urine sample. A grey area HPV result was reported for two urine samples, multiple with at least one more genotype. These samples were regarded positive for the other genotype.

Concordance between Luminex and NGS varied between the sample materials. Cohen's Kappa was calculated to assess the concordance. There were 12 buccal brush samples with positive HPV detection, four with Luminex and 11 with NGS. Identifications with NGS of genotypes outside the Luminex repertoire were excluded from comparison. For buccal brush, there were three samples with such identifications; leaving eight samples for comparison. Luminex and NGS made the same identification in two samples. There were 123 samples negative with both methods. With these numbers, we find a kappa-value of 0.31, which is a fair concordance. Considering how the MGP primers have been designed to be especially sensitive towards HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 59, 66 and 68[63], we should consider the possibility of the concordance increasing if we review only these genotypes, as the detection methods are more efficient in identifying these. Calculating kappa-value for these we find 0.39 which nearly the same as the kappa-value for any genotype. Due to the special association of HPV 16 to OPSCC, it would be interesting focus on findings of this genotype. Four buccal brush samples had HPV 16 identified; only one of these four samples had HPV 16 identified with both detection methods.

There were 11 oral rinse samples with positive HPV detection, 6 with Luminex and 9 with NGS, two of which were genotypes outside the Luminex repertoire; leaving 7 for comparison. The two detection methods made the same identification in four samples. There were 122 samples negative with both methods. Calculating the Cohen's Kappa with these values, we find a kappa-value of 0.60 which is a moderate concordance. Recalculating with only MGP sensitive genotypes, we find kappa-value 0.58, which is nearly identical. Four oral rinse samples with HPV 16 were identified, two of which had HPV 16 identified with both detection methods.

There were 25 urine samples with positive HPV detection, 17 with Luminex and 17 with NGS, two of which were outside the Luminex repertoire; leaving 15 samples for comparison. The two detection methods made the same identification in eight samples. There were 72 samples negative with both methods. The kappa-value is 0.40, a fair concordance. Recalculating with only MGP sensitive genotypes, we find kappa-value 0.75, which is a substantial concordance and a clear increase from the kappa-value seen for any genotype.

Of the 12 buccal brush and 11 oral rinse samples with positive HPV detection, only four were found positive with the same genotype in both buccal brush and oral rinse. Of all HPV positive oral samples, 42% (8/19) were exclusively detected in buccal brush, 37% (7/19) exclusively in oral rinse and 21% (4/19) detected in both. Patients with negative buccal brush and oral rinse numbered 114. Calculating concordance between the sample types finds kappa-value 0.29; adjusting for MGP primer sensitivity gives kappa-value 0.34.

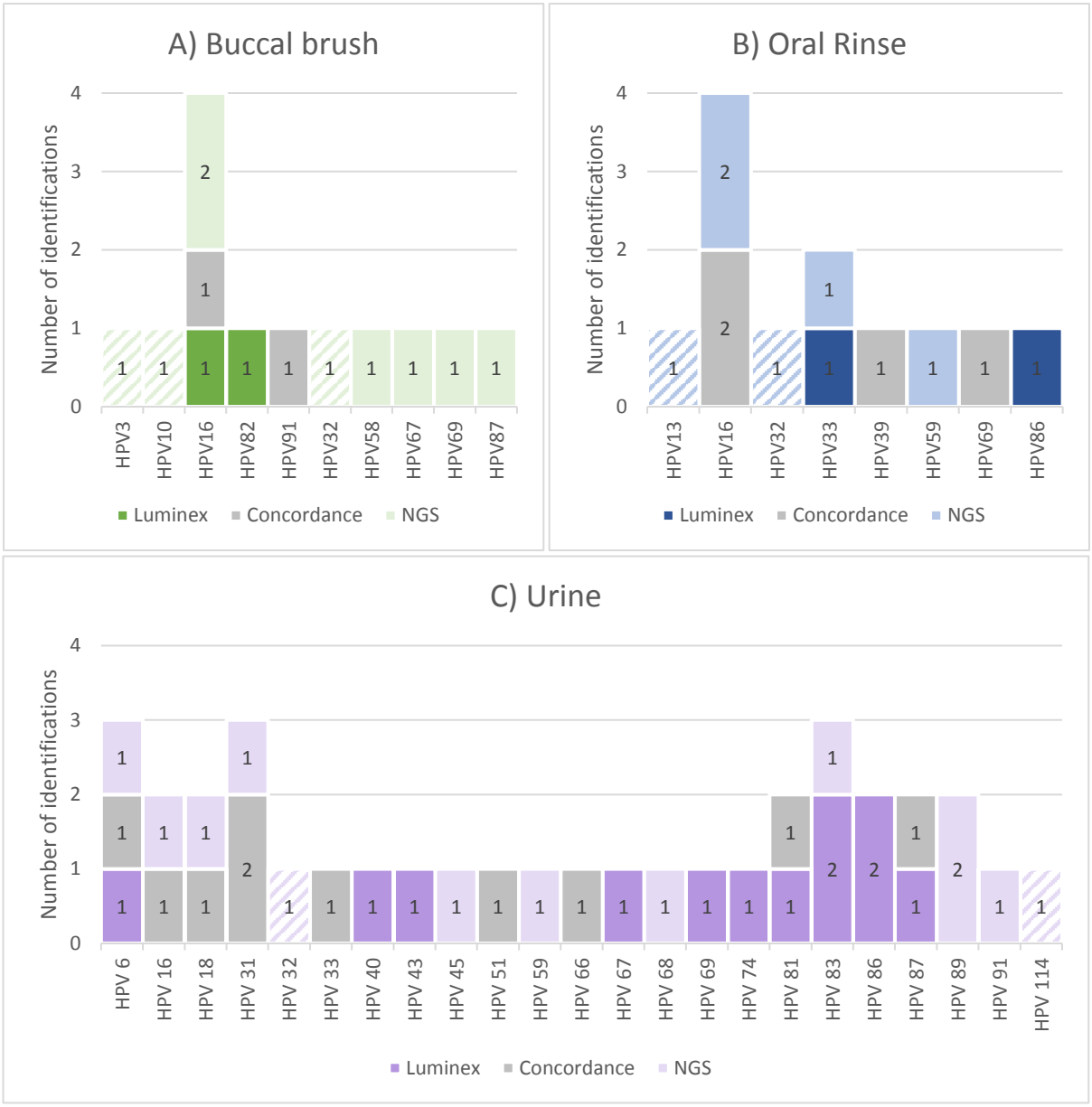


Figure 17 Genotype results from Luminex and NGS. A) Green represents buccal brush samples, white and green stripes are HPV genotypes detectable by NGS only. Dark green represents Luminex, light green NGS and grey concordance. B) Blue represents oral rinse samples. C. Purple represents the urine samples. The colour scheme for oral rinse and urine (Dark, light, grey and stripes) is as for A).

4.7 NGS reads and quality

When examining all NGS data from the 136 buccal brushes, 135 oral rinse and 98 urine samples analysed regardless of HPV detection, and assess the amount of sequences obtained we find a mean value of 6610 sequences (95% CI, 6097.7 to 7122.3) in buccal samples, 3639 sequences (95% CI, 3127.5 to 4150.5) in oral rinse, and 1593 sequences (95% CI, 1196 to 1989) in urine samples. These numbers are the same for unprocessed forward sequences and reverse sequences.

