Diversity of Skin Infections

Ekström, Johanna

2011

Link to publication

Citation for published version (APA):
Ekström, J. (2011). Diversity of Skin Infections Avd för medicinsk mikrobiologi

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain.
• You may freely distribute the URL identifying the publication in the public portal.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
DIVERSITY OF SKIN INFECTIONS

Johanna Ekström

Avdelningen för medicinsk mikrobiologi
Skånes universitetssjukhus
Lunds universitet

AKADEMISK AVHANDLING

som med vedebörligt tillstånd av Medicinska Fakulteten vid Lunds universitet för avläggande av doktorsexamen i medicinsk vetenskap kommer att offentligen försvaras i Patologiska Institutionens föreläsningssal, ingång 78, Skånes universitetssjukhus, Malmö, torsdagen den 17 februari 2011 kl. 09.00

Fakultetsopponent:
Professor Jan Albert
Avdelningen för mikrobiologi, tumör och cellbiologi
Karolinska institutet
**Abstract**

The identification of infectious agents in cancer has been one of the most rewarding endeavors in cancer research. Currently about 20% of the global cancer burden is linked to an infection. A common characteristic of virus-induced cancer is an increased incidence in immunosuppressed patients, presumptively because of impaired host control of virus. Yet non-melanoma skin cancer (NMSC), the cancer that increases most among the immunosuppressed, does not have an established link to infection. NMSC, including squamous cell carcinoma (SCC) and basal cell carcinoma, is the most common cancer among Caucasians. Ultraviolet radiation is an established risk factor.

Human papillomaviruses (HPVs) have been established as the major cause of cervical cancer. Many NMSCs contain one or several cutaneous types of HPV. Exploration of a possible infectious etiology of NMSC requires an unbiased and comprehensive approach for detection of as many infections as possible in the tumor.

We examined NMSCs and other presumably HPV-associated lesions for the presence of unidentified HPV types or other microorganisms, using a combination of multiple displacement amplification (MDA), which amplifies all DNA in a sample without any requirement of prior knowledge of the nucleotide sequence, degenerate "general HPV primers" PCR and high-throughput sequencing. The most common microbial DNA in NMSC was Staphylococcus aureus (S. aureus). We also identified sequences from at least 40 previously not described putative HPV types, of which three novel types (HPV 109, 112 and 114) and an HPV 88 isolate were cloned and completely sequenced. Prevalences and viral loads were investigated in skin and genital samples from different patient groups. S. aureus DNA was more commonly detected in SCC compared to healthy skin (odds ratio, 6.23; 95% confidence interval, 3.10 – 12.53). However, the study design could not determine the causality of the association. HPV 88, 109 and 112 were almost only found in their index patients, whereas HPV114 was found in 1.7% of the female genital samples.

In summary, we find that there is a wide diversity of HPV types in the skin. The association of S. aureus with SCC raises the possibility of general susceptibility to infection in SCC. An association of NMSC with a specific infection remains to be found.

**Key words:** Human papillomavirus, non-melanoma skin cancer, Staphylococcus aureus, multiple displacement amplification, high-throughput sequencing

---

**Diversity of skin infections**
DIVERSITY OF SKIN INFECTIONS

Johanna Ekström

Doctoral Thesis

Malmö 2011

Department of Laboratory Medicine, Medical Microbiology, Lund University, Skåne University Hospital, Malmö, Sweden
FOR MY FAMILY
# Table of Contents

**SUMMARY** ...................................................................................................................... 7

**POPULÄRVETENSKAPLIG SAMMANFATTNING** ................................................................. 8

**LIST OF PAPERS** ............................................................................................................. 10

**ABBREVIATIONS** ........................................................................................................... 11

**INTRODUCTION** ............................................................................................................ 12

- **INFECTIONS AND CANCER** ........................................................................................ 12
- **NON-MELANOMA SKIN CANCER** ................................................................................. 15
  - **The skin** ..................................................................................................................... 15
  - **Skin cancer** ................................................................................................................ 16
  - **Non-melanoma skin cancer** ....................................................................................... 16
    - Risk factors for non-melanoma skin cancer ............................................................... 17
- **PAPILLOMAVIRUSES** .................................................................................................. 19
  - **Generalities** .............................................................................................................. 19
  - **Evolution** .................................................................................................................. 19
  - **Classification** .......................................................................................................... 20
  - **Genomic organization** ............................................................................................. 23
    - **E1** ............................................................................................................................ 23
    - **E2** ............................................................................................................................ 24
    - **E4** ............................................................................................................................ 24
    - **E5** ............................................................................................................................ 24
    - **E6 and E7** ................................................................................................................ 25
    - **L1 and L2** ............................................................................................................... 25
    - Non coding region ...................................................................................................... 26
  - **The papillomavirus life cycle** ................................................................................... 26
- **HPV and mucosal infections** ...................................................................................... 27
  - **Genital infections** ...................................................................................................... 27
  - **Oral infections** .......................................................................................................... 28
- **HPV and cutaneous infections** .................................................................................. 28
  - **Warts** ........................................................................................................................ 28
  - **Psoriasis** .................................................................................................................... 28
  - **Epidermodysplasia verruciformis** ........................................................................... 29
  - **Non-melanoma skin cancer** .................................................................................... 29
  - **Identification of novel human papillomavirus types** ............................................. 30
- **STAPHYLOCOCCUS AUREUS** ..................................................................................... 31
  - **Generalities** .............................................................................................................. 31
  - **Pathogenesis and virulence factors** ........................................................................ 31
  - **Staphylocooccus aureus associated diseases** ....................................................... 32

---

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>17</td>
</tr>
<tr>
<td>19</td>
</tr>
<tr>
<td>19</td>
</tr>
<tr>
<td>19</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>23</td>
</tr>
<tr>
<td>23</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>25</td>
</tr>
<tr>
<td>26</td>
</tr>
<tr>
<td>27</td>
</tr>
<tr>
<td>27</td>
</tr>
<tr>
<td>28</td>
</tr>
<tr>
<td>28</td>
</tr>
<tr>
<td>28</td>
</tr>
<tr>
<td>28</td>
</tr>
<tr>
<td>29</td>
</tr>
<tr>
<td>29</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>31</td>
</tr>
<tr>
<td>31</td>
</tr>
<tr>
<td>31</td>
</tr>
<tr>
<td>32</td>
</tr>
</tbody>
</table>
METHODS FOR DETECTION OF POTENTIAL PATHOGENS

Multiple displacement amplification

Applications of MDA

High-throughput sequencing

High-throughput sequencing using GS FLX

Applications of GS FLX

SUMMARY OF PAPERS

AIMS

Paper I

Paper II

Paper III

Paper IV

MATERIALS AND METHODS

Human papillomavirus-modified multiple displacement amplification

Real-time PCR

Paper I

Paper II

Paper III

Paper IV

RESULTS AND DISCUSSION

Paper I

Paper II

Paper III

Paper IV

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

ACKNOWLEDGEMENTS

REFERENCES
Summary

The identification of infectious agents in cancer has been one of the most rewarding endeavors in cancer research. Currently about 20% of the global cancer burden is linked to an infection. A common characteristic of virus-induced cancer is an increased incidence in immunosuppressed patients, presumably because of impaired host control of virus. Yet non-melanoma skin cancer (NMSC), the cancer that increases most among the immunosuppressed, does not have an established link to infection. NMSC, including squamous cell carcinoma (SCC) and basal cell carcinoma, is the most common cancer among Caucasians. Ultraviolet radiation is an established risk factor.

Human papillomaviruses (HPVs) have been established as the major cause of cervical cancer. Many NMSCs contain one or several cutaneous types of HPV. Exploration of a possible infectious etiology of NMSC requires an unbiased and comprehensive approach for detection of as many infections as possible in the tumor.