The e-gel (appendix 9) show weak and inconsistent bonds as compared to the controls, Hela and Siha. The e-gel after index PCR (appendix 10) show stronger bonds, but still inconsistent and diffuse. With Agilent we see multiple peaks in the undiluted samples, and no sample peak in the dilutions. The amount of material is also higher than expected (appendix 11).

4.8 Centrifugation of buccal brush and oral rinse during sample preparation

During this study, the speed of the centrifuge was altered two times, applying three different spin speeds; Nanodrop and beta globin values are shown in Table 7. Examining the effect of the alteration in the buccal brushes a f-test was conducted to assess the variance, finding p-value 0.04 and 0.31 in Nanodrop and beta globin, respectably. The Nanodrop p-value of 0.04 is significant and demands a t-test assuming unequal variance, giving a p-value of 0.01,

Table 7: Average, mean and 95% CI for nucleic acid concentration values (Nanodrop), and beta globin counts for buccal brush and oral rinse at 450 × g and 1000/1500 × g.

		450 × g			1000/ 1500 × g		
		Average	Mean	95% CI	Average	Mean	95% CI
Buccal brush	Nanodrop	12.8	11.8	1.9	16.5	13.9	2.1
	Beta globin	721	630	137	732	567	133
Oral rinse	Nanodrop	45.7	45.9	7.9	47.4	43.5	6.9
	Beta globin	3352	2979	657	3501	2545	646

which is of statistical significance. For beta globin, with equal variance, a t-test finds the p-value of 0.91, far from statistical significance.

Calculating significance on the centrifugation change in oral rinse, finds through f-testing that the variance is equal; p-value 0.76 and 0.20 for Nanodrop and beta globin, respectively. Performing t-test with equal variance finds p-value 0.76 and 0.76 for Nanodrop and beta globin, respectively. Indicating no statistical significance.

We saw indications of variation between the different centrifugation speeds. The alteration of centrifuge speed was originally not a result of the amount of negative Luminex results. However, the possibility of uncovering a significant weakness in the procedure led to an examination of this possibility. Once results from NGS were added we see no indication of relevance for centrifugation. The new prevalence calculated for the alteration in speed is seen in Table 8 below. The 3 speeds (450 × g, 1000 × g and 1500 × g) are separated in above and below 1000 × g.

Table 8: Number of samples and their positive in buccal brush and oral rinse, for 450 × g and 1000/1500 × g, respectively.

		Buccal brush		**Oral Rinse**	
		Positives of total samples	**Prevalence**	**Positives of total samples**	**Prevalence**
Luminex	**450 × g**	1 of 57	1.6%	1 of 56	1.8%
	1000/1500 × g	3 of 79	3.8%	5 of 79	6.3%
NGS	**450 × g**	4 of 57	7.0%	4 of 56	7.1%
	1000/1500 × g	7 of 79	8.9%	5 of 79	6.3%

5 Discussion

5.1 Establishment of DNA extraction

The Tromsø samples (n = 50) in this study were used to determine the preferred method for extraction of nucleic acids, and to establish the extraction protocol.

The question of whether the procedure for cervical samples would be suited for oral samples needed examining. A total of four approaches to extraction were completed to assess the best method for extraction of our samples, prior to the collection of the primary samples (Table 1). T-tests indicate no statistical significance, apart from the beta globin values between QIAamp and easyMag, p-value 0.02. There is the argument of whether 10 samples are enough to make an effective comparison; additionally, when examining the boxplots (Figure 14), we see beta globin values are higher with QIAamp than easyMAG, but the opposite is true for the Nanodrop, which favours easyMAG; and the p-value is close to the statistical significance level 0.05. This raises a level of uncertainty as to the weather the yield when extracting with QIAamp always will be higher, or if this finding is stochastic. However, the statistical significance cannot be ignored; ideally, a repetition with higher sample pool should be conducted if manual extraction is considered. For this study, the choice was extraction with easyMag eluting with 100 µl, due to sufficient yield, practicality in the laboratory and the findings with our different approaches.

5.2 DNA Extraction and quality of buccal brush, oral rinse and urine samples

Statistical analysis of our Nanodrop and beta globin values found statistical significance between buccal brush and oral rinse samples, oral rinse and urine samples, and buccal brush and urine samples. The observed difference is of statistical significance: oral rinse provided the highest amounts of DNA, followed by buccal brush, and the least amount of DNA was acquired in the urine samples (Figure 15).

While no samples were negative in beta globin gene amplification in our study, suggesting an adequate extraction protocol, the copy numbers could differ greatly between samples. No cut-off for beta globin was not assigned, because of the exploratory nature of the study. It is important to note that beta globin is a human gene and therefore reflects the amount of cells or cellular debris in the samples. Beta globin cannot be used to determine the successful HPV sample collection, since HPV might be in the form of free viral particles, but it will determine

the success in acquiring epithelial cells. We found that several HPV positive samples were among those with low overall beta globin; four samples with less than 20 beta globin copies were HPV positive with nine identifications among them, none were HR-HPV. This finding show that HPV detection can be achieved in samples with low beta globin. Similarly, beta globin positive samples may be false negatives, as high copy number of a human gene does not promise the detection of HPV present in the sample[60]. This is concerning considering studies were beta globin cut-offs are assigned and samples are excluded from analysis, possibly falsely excluding sufficient sample material.

5.3 HPV prevalence, detection and genotyping of buccal brush, oral rinse and urine samples

In this study, we wanted to evaluate three different sample materials as well as two different HPV detection and genotyping methods, to ascertain their suitability for future studies. Due to low HPV prevalence in the oral samples, we decided to include the urine samples in order to obtain more HPV positive samples for comparison between the two genotyping methods, leaving us with three sample materials and two methods. The nature of these different variables created several layers of comparison.

5.3.1 HPV Prevalence in buccal brush, oral rinse and urine samples

The primers used for both Luminex and NGS were the MGP system developed by Söderlund-Strand et al., a modified rendition of the primer set GP5+/6+. These primers are designed to amplify mucosal HPV types, and have been further optimised for increased sensitivity of 14 oncogenic HPVs[63]. This is ideal for research focussing on genotypes with an established relation to cancers. It is however important to remember when using MGP, that the sensitivity of these primers to oncogenic mucosal HPV makes the detection of these genotypes more likely than detection of other genotypes.

Luminex is based on hybridisation of probes, which have been meticulously designed to detect 37 specific HPV genotypes. NGS genotyping amplicons, from MGP primers, are matched to a reference, making it possible to detect additional genotypes. When comparing the two methods and their performances for oral and urine samples we need to account for the limitations of the Luminex probe set. There is a probe for universal HPV detection in Luminex, however no detections with this probe were made. Apart from this possible universal probe detection, we know that Luminex will not detect additional genotypes. From

here on the prevalence of genotypes detected with NGS that is included in the probe-set in Luminex will be referred to as “adapted”.