We examined NMSCs and other presumably HPV-associated lesions for the presence of unidentified HPV types or other microorganisms, using a combination of multiple displacement amplification (MDA), which amplifies all DNA in a sample without any requirement of prior knowledge of the nucleotide sequence, degenerate “general HPV primers” PCR and high-throughput sequencing. The most common microbial DNA in NMSC was Staphylococcus aureus (S. aureus). We also identified sequences from at least 40 previously not described putative HPV types, of which three novel types (HPV 109, 112 and 114) and an HPV 88 isolate were cloned and completely sequenced. Prevalences and viral loads were investigated in skin and genital samples from different patient groups. S. aureus DNA was more commonly detected in SCC compared to healthy skin (odds ratio, 6.23; 95% confidence interval, 3.10 – 12.53). However, the study design could not determine the causality of the association. HPV 88, 109 and 112 were almost only found in their index patients, whereas HPV114 was found in 1.7% of the female genital samples.

In summary, we find that there is a wide diversity of HPV types in the skin. The association of S. aureus with SCC raises the possibility of general susceptibility to infection in SCC. An association of NMSC with a specific infection remains to be found.
Det är väl känt att infektioner kan orsaka eller medverka till utveckling av vissa former av cancer hos människan. Cirka 20% av den globala cancerincidensen orsakas av infektioner. Patienter med nedsatt immunförsvar har en ökad förekomst av ett flertal virus-orsakade cancerformer, t.ex. livmoderhalscancer som orsakas av humant papillomvirus (HPV) och levercancer som orsakas av Hepatitis B och C. Upptäckten av dessa samband har möjliggjort att man numera kan vaccineras mot dessa sjukdomar. Även för icke-melanom hudcancer (NMSC) kan man påvisa en kraftig ökning hos dessa patienter. Hittills har inget samband mellan infektion och NMSC kunnat säkerställas. NMSC, som huvudsakligen består av diagnoserna skivepitelcancer och basalkelncancer, är den vanligast förekommande cancern bland vithyade, och ultraviolett ljus är en känd riskfaktor.

HPV orsakar ett flertal sjukdomar; utöver livmoderhalscancer t.ex. även kondylom och cancer i munhålan. I NMSC hittas vanligen flera olika HPV-typer, men de påträffas även i frisk hud. För att kunna undersöka en möjlig association mellan en infektion och NMSC krävs därför en mångsidig objektiv metod som kan upptäcka maximalt antalet pathogenar i en tumör.

Vi undersökte NMSC och andra möjliga HPV-relaterade lesioner efter förekomst av nya HPV-typer och andra mikroorganismer med en metod som amplifierar allt DNA i ett prov, samt amplifiering med en teknik som kan påvisa många olika HPV-typer följt av sekvensering med en effektiv sekvenseringsteknik. Vi identifierade sekvenser från minst 40 tidigare ej kända HPV typer. Av dessa klonades och helgenomsekvenserades tre typer, HPV 109, 112 och 114, samt den sedan tidigare kända typen HPV 88. HPV 88 och 109 hittades bägge i skivepitelcancer, HPV 112 upptäcktes i ett kondylom och HP114 i en lätt cellförändring i livmoderhalsen. När vi undersökte skivepitelcancer för förekomst
av nya HPV-typer och andra mikroorganismer fann vi även flera sekvenser som tillhörde bakterien *Staphylococcus aureus* (*S. aureus*).

Vidare undersöktes förekomsten av *S. aureus* samt HPV 88, 109, 112 och 114 i olika hud- samt genitala prover. Vi fann att DNA från *S. aureus* var betydligt vanligare i skivepitelcancer än i frisk hud. HPV 88, 109 och 112 är sällsynta virus, medan HPV 114 återfanns i 1.7% av de genitala proverna från kvinnor.