The prevalence of any HPV genotype detected in the buccal brush samples, was 2.9 % (4/136) with Luminex, and 8.1 % (11/136) with NGS (Table 4). The adapted prevalence is 5.9 % (8/136). The adapted prevalence with NGS is higher than with Luminex with twice as many identifications. Only two identifications were the same by the two methods.
The oral rinse samples provided a prevalence of 4.5 % (6/135) with Luminex. With NGS the prevalence was 6.7 % (9/135), and the adapted prevalence 5.2 % (7/135) (Table 4). Four of the identifications were the same by the two methods. Although the number of positives by both two oral sampling methods is quite low, it is interesting to note that the prevalence detected with Luminex is higher in oral rinse than buccal brush, 4.5% vs 2.9%. NGS on the other hand indicates the buccal brush to be more HPV rich than oral rinse; 8.1% vs 6.7%. The prevalence of genotypes only within the Luminex probe-set detected with NGS is also higher in buccal brush than oral rinse, 5.9% vs 5.2%.
The prevalence of our oral samples are entirely within the reported range for oral HPV in MSM, 2-14%[26,34-36], which does not differ much from the oral prevalence in non MSM 4.0 - 11.5% [17,24-30].

The urine samples provided a prevalence of 17.3 % (17/98) by Luminex and a prevalence of 17.3 by NGS (17/98). The adapted prevalence with NGS was 15.3% (15/98) (Table 4). Our finding are within the range of HPV prevalence reported in male urine 5.8 – 36.7%, as reviewed by Enerly et al.[67]. In our urine samples, Luminex is shown as more sensitive than the adapted prevalence with NGS, as opposed oral samples where the prevalence by NGS was higher in all instances. There is a level of uncertainty as argued by Enerly et al. [67], in spite of the prevalence reported in their review, and by Giuliano et al. that ended the analysis of HPV from male urine due to low positives[21.]

It is believed that 70-90% of OPSCCs are caused by HPV 16[14,16-19], thus creating an interest of individually examining this genotype. In the present study, 32% (6/19) of the patients with positive HPV detection in one of the oral samples were positive for HPV 16. That is a prevalence of 4.3% (6 of 138) of patients in the study. This is higher than other reported oral prevalence of HPV 16, which is at 0.6% - 1.8%[17,26,30]. For the urine samples, 8 % (2 of 25)

of HPV positives were positive for HPV 16. That is 2 % (2 of 99) of patients in the study who provided a urine sample.

It appears that our population of MSM does not exhibit a higher prevalence of oral HPV than what is expected from heterosexual men. This could be due to lower transmission rates when performing oral sex on men, or the effective clearance of oral HPV in immunocompetent individuals. This is in line with findings by D'Souza et al.[24], who suggest that preforming oral sex on women constitute a higher risk of acquiring oral HPV than preforming oral sex on men due to more efficient transfer from female genitalia. Another study conducted by D'Souza et al.[51] examined risk of oral HPV infection in partners of HPV positive OPSCCs. They found no higher incidence of oral HPV in these partners suggesting low oral-to-oral transmission, or the effective clearance of any infections. They did find that several of the OPSCC patients had current or previous partners with a history of cervical dysplasia or cervical invasive cancer[51]. It would be sensible suggesting that heterosexual men, especially partners of women with oncogenic cervical HPV, is an at-risk group.

5.3.2 Concordance between Luminex and NGS

The concordance between the methods defined with kappa-value were, 0.31, 0.60 and 0.40 for buccal brush, oral rinse and urine, respectively. It is interesting how Luminex and NGS show a concordance in oral rinse that is much better than in buccal brush and urine samples. Especially if we considered the findings discussed from our reported prevalence where there was disagreement between the methods as to which sample material was superior. The difference between the methods was small, with only a few positive HPV detections making out the difference. It is clear that the reproducibility between the methods is highest in the oral rinse samples, which would also indicate oral rinse as a more suited material.

As discussed, a focus on MGP sensitive genotypes is pertinent; we expect detection of these to be higher. The concordance between the methods on MGP sensitive genotypes was; kappa-value 0.39, 0.58 and 0.75 for buccal brush, oral rinse and urine samples, respectively. Concordance did not change by much for the oral samples, which is not surprising as the number of HPV positives were low regardless. However, the urine samples stand out here with nearly a doubling in concordance as per kappa-value. Indeed, NGS detected MGP sensitive genotypes in the same samples as Luminex, with three additional detections not

made by Luminex. This shows that the increases sensitivity of the MGP primers increased the reproducibility of these genotypes in the urine samples.

It would not be correct to evaluate one of our detection methods against the other, as we are exploring the possible usage of both methods; and neither is a gold-standard. Furthermore, both methods made detections not identified with the other. All samples with an HPV detection not made by both detection methods should ideally be tested against a third method. This could bring some clarity to the discrepancy stemming from e.g. DNA degradation over time, false or low-level positives and false negatives.

Granted the positive identifications are few and do not support hard conclusions, when calculating kappa-value the difference of one sample changes the outcome considerably; the findings could be stochastic.

In the oral samples NGS made more detections than Luminex, in urine samples the number of detections were the same, but NGS made more MGP sensitive detections. The difference here is also not substantial. All variation observed between Luminex and NGS for the three sample materials shows little difference, but favouring NGS. To clearly establish a preferred detection method, further studies with larger sample size is needed. Meisal et al.[68] comparing HPV genotyping with Luminex and NGS in urine samples found NGS to be more sensitive, which is in line with the inclinations of this study. In addition, as this study do not eliminate one of the methods for future use in oral HPV detections, other factors such as an assessment of cost, labour intensity, time etc. may be considered.

A possible factor affecting the difference between the two methods could be the time between the two methods and possible degradation of sample material. Luminex was conducted in several stages after acquiring enough samples for a full run. NGS was preformed after all the samples had been collected and processed, as this method could analyse a larger number of samples simultaneously. All samples were extracted the day after collection and then frozen, degradation in this stage should affect both detection methods. However, due to NGS being conducted later, all samples analysed with NGS had been thawed and stored in fridge to conduct Luminex then refrozen. This additional freeze-thaw cycle and subsequent wait to be re-frozen introduces an element of possible degradation that did not affect Luminex. This could explain the samples in which Luminex detected HPV and NGS did not. Most Luminex detections not made by NGS represented the low end of Luminex probe counts (data not

shown). These samples could have degraded to undetectable levels, and therefore not detected with NGS. However, NGS still made more HPV detections than Luminex, this could seem to contradict the theory. To obtain additional evidence of this theory, a second analysis of beta globin could be conducted; a substantial fall in beta globin could corroborate the theory of DNA degradation.
Detections made by NGS and not Luminex could be a result of internal competition for probes and signalling chemistry leading to none of the genotypes being detected.

5.3.3 Concordance between sample materials

5.3.3.1 Buccal brush and oral rinse

As this study aims to find a suitable method for future studies, we wanted to examine oral HPV sampling with more than one approach. Hence, we included both buccal brushes and oral rinse. In two studies reviewed both sampling methods had been used on a normal healthy population[25,36], however both of them also swabbed the throat/oropharynx with the brushes, and so not directly comparable to our approach. Commonly oral rinse/saliva was the only sample material[17,24,26-29,32,34,51,59,60]. Any form of brush sampling is typically used for sampling patients with identifiable malignant lesions in the oral cavity[47]. This direct brushing on lesions is not comparable to brushing of healthy mucosa, primarily due to the criteria for inclusion: healthy vs identifiable malignancy. Furthermore, a targeted swab of a lesion is expected to pick up the cells of interest, where HPV should be abundant. This contrasts with brushing healthy mucosa where an infection not visible to the naked eye could be present outside the brushed area, and HPV DNA not present in the sample or at very low levels.