Sammanfattningsvis visar dessa resultat på att mångfalden av HPV i huden är mycket stor. Det faktum att *S. aureus* var associerat med skivepitelcancer kan visa på en allmän mottaglighet för infektioner av skivepitelcancer och att ett samband mellan skivepitelcancer och en specifik infektion återstår att finna.
This thesis is based on the following papers:

I. Cutaneous human papillomavirus 88: remarkable differences in viral load. 

II. *Staphylococcus aureus* and squamous cell carcinoma of the skin.  
    **Kullander J**, Forslund O, Dillner J.  

III. Three novel papillomaviruses (HPV 109, HPV 112 and HPV 114) and their presence in cutaneous and mucosal samples.  
    **Ekström J**, Forslund O, Dillner J.  

IV. High-throughput sequencing reveals diversity of Human Papillomaviruses in cutaneous lesions.  
    **Ekström J**, Bzhalava D, Svenback D, Forslund O and Dillner J.  
    Manuscript.

Paper I and II were published in the maiden name Kullander.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIDS</td>
<td>Acquired immune deficiency syndrome</td>
</tr>
<tr>
<td>AK</td>
<td>Actinic keratosis</td>
</tr>
<tr>
<td>ASCUS</td>
<td>Atypical cell of undetermined significance</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>BCC</td>
<td>Basal cell carcinoma</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>CCD</td>
<td>Colony collapse disorder</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CIN</td>
<td>Cervical intraepithelial neoplasia</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein Barr virus</td>
</tr>
<tr>
<td>EV</td>
<td>Epidermodysplasia verruciformis</td>
</tr>
<tr>
<td>GS FLX</td>
<td>Genome sequencer FLX</td>
</tr>
<tr>
<td>HCC</td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td>HHV</td>
<td>Human herpes virus</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>HPV</td>
<td>Human papillomavirus</td>
</tr>
<tr>
<td>KA</td>
<td>Keratoachantoma</td>
</tr>
<tr>
<td>kb</td>
<td>Kilo base</td>
</tr>
<tr>
<td>MCC</td>
<td>Merkel cell carcinoma</td>
</tr>
<tr>
<td>MCV</td>
<td>Merkel cell poylomavirus</td>
</tr>
<tr>
<td>MDA</td>
<td>Multiple displacement amplification</td>
</tr>
<tr>
<td>NF-kB</td>
<td>Nuclear factor kappa B</td>
</tr>
<tr>
<td>NMSC</td>
<td>Non-melanoma skin cancer</td>
</tr>
<tr>
<td>OR</td>
<td>Odds ratio</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>pRb</td>
<td>Retinoblastoma protein</td>
</tr>
<tr>
<td>PV</td>
<td>Papillomavirus</td>
</tr>
<tr>
<td>RDA</td>
<td>Representational differences analysis</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RRP</td>
<td>Recurrent respiratory papillomatosi</td>
</tr>
<tr>
<td>SCC</td>
<td>Squamous cell carcinoma</td>
</tr>
<tr>
<td>SISPA</td>
<td>Sequence-independent, single primer amplification</td>
</tr>
<tr>
<td>SK</td>
<td>Seborrhoec keratosis</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolett</td>
</tr>
</tbody>
</table>
Introduction

Infections and cancer

Identification of unknown pathogens is an urgent task of importance for outbreak preparedness as well as for studies of chronic diseases of unknown etiology (11, 61, 201). Currently, about 20% of the global cancer incidence can be linked to infectious agents (270).

The search to identify infectious agents as causative factors for human cancers is difficult due to of several reasons as summarized in the Nobel lecture 2008, by Harald zur Hausen (270):

1) No human cancer arises as the acute consequence of infection. The latency period between primary infection and development of cancer is usually in the range of 15 to 40 years (270).

   This increases the importance of longitudinal epidemiological studies that follow patients over time, preferably over many decades. Infected healthy subjects may not develop cancer until many years later. Also, it is possible that a causative infection could have disappeared long before the cancer has developed (so called hit-and-run mechanism) (10, 229).

2) No synthesis of the infectious agents occurs in the cells, besides some exceptions (270).

   Nowadays, virus detection by isolation in tissue culture is not commonplace, but indeed detection of the infectious agent by other molecular methods could also be complicated when there is no ongoing virus production. E.g., immunohistochemistry might be false negative if it is using antibodies to a protein that is only expressed in the productive phase of the viral life cycle or a polymerase chain reaction (PCR) might be
false negative if it uses primers targeting a piece of deoxyribonucleic acid (DNA) that could be lost if the genome of a virus is integrated.

3) **Most of the infections linked to human cancers are common in the whole human population, while only a proportion develops cancer (270).**

4) **Mutations in host cell genes or within the viral genome are mandatory for malignant conversion (270).**

5) **Mutations caused by chemical and physical carcinogens act synergistically with carcinogenic infectious agents (270).**

6) **Some infectious agents act as indirect carcinogens, without persistence of their genes within their respective cancer cells (270).**

E. g. Helicobacter Pylori that presumably induces cancer by causing a chronic inflammation (164), and is not present in all parts of the carcinogenic tissue. This indirect carcinogenic activity is more difficult to link to cancer than case for HPV where part of the viral DNA is always present in malignant cells and production of oncogenic proteins E6 and E7 occurs (67).

An important issue in investigations of cancer causes is the direction of causality, one of nine criteria required to establish a causal relationship between an exposure and a disease, as formulated by Bradford Hill (123). If an infection is more common in cancers than in controls what came first? Does a particular lesion attract (or activate) a microbe or does the microbe contribute to the development of the lesion?

With appropriate molecular and epidemiological approaches to the issue, several pathogens have been identified to be implicated in different types of cancer. The first human tumor virus, Epstein-Barr virus (EBV), was 1965 linked to Burkitt’s lymphoma (77, 122). Another herpesvirus associated to a human cancer, Kaposi’s sarcoma, is human herpes virus (HHV) -8, also called Kaposi sarcoma virus, discovered by Yang Chang and Patrick Moore in 1994 (45). Chang and Moore also discovered a new polyomavirus in Merkel cell carcinoma, Merkel cell polyomavirus (85). The revelation of a relationship between hepatitis B and hepatocellular carcinoma (HCC) in 1975 (34) led to the development of the first
vaccine shown to be effective in preventing a human cancer. Later also hepatitis C was linked to HCC (9, 146). The second human cancer vaccine, that became available in 2006, prevents the majority of cervical cancer cases caused by human papillomavirus (HPV) (254, 269). The discovery of the oncogenic anogenital HPVs that cause cervical cancer was made by Harald zur Hausen and co-workers and was honoured with the Nobel Prize in Physiology or Medicine 2008 (254, 269). Besides viruses, also other pathogens have been identified as causes of cancer, such as the bacterium, Helicobacter Pylori, the major cause of gastric cancer (164), and parasitic infections, e.g. Schistosoma hematobium associated with bladder cancer (35).

A few cancers occur at an increased incidence in immunosuppressed patients compared to the general population (76, 113). Most of them have an established infectious etiology, e.g. cervical cancer and HPV (254, 269), Burkitts lymphoma and EBV (77, 122) and Kaposi’s sarcoma and HHV-8 (45). Non-melanoma skin cancer is the cancer form that increases most after immunosuppression, but still does not have an established infectious etiology. Interestingly, several additional cancer forms in addition to the above mentioned have also been reported to have some increased incidence in immunosuppression, e.g. cancer of the kidney, leukemia and colon cancer (113), warranting a search for infections also in these cancer forms.
Non-melanoma skin cancer

The skin

The skin is the largest organ in the body. It protects the body against foreign microorganisms and prevents loss of too much water. The skin consists of three layers; epidermis, dermis and the hypodermis (the subcutaneous fatty tissue) (Figure 1). The purpose of the hypodermis is to attach the skin to underlying bone and muscles as well as supply it with blood vessels and nerves. The dermis contains hair follicles, sweat glands, lymphatic vessels and blood vessels. The blood vessels in the dermis provide nourishment and waste removal from its own cells as well as from the epidermis. The epidermis consists of three layers with the bottom layer made up of basal cells. The basal cells divide to form keratinocytes. As new keratinocytes are formed the old ones move up the epithelia, change shape, die and form the outermost layer (Figure 1). Also melanocytes, Langerhans cells and Merkel cells are found in the epidermis.