A total of 10 different genotypes were detected with any of the two methods in the buccal brush samples, two of which were HR-HPV (HPV 3, 10, **16**, 32, **58**, 67, 69, 82, 87 and 91) (Figure 17a.). Eight genotypes was detected in the oral rinse samples, four of which were HR-HPV (HPV 13, **16**, 32, **33**, **39**, **59**, 69 and 86) (Figure 17b.). Only three genotypes were detected in both sample materials (HPV 16, 32 and 69). Our question on whether there would be different genotypes in the buccal brush and oral rinse might be of merit. This study shows a possible indication of a difference, whether this translate to a larger scale is not certain. All genotypes identified from the oral samples belong to the alpha-papillomaviruses, which is expected as we sample mucosal epithelium and use MGP. It is interesting to point out that four of the genotypes (HPV 3, 10, 13 and 32) detected by NGS in the oral samples are not

part of the probe-set in Luminex. HPV 3 and 10 cause flat warts that have been identified in lesions on the face[11], HPV 13 is associated with focal epithelial hyperplasia[69], and HPV 32 cause genital warts[11]. All are alpha papillomaviruses[10], and finding these genotypes in our buccal brush samples is not surprising. The interest and importance of detecting these genotypes can be argued. It should be considered in future studies if it is important to detect genotypes other than the 37 genotypes in the Luminex repertoire. The detections of genotypes outside the Luminex probe set in this study constituted 25% (3/12) of the positive buccal brush samples, and 18% (2/11) of the oral rinse samples.

Only four of the 19 patients with a positive oral sample were positive with both the oral samples. Of all positive oral samples, 42% were only positive in buccal brush, 36% in oral rinse and 21% in both. Our concordance between buccal brush and oral rinse had a kappa-value 0.29. Our findings are similar with a study by Edelstein et al. who found that in all their positive samples; 49% were found in only oral rinse samples, 39% in only self-collecting brushes, and 12% were detected with both methods[25]. Their genotyping platform was Luminex although they used different extraction platform and primers[25].

A number of factors could explain the reason for this discrepancy. The site of sample collection is not identical and complete concordance between buccal brush and oral rinse was not expected. Oral rinse could dislodge HPV either in free viral form or associated to epithelial cells, it is not a targeted sample[48], there is no telling from where in the oral cavity or oropharynx the HPV originates. Because of this an oral rinse could in theory sample the same infection as buccal brushes due to the overlap in sample site. The concordance between the sample materials is likely because of this overlap; however, this was not the case in all patients. It is possible that oral rinse samples not confirmed with a buccal brush is the result of an infection in the oropharynx, although this is highly likely, it cannot be confirmed without a brush sample from the oropharynx. Buccal brush positives not confirmed with oral rinse samples can indicate infections in or around the buccal lining, or the ability of the brushes to collect sells and make micro biopsies, therefore collecting epithelial cells not loose enough for an oral rinse to collect.

PCR inhibitors might also be present. Oral samples might have a bigger problem with PCR inhibitors than do other anogenital samples.[60] There is no certainty that buccal brush and oral rinse from the same patient were extracted with equal quality, or that PCR amplifications of these extracted products were of the same quality.

Our buccal brush samples were collected prior to the oral rinse, this would prevent the oral rinse from possibly collecting all lose epithelial cells so that the buccal brush would underperform. Interestingly it has been reported that recent tooth brushing increases detection of HPV, which could mean that brushing the oral cavity and oropharynx prior to oral rinse would improve HPV detection in oral rinse samples[36]. It is therefore possible that our buccal brushes aided collection of the oral rinse samples. Light abrasion of the oral mucosa prior to oral rinse could be considered as a possible approach to improve sample collection.

The low concordance between buccal brush and oral rinse sample would indicate that these two sampling methods cannot be rendered obsolete by the other. Each of them found distinct positives not confirmed by the other. Oral rinse however is more appropriate to the topic of OPSCCs, with the tonsils having the higher prevalence of HPV of the oral sites[11].
Considering how oral rinse found more high-risk HPV, and greater concordance between the two detection methods; it would appear in this study that oral rinse is a more appropriate sample material. Which is in line with others conclusions[36], and the research on oral HPV were oral rinse seems to be the chosen material as discussed.
Our buccal brushes are not 'site accurate' to the primary location for oral HPV lesions, and although oral rinse is not site specific it included these sites. Buccal brush has in this study been showed as a sufficient sample material for oral HPV detection, finding nearly as many identifications as oral rinse. It would be interesting to attempt brushings of the oropharynx in a future study, to assess whether sampling involving scraping improves detection, as theorised[36].

5.3.3.2 Oral- and urine samples

Of the 29 genotypes identified from either oral or urine samples, nine were found in both places, giving a genotype agreement of 31% (9/29). Only two patients were HPV positive in both an oral sample and a urine sample, only one of them had the same genotype in both sample types (Table 6). There are not enough findings in this present study to evaluate agreement between oral and anogenital HPV infections. One study collecting oral rinse and anogenital samples (urine sample, intra-anal and external genital swabs), found none of the same genotypes identified at both sites in their participants[34]. As discussed, it has been questioned if urine samples are adequate for HPV detection in men. To examine whether patients with oral HPV exhibit the same infections at anogenital sites, sample material other

than urine might be more suited. Deshmunkh et al.[30] using oral rinse and penile swabs, found that the prevalence of oral HPV infections among men with genital HPV infection, regardless of genotype, was 19.3%. It might be worth considering examining anogenital and oral HPV in future studies to assess angoential HPV infection as a risk factor for oral HPV, but this should be examined using anogenital swabs rather than urine, if examining men.

5.4 Quality of samples during HPV detection with NGS

After PCR amplification, a selection of 10 random samples were run with gel electrophoreses on an E-gel to assess the quality after MGP amplification of the product. These samples all contained noise (background fragments). After the indexation PCR, a new run on E-gel was performed with 10 different samples, and again considerable amounts of noise were present. The reason for this is difficult to pinpoint. The SiHa and HeLa cell line controls were also run on the gel, showing clear concise bands, this excludes contamination as the source of the background noise. No negative controls (water blank) were included, based on the knowledge that the majority of the samples were HPV negative. In addition, contamination due to laboratory conditions has been excluded in previous studies. Still, we realize that including negative controls would be valuable in order to remove any doubt. The running theory for the noise is the use of generalised primers amplifying off-target and the impurity of the sample material. The primers, although specific for HPV, might pick up and amplify bacteria and other genetic material present in the oral samples. This is strengthened by the amount of genetic material present in the sample. Agilent Bioanalyzer was utilized to assess the fragment size and found the amount of DNA, our samples showed the DNA amount to be about 10 times greater than commonly found in-house with this sequencing protocol. Furthermore, the size of the fragments varied considerably showing no uniform length.

When examining the statistical data of the NGS samples, we see an inconsistent degree of quality at the sample level with good quality of the sequences overall. Numerous samples showed less than 1000 reads which even for HPV negative samples is unexpectedly low. It appears however not to be any pattern for these samples. A reconstruction of the sample plates with their reads was created to asses any possible patterns; e.g. if a whole row was poor in quality, it could be errors in pipetting. No such patterns emerged. The only way of pattern is a steady increase in low quantity samples for each consecutive plate, the best being the first and the worst being the third and fourth. If the decrease in quality per plate is a result

of the work conducted in the lab is difficult to assert. An alternative explanation could be that the decrease in quality coincides with the transition from buccal brush to oral rinse to urine. It is the oral rinse samples that seems to exhibit the most samples with a low quality, followed by urine, and then buccal brushes having nearly none. If this is the case, then amount of DNA identified would indicate buccal brushes stored in PreservCyt as superior sample material for DNA collection according to NGS. Whether it is the buccal brush sample technique or the PreservCyt that constitutes the vital factor is not possible to tell from the data at hand.