Figure 1. Schematic picture of a cross section of the skin. Image reprinted with permission from Medscape.com, 2010.
Skin cancer

There are several different types of cancer originating in the skin. Merkel cell carcinoma (MCC), from Merkel cells, is very rare but has the highest mortality rates among skin tumors (37, 196). Risk factors for MCC are ultraviolet (UV) radiation and immunosuppression (169). Since the discovery of Merkel cell polyoma virus (MCV) in MCC (85) several studies have investigated if MCV is associated with MCC (14, 42, 85, 190, 261). MCV DNA is more common in MCC than in healthy controls or other skin lesions (14, 85, 261). Antibodies to MCV are common in the general population, but MCC patients have higher antibody levels (42, 190). Another aggressive skin tumor is malignant melanoma that develops from the melanocytes. It is curable at early stages, but can spread to other parts of the body and be highly mortal (37). Non-melanoma skin cancer (NMSC) is the most common cancer among Caucasians and is described in more detail below (158, 235). Another common neoplasm is keratoachantoma (KA), a skin tumor characterized by a rapid onset most often followed by spontaneous regression within a few months (137, 217). The histological pattern is often difficult to distinguish from squamous cell carcinoma (SCC), a form of NMSC.

Non-melanoma skin cancer

The predominating NMSCs are SCC and basal cell carcinoma (BCC). BCC is more common than SCC, with a ratio of approximately 4:1 (102). The increased incidence of NMSC in Sweden from 1970 to 2008 can be seen in Figure 2. The mortality rate is fortunately very low, but the lesions can cause disfigurement and treatment costs are high (197). Prevention and early detection would reduce both morbidity and cost.

BCC is a slowly growing tumor that has a low degree of malignancy (208, 244). However, BCC can be locally invasive and can cause massive tissue damage (244). The lesion is most frequently found on areas exposed to the sun as the head and neck, followed by the trunk, arms and legs (208). Treatment can be both surgical, e.g. excision of the lesion, cryosurgery or electrodesiccation, and nonsurgical, e.g. radiotherapy (208).

SCC is the second most common skin cancer and, as for BCC, the tumors are most common on sun-exposed areas such as the head (37). In contrast to BCC, which is
believed to arise de novo (i.e. it has no known precursor), a small proportion of SCCs are suggested to arise from actinic keratosis (AK) even though the rate of malignant conversion varies greatly in different studies (148, 162, 163). SCCs are generally slow growing but have the capability to metastasize (49, 66). Nevertheless, most patients have an excellent prognosis and most tumors can be eliminated by electrodesiccation, curettage, excision or cryosurgery, with a low risk of metastasis (7).

Figure 2. Incidence of non-melanoma skin cancer per 100 000 for males and females. Adapted from Cancer incidence in Sweden 2008, published by Socialstyrelsen.

**Risk factors for non-melanoma skin cancer**

The main risk factor for developing NMSC is exposure to ultraviolet (UV) light (159, 198). In the case of SCC, it appears that long-term exposure to UV light is responsible for the progression, but for BCC short term burning episodes seem more important (37, 159). Other risk factors include older age, male sex, fair skin that tan poorly, red, blond or light-brown hair, blue or light-colored eyes and a
number of inherited genetic skin conditions, e.g. epidermodysplasia verruciformis (197).

There are several lines of evidence suggesting that there may also be other preventable risk factors for this cancer. There is a greatly increased incidence of NMSCs, varying from 10- to 250-fold, in patients receiving immunosuppressive therapy because of organ transplantation, (28, 117, 132, 154). The BCC:SCC ratio is reversed in immunocompromised patients, with SCC being the most common (78, 117). Among these patients, the impaired immune surveillance against viral antigens results in increased incidence of skin warts as well as several virus-associated cancers (113) such as Kaposi’s sarcoma, Epstein-Barr virus associated lymphoma and HPV-associated cancers (32, 40, 232). An association of NMSC with HPV has been investigated (reviewed in 118, 177) and will be discussed further in the section “HPV and cutaneous infections”.
Papillomaviruses

Generalities

Papillomaviruses (PVs) are small, nonenveloped viruses that belong to the family Papillomaviridae. Their genome consists of double-stranded circular DNA, about 8 kilo bases (kb), surrounded by an icosahedral capsid. The capsid consists of 72 capsomers, resembling a golf ball when viewed by an electron microscope (27). The first PV was identified by Richard Shope in the 1930s in wild cottontail rabbits (221). Soon thereafter Rous and Beard showed that cottontail rabbit papillomavirus caused skin cancer in domestic rabbits (206, 207). In 1949 the first human PVs (HPVs) were visualized in skin warts by electron microscope (238) and in 1972 Jablonska et al. demonstrated that HPV is responsible for cutaneous lesions in patients with the disease epidermodysplasia verruciformis (129). The association of genital HPVs with cervical cancer was proposed by Harald zur Hausen in the late 1970s (269) and the first major oncogenic HPV (HPV 16) was cloned in 1983 (72). In 2006, after sufficient amounts of epidemiological data had accumulated, the US Food and Drug Administration approved the first vaccine for prevention of cervical cancer (168).

Evolution

The PV genome is very stable and as they utilize the host cell replication machinery they evolve at the same rate as the human genome, mainly through point mutations (31). Papillomaviruses have been detected worldwide and in a wide variety of different species (16, 199, 200, 236). The extensive genomic diversity of Papillomaviridae has taken millions of years to arise. By comparison Human Immunodeficiency Virus (HIV) can diverge to a similar extent during a 10-year infection (31). This difference is due to the high-fidelity proof-reading capacity of the DNA-dependent DNA polymerases used by HPV.

The current hypothesis is that PVs are ancient and existed already at the evolutionary origin of humans. This theory is based on the wide geographic distribution of PVs that cannot be explained by airborne transmission as close physical mucosal or cutaneous contact is required for transmission of PV types (31, 160). Studies of different HPV 16 and 18 isolates indicates co-evolution with humans since the origin of man-kind (Ho, Ong, Chan).
Classification

PVs comprise a large group of viruses with at least 200 completely characterized types divided into 29 genera infecting a wide variety of mammals and birds, including e.g. cows, parrots and dolphins in addition to humans (30). Classification of PVs is based on the sequence of the capsid protein L1 (58). PV types within a genus show less than 60% similarity in the L1 gene to PV types of other genera and different viral species within a genus share between 60 and 70% similarity. A novel PV types has less than 90% similarity to any known type. Subtypes and variants differ 2-10% and maximum 2% respectively from any PV type. The PV types infecting humans are found in five different genera: Alpha-, Beta-, Gamma-, Mu- and Nu-PVs (Figure 3).

The majority of the at least 74 types found in genus Alpha-PV infect the human mucosa, but a few types that infect animals are also found, e.g. Rhesus Macaque and Colobus monkey, in addition to a small number of types infecting human skin (Figure 3). The PV types with mainly cutaneous tissue tropism are found in species 2 (HPV 3, 10, 28, 29, 77, 78, 94 and 117), 4 (HPV 2, 27 and 57) and 8 (HPV 7, 40, 43 and 91) (58). The mucosal types are further divided into high- and low-risk types depending on their ability to cause cancer (173). Most of the high-risk types cluster within species 7 and 9, e.g. HPV 16 and 18 (51, 173), which are the two most common types in cervical cancer (173). In species 10 we find the low-risk types 6 and 11 which are the main etiological agents of condyloma acuminata (genital warts) (112). The two different HPV vaccines approved, Gardasil® (Merck and Co) and Cervarix® (GlaxoSmithKline) both protect against HPV 16 and 18, while Gardasil also protects against HPV 6 and 11 (12).

Genus Beta-PV contains six different species, with at least 42 different