There is no general recommendation for cut-off with the NGS genotyping method and sample material in this study. In future, more genotyping of HPV in oral samples with NGS should be conducted to determine appropriate cut-of. In this study cut-off was assigned after optioning the sequenced reads. As we explore what can be detected, it is reasonable to have a low cut-off to assess what we actually detect, still we did not want to include all the samples with low counts that might be false positives. A cut-off of 50 reads was chosen for this study. In future studies the cut-off limit can be increased to serve the appropriate purpose. Ideally, more research will be available in future, providing recommendations for cut-offs specific genotypes and sample materials. When comparing to detection methods, Meisal et. al[68] used a cut-off of 20 reads, but also evaluated the changes in agreement with cut-offs 10, 20, 50 and 100. It is desired to have a cut-off low enough to include all true positives, while keeping it high enough to avoid false positives[68]. Finding this balance is difficult, and the ideal cut-off might change between the genotypes. Having a third detection method, or confirmed detections to compare with could help assert the ideal cut-off. Confirmation of our detections with low read counts with a third method could also help separate the true positives and the false positives.

5.5 Thoughts for future studies

We used 50ml sample tubes when collecting oral rinse samples. This was to allow the patients to redeposit the saline into the tube, and so the diameter of the opening should be wide enough to serve this purpose. The tubes we used were also designed with a v-shaped bottom to allow centrifugation and supernatant discarding by inversion of the tube. However due to high amounts of mucous the pellets did not stick when inverting the tubes, leading to possible loss of sample material.

There was no statistical significance between the different centrifugation speeds when examining both the Nanodrop and beta globin results for the oral rinse. The buccal brush showed statistical significance in its Nanodrop values ($p = 0.01$), indicating statistical significance. This is in contrast to the beta globin values, which are far from statistically significant (p = value 0.91). Due to statistically significant Nanodrop values, we cannot exclude the possibility that the change in centrifugation did increase the genetic material output. Considering how beta globin is a more reliable measure on the presence of genomic DNA, it is reasonable to doubt if the finding of statistical significance in the Nanodrop values should be trusted. However, the samples themselves proved far easier to work with when increasing the centrifugation speed, as the pellets formed harder and more compact, making discarding of supernatant less problematic. Mucus was still a challenge, especially for oral rinse, but was usually solved with extra time manually pipetting/mixing. When larger numbers of samples were analysed on Luminex, it appeared as if the change in centrifugation speed was resulting in an increase in HPV positive samples, suggesting that the speed of centrifugation might be an essential criteria for successful detection of HPV. We pooled the HPV positives in groups of less than and greater than 1000 × g centrifugation and calculated their identical prevalence based on total amount of sample for the different speeds (Table 8). Although there was an increase in the number of positives after the centrifugation increase with Luminex, results from NGS showed no pattern to indicate a difference in the number of positives before and after speed alteration. The findings for Luminex could be stochastic. Another explanation could be that the sensitivity on Luminex is poorer, and so low centrifugation resulted in acquiring viral DNA at levels undetectable with Luminex, but detectable with NGS.

6 Conclusions

In order to recommend a sample material, buccal brush or oral rinse, and a detection method, Luminex or NGS, we see a need for further testing. Using a standard DNA extraction protocol proved satisfactory. All three sample materials have been successful in terms of providing HPV DNA. Both HPV detection methods were found suitable for the detection and genotyping of HPV from oral and urine samples from men. No approach was uniformly superior to the other.

It is not feasible to assert one oral sample material as superior to the other, but there are indications of oral rinse being a better approach. Oral rinse show higher DNA amounts after extraction and better reproducibility of detections between Luminex and NGS. However, with regards to HPV detection, it performed similar to the buccal brush samples.

The discrepancy between the two HPV genotyping methods is evident. The correlation is poor, but the number of identifications is similar. The correlation increased considerably in the urine samples when focusing on the genotypes the MGP are designed most sensitive to, but increased with less in the oral samples. The concordance constituted a small size of the overall identifications, and both methods made detections not confirmed by the other. NGS has the advantage in its ability to detect additional genotypes outside the panel included for Luminex. If an extended panel of detectable genotypes is important, then the advantage of NGS is clear. Inhibitors or degradation of samples could have reduced the sensitivity of our NGS.

The HPV prevalence observed is within the documented area, somewhat lower than originally expected. MSM is expected to have higher HPV prevalence in the anogenital region than heterosexual men, it is clear that this is not the case for the oral region. The MSM population at Olafiaklinikken were willing to provide sample material. Making future recruitment feasible.

7 References

1. zur Hausen H. Papillomaviruses in the causation of human cancers - a brief historical account. Virology. 2009;384(2):260-5.
2. The Nobel Prize in Physiology or Medicine 2008 [press release]. Karolinska Institutet, 06 Oct 2008 2008.
3. Forman D, de Martel C, Lacey CJ, Soerjomataram I, Lortet-Tieulent J, Bruni L, et al. Global burden of human papillomavirus and related diseases. Vaccine. 2012;30 Suppl 5:F12-23.
4. Bzhalava D, Eklund C, Dillner J. International standardization and classification of human papillomavirus types. Virology. 2015;476:341-4.
5. Doorbar J, Quint W, Banks L, Bravo IG, Stoler M, Broker TR, et al. The biology and life-cycle of human papillomaviruses. Vaccine. 2012;30 Suppl 5:F55-70.
6. Egawa N, Egawa K, Griffin H, Doorbar J. Human Papillomaviruses; Epithelial Tropisms, and the Development of Neoplasia. Viruses. 2015;7(7):3863-90.
7. de Villiers EM. Cross-roads in the classification of papillomaviruses. Virology. 2013;445(1-2):2-10.
8. Muhr LSA, Eklund C, Dillner J. Towards quality and order in human papillomavirus research. Virology. 2018;519:74-6.
9. International Human Papillomavirus (HPV) Reference Center. Human Reference Clones 2019 [updated 11 Mar 2019; cited 10 May 2019. Available from: https://www.hpvcenter.se/human_reference_clones/.
10. International Agency for Research on Cancer. IARC Monographs Human Papillomaviruses. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Biological Agents Volume 100 B A Review of Human Caricinogens. Lyon France: International Agency for Research on Cancer; 2012.
11. Doorbar J, Egawa N, Griffin H, Kranjec C, Murakami I. Human papillomavirus molecular biology and disease association. Rev Med Virol. 2015;25 Suppl 1:2-23.
12. Burk RD, Harari A, Chen Z. Human papillomavirus genome variants. Virology. 2013;445(1-2):232-43.
13. Mirabello L, Yeager M, Cullen M, Boland JF, Chen Z, Wentzensen N, et al. HPV16 Sublineage Associations With Histology-Specific Cancer Risk Using HPV Whole-Genome Sequences in 3200 Women. Journal of the National Cancer Institute. 2016;108(9).
14. Fossum GH, Lie AK, Jebsen P, Sandlie LE, Mork J. Human papillomavirus in oropharyngeal squamous cell carcinoma in South-Eastern Norway: prevalence, genotype, and survival. Eur Arch Otorhinolaryngol. 2017;274(11):4003-10.
15. Pinatti LM, Walline HM, Carey TE. Human Papillomavirus Genome Integration and Head and Neck Cancer. Journal of dental research. 2018;97(6):691-700.
16. Rettig EM, D'Souza G. Epidemiology of head and neck cancer. Surg Oncol Clin N Am. 2015;24(3):379-96.
17. Gillison ML, Broutian T, Pickard RK, Tong ZY, Xiao W, Kahle L, et al. Prevalence of oral HPV infection in the United States, 2009-2010. Jama. 2012;307(7):693-703.
18. Gillison ML, D'Souza G, Westra W, Sugar E, Xiao W, Begum S, et al. 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):407-20.
19. Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, 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):4294-301.
20. ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre). Human Papillomavirus and Related Diseases Report. 22 jan 2019.

21. Giuliano AR, Nielson CM, Flores R, Dunne EF, Abrahamsen M, Papenfuss MR, et al. The optimal anatomic sites for sampling heterosexual men for human papillomavirus (HPV) detection: the HPV detection in men study. The Journal of infectious diseases. 2007;196(8):1146-52.
22. Cuschieri K, Nandwani R, McGough P, Cook F, Hogg L, Robertson C, et al. Urine testing as a surveillance tool to monitor the impact of HPV immunization programs. Journal of medical virology. 2011;83(11):1983-7.
23. Koene F, Wolffs P, Brink A, Dukers-Muijrers N, Quint W, Bruggeman C, et al. Comparison of urine samples and penile swabs for detection of human papillomavirus in HIV-negative Dutch men. Sexually transmitted infections. 2016;92(6):467-9.
24. D'Souza G, Cullen K, Bowie J, Thorpe R, Fakhry C. Differences in oral sexual behaviors by gender, age, and race explain observed differences in prevalence of oral human papillomavirus infection. PloS one. 2014;9(1):e86023.
25. Edelstein ZR, Schwartz SM, Hawes S, Hughes JP, Feng Q, Stern ME, et al. Rates and determinants of oral human papillomavirus infection in young men. Sexually transmitted diseases. 2012;39(11):860-7.
26. Kreimer AR, Villa A, Nyitray AG, Abrahamsen M, Papenfuss M, Smith D, et al. The epidemiology of oral HPV infection among a multinational sample of healthy men. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2011;20(1):172-82.
27. Du J, Nordfors C, Ahrlund-Richter A, Sobkowiak M, Romanitan M, Nasman A, et al. Prevalence of oral human papillomavirus infection among youth, Sweden. Emerg Infect Dis. 2012;18(9):1468-71.
28. Conway DI, Robertson C, Gray H, Young L, McDaid LM, Winter AJ, et al. Human Papilloma Virus (HPV) Oral Prevalence in Scotland (HOPSCOTCH): A Feasibility Study in Dental Settings. PloS one. 2016;11(11):e0165847.
29. Ortiz AP, Gonzalez D, Vivaldi-Oliver J, Castaneda M, Rivera V, Diaz E, et al. Periodontitis and oral human papillomavirus infection among Hispanic adults. Papillomavirus research (Amsterdam, Netherlands). 2018;5:128-33.
30. Sonawane K, Suk R, Chiao EY, Chhatwal J, Qiu P, Wilkin T, et al. Oral Human Papillomavirus Infection: Differences in Prevalence Between Sexes and Concordance With Genital Human Papillomavirus Infection, NHANES 2011 to 2014. Annals of internal medicine. 2017;167(10):714-24.
31. Rautava J, Willberg J, Louvanto K, Wideman L, Syrjanen K, Grenman S, et al. Prevalence, genotype distribution and persistence of human papillomavirus in oral mucosa of women: a six-year follow-up study. PloS one. 2012;7(8):e42171.
32. de Souza MMA, Hartel G, Whiteman DC, Antonsson A. Detection of oral HPV infection - Comparison of two different specimen collection methods and two HPV detection methods. Diagnostic microbiology and infectious disease. 2018;90(4):267-71.
33. Hearnden V, Murdoch C, D'Apice K, Duthie S, Hayward NJ, Powers HJ. Oral human papillomavirus infection in England and associated risk factors: a case-control study. BMJ Open. 2018;8(8):e022497.
34. King EM, Gilson R, Beddows S, Soldan K, Panwar K, Young C, et al. Oral human papillomavirus (HPV) infection in men who have sex with men: prevalence and lack of anogenital concordance. Sexually transmitted infections. 2015;91(4):284-6.
35. Zou H, Tabrizi SN, Grulich AE, Garland SM, Hocking JS, Bradshaw CS, et al. Early acquisition of anogenital human papillomavirus among teenage men who have sex with men. The Journal of infectious diseases. 2014;209(5):642-51.
36. Read TR, Hocking JS, Vodstrcil LA, Tabrizi SN, McCullough MJ, Grulich AE, et al. Oral human papillomavirus in men having sex with men: risk-factors and sampling. PloS one. 2012;7(11):e49324.
37. Machalek DA, Poynten M, Jin F, Fairley CK, Farnsworth A, Garland SM, et al. Anal human papillomavirus infection and associated neoplastic lesions in men who have sex with men: a systematic review and meta-analysis. The Lancet Oncology. 2012;13(5):487-500.

38. Morgan IM, DiNardo LJ, Windle B. Integration of Human Papillomavirus Genomes in Head and Neck Cancer: Is It Time to Consider a Paradigm Shift? Viruses. 2017;9(8).
39. McBride AA, Warburton A. The role of integration in oncogenic progression of HPV-associated cancers. PLoS pathogens. 2017;13(4):e1006211.
40. Williams VM, Filippova M, Soto U, Duerksen-Hughes PJ. HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress. Future virology. 2011;6(1):45-57.
41. Christiansen IK, Sandve GK, Schmitz M, Durst M, Hovig E. Transcriptionally active regions are the preferred targets for chromosomal HPV integration in cervical carcinogenesis. PloS one. 2015;10(3):e0119566.
42. D'Souza G, Kreimer AR, Viscidi R, Pawlita M, Fakhry C, Koch WM, et al. Case-control study of human papillomavirus and oropharyngeal cancer. The New England journal of medicine. 2007;356(19):1944-56.
43. Rettig EM, Wentz A, Posner MR, Gross ND, Haddad RI, Gillison ML, et al. Prognostic Implication of Persistent Human Papillomavirus Type 16 DNA Detection in Oral Rinses for Human Papillomavirus-Related Oropharyngeal Carcinoma. JAMA Oncol. 2015;1(7):907-15.
44. Gillison ML, Chaturvedi AK, Anderson WF, Fakhry C. Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(29):3235-42.
45. Giovannelli L, Campisi G, Lama A, Giambalvo O, Osborn J, Margiotta V, et al. Human papillomavirus DNA in oral mucosal lesions. The Journal of infectious diseases. 2002;185(6):833-6.
46. Blioumi E, Chatzidimitriou D, Pazartzi C, Katopodi T, Tzimagiorgis G, Emmanouil-Nikoloussi EN, et al. Detection and typing of human papillomaviruses (HPV) in malignant, dysplastic, nondysplastic and normal oral epithelium by nested polymerase chain reaction, immunohistochemistry and transitional electron microscopy in patients of northern Greece. Oral oncology. 2014;50(9):840-7.
47. Termine N, Giovannelli L, Rodolico V, Matranga D, Pannone G, Campisi G. Biopsy vs. brushing: comparison of two sampling methods for the detection of HPV-DNA in squamous cell carcinoma of the oral cavity. Oral oncology. 2012;48(9):870-5.
48. Kreimer AR, Pierce Campbell CM, Lin HY, Fulp W, Papenfuss MR, Abrahamsen M, et al. Incidence and clearance of oral human papillomavirus infection in men: the HIM cohort study. Lancet (London, England). 2013;382(9895):877-87.
49. Shigeishi H, Sugiyama M. Risk Factors for Oral Human Papillomavirus Infection in Healthy Individuals: A Systematic Review and Meta-Analysis. J Clin Med Res. 2016;8(10):721-9.
50. Cancer Registry of Norway. Cancer in Norway 2017 - Cancer incidence, mortality, survival and prevalence in Norway. Oslo: Cancer Registry of Norway.
51. D'Souza G, Gross ND, Pai SI, Haddad R, Anderson KS, Rajan S, et al. Oral human papillomavirus (HPV) infection in HPV-positive patients with oropharyngeal cancer and their partners. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2014;32(23):2408-15.
52. WHO. Immunization, Vaccines and Biological database2018 [cited 2019 01 Feb].
53. Molden T, Feiring B, Ambur OH, Christiansen IK, Hansen M, Laake I, et al. Human papillomavirus prevalence and type distribution in urine samples from Norwegian women aged 17 and 21 years: A nationwide cross-sectional study of three non-vaccinated birth cohorts. Papillomavirus research (Amsterdam, Netherlands). 2016;2:153-8.
54. Feiring B, Laake I, Christiansen IK, Hansen M, Stalcrantz J, Ambur OH, et al. Substantial Decline in Prevalence of Vaccine-Type and Nonvaccine-Type Human Papillomavirus (HPV) in Vaccinated and Unvaccinated Girls 5 Years After Implementing HPV Vaccine in Norway. The Journal of infectious diseases. 2018;218(12):1900-10.
55. HPV - vaksine (Humant papillomavirus) - veileder for helsepersonell. Updated 15 feb 2019 [cited 25 apr 2019]. In: Vaksinasjonsveilederen for helsepersonell (Vaksinasjonsboka) [Internet].

Oslo: Folkehelseinstituttet, [cited 25 apr 2019]. Available from: https://www.fhi.no/nettpub/vaksinasjonsveilederen-for-helsepersonell/vaksiner-mot-de-enkelte-sykdommene/hpv-vaksinasjon-humant-papillomavir/#hpvvaksiner.
56. Folkehelseinstituttet. Tilbud om gratis HPV-vaksine (Gardasil) fra folkehelseinstituttet til utvalgte grupper 03 sep 2018 [cited 25 apr 2019]. Available from: https://www.fhi.no/globalassets/bilder/vaksine/18-11440-2-gardasil-345584_1_1.pdf.
57. Folkehelseinstituttet. Vaksine mot humant papillomavirus (HPV) Rapport fra arbeidsgruppe nedsatt av Nasjonalt folkehelseinstitutt for å vurdere om HPV vaksine til gutter skal tilbys i program i Norge. Oslo: Folkehelseinstituttet; 2016.
58. Chesson HW, Dunne EF, Hariri S, Markowitz LE. The estimated lifetime probability of acquiring human papillomavirus in the United States. Sexually transmitted diseases. 2014;41(11):660-4.
59. Antonsson A, Cornford M, Perry S, Davis M, Dunne MP, Whiteman DC. Prevalence and risk factors for oral HPV infection in young Australians. PloS one. 2014;9(3):e91761.
60. 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):5526-35.
61. Pierce Campbell CM, Kreimer AR, Lin HY, Fulp W, O'Keefe MT, Ingles DJ, et al. Long-term persistence of oral human papillomavirus type 16: the HPV Infection in Men (HIM) study. Cancer prevention research (Philadelphia, Pa). 2015;8(3):190-6.
62. Qiagen. QIAamp DNA Mini and Blood Mini Handbook. Fifth ed: Qiagen; 2016 May 2016. 70 p.
63. Soderlund-Strand A, Carlson J, Dillner J. Modified general primer PCR system for sensitive detection of multiple types of oncogenic human papillomavirus. Journal of clinical microbiology. 2009;47(3):541-6.
64. Schmitt M, Bravo IG, Snijders PJ, Gissmann L, Pawlita M, Waterboer T. Bead-based multiplex genotyping of human papillomaviruses. Journal of clinical microbiology. 2006;44(2):504-12.
65. Illumina Inc. An introduction to Next-Generation Sequencing Technology. 2017. Available from: https://www.illumina.com/documents/products/illumina_sequencing_introduction.pdf.
66. Van Doorslaer K, Tan Q, Xirasagar S, Bandaru S, Gopalan V, Mohamoud Y, et al. The Papillomavirus Episteme: a central resource for papillomavirus sequence data and analysis. Nucleic acids research. 2013;41(Database issue):D571-8.
67. Enerly E, Olofsson C, Nygard M. Monitoring human papillomavirus prevalence in urine samples: a review. Clinical epidemiology. 2013;5:67-79.
68. Meisal R, Rounge TB, Christiansen IK, Eieland AK, Worren MM, Molden TF, et al. HPV Genotyping of Modified General Primer-Amplicons Is More Analytically Sensitive and Specific by Sequencing than by Hybridization. PloS one. 2017;12(1):e0169074.
69. Ruiz R, Silva GR, Menchaca HRM. Focal epithelial hyperplasia. The Lancet. 2014; 384: 173

8 Appendix

8.1 EasyMAG

Appendix 1

Table and chemicals used in easyMAG exstraction. Information taken from procedure.[69]

Name	Storage	Expiration	Supplier	Producer	Article number
Lysis buffer	Room temperature	See bottle	BioMerieux	NucliSENS	280134
Ekstraction buffer 1	Room temperature	See bottle	BioMerieux	NucliSENS	280130
Ekstraction buffer 2	Room temperature	See bottle	BioMerieux	NucliSENS	280131
Ekstraction buffer 3	Refrigerated 2 – 8 ºC	See bottle	BioMerieux	NucliSENS	280132
NucliSENS easyMAG Magnetic Silica	Refrigerated 2 – 8 ºC	See tube / opened 14 days	BioMerieux	NucliSENS	280133
NucliSENS easyMAG Disposables	Room temperature	Unlimited	BioMerieux	NucliSENS	280135
Biohit tips	Room temperature	Unlimited	BioMerieux	NucliSENS	N063783222

8.2 Beta globin

Appendix 2.

Supplies and equipment used for beta globin PCR. Information taken from prossedure.[70]

Equipment not listed; micropipettes and pipette tips with filter.

Name	Supplier	MTU number
Stratagene Mx3005P	Agilent Technologies	55228-15
Stratagene Mx3005P	Agilent Technologies	51803-13
Centrifuge	Heareaus Multifuge IS-R	42217-08
Vortexer	IKA MS 3 digital	43941-08
PCR-plates	Agilent Technologies / Matrix	410088
PCR-striplids Optical Cap	Agilent Technologies / Matrix	401425

8.3 Luminex

Appendix 3.

Supplies and equipment used for Luminex genotyping. Information taken from prosedure.[71]

Name	Supplier	MTU-number
ABI9700 – PCR	Applied Biosystems	51401-11
ABI9700 – PCR	Applied Biosystems	19099-05
Thermomixer Comfort #1	VWR	42684-08
Thermomixer Comfort #2	VWR	42685-08
Luminex	Luminex Corp	50405-08
PCR-plater (Abgene AB-600)	VWR	732-4828
PCR-stripslokk	VWR	732-4830
Biorad flat bottom plates	BioRad	171-025001

Appendix 4.

Reagents used in PCR for Luminex. Information taken from prossedure.[71]

Name	Storage	Supplier	Product number
Deoxynucleoside Thriphosphate (dNTP)	-20 ºC	Roche	1969064
AmpliTaq Gold with GeneAmp 10 x PCR Buffer II (without $MgCl_2$) and separate $MgCl_2$ solution	-20 ºC	Applied Biosystems	N808-0243

Appendix 5.

Reagents used in hybridisation and washing steps. Information taken from prosedure.[71]

Name	Storage	Expiration	Supplier
Hybridising buffer	Room temperature	6 months after production	Medieproduksjon og glassvask, MIKS, Ahus
Colour buffer	Room temperature	6 months after production	Medieproduksjon og glassvask, MIKS, Ahus
Luminex Wash solution	Refrigerated at 4 ºC	1 months after production	Medieproduksjon og glassvask, MIKS, Ahus
TE-Buffer	Refrigerated at 4 ºC	6 months after production	Medieproduksjon og glassvask, MIKS, Ahus

8.4 NGS

Appendix 6.

Equipment and reagents utilized for gelelectrophereses with Invitrogen E-gel.

Equipment: E-gel station, Vortexer, 8-well PCR strips, spinner for strips, GelDoc station.

Table of chemicals and reagents:

Name	Storage	Producer	Supplier	Product number
50 bp ladder	-20 ºC	Invitrogen	Thermo Fischer Scientific	10416-014
Water, Mol Bio grade	Room temperature	5PRIME	VWR	733-0153
E-gel, 4% Agarose	Room temperature	Invitrogen	Thermo Fischer Scientific	G401004

Appendix 7.

Equipment and reagents utilized for Agilent 2100 Bioanalyzer.

Equipment: Agilent 2100 Bioanalyzer, Chipp priming station, Vortexer, Mixer with adapter for chip, timer.

Table of chemicals and reagents:

Name	Storage	Producer	Supplier	Product Number
Agilent High Sensitivity DNA kit	Room temperature/ Refrigerated at 4 ºC	Agilent	Matrix	5067-4626
Agilent DNA1000 kit	Room temperature/ Refrigerated at 4 ºC	Agilent	Matrix	5067-1504

Appendix 8.

Equipment and reagents utilized for library generation prior to NGS. Equipment: Pipettes, Vortexer, Spinner, Magnetic holder for 1,5ml Eppendorf-tubes, Lo-Bind Eppendorf tubes 1,5nl, PCR tubes 200µ in 8-well strips.

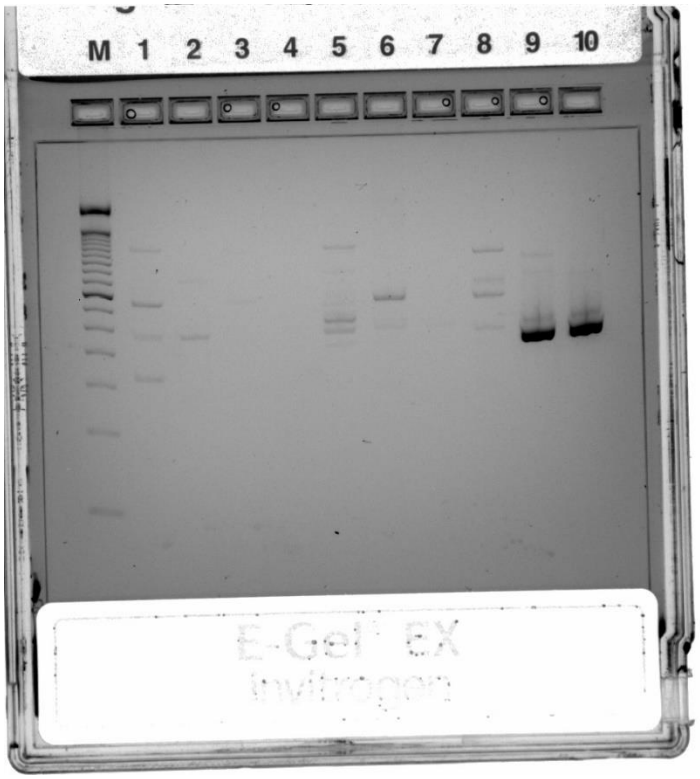
Table of chemicals and reagents.

Name	Storage	Producer	Supplier	Product number
Agencourt AMPure XP beads	2-8 ºC	Beckman Coulter	Nerliens Meszansky	A603881
Water	Room temperature	5PRIME	VWR	733-0153
Etanol, absolute 100%	Room temperature			

Appendix 9

E-gel picture after MGP amplification.

M = ladder, 1 - 8 = samples and 9 and 10 is SiHa and HeLa, respectively.



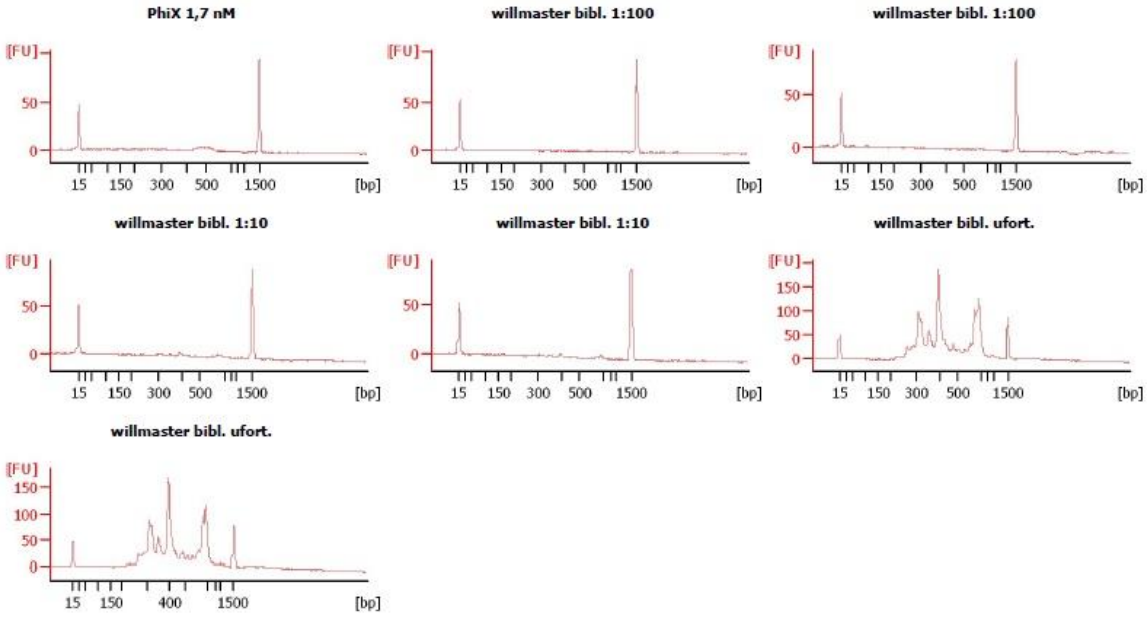
Appendix 10

E-gel after index PCR

M = ladder, 1 - 10 = samples

Appendix 11

Agilent charts. Last two charts are undiluted.



Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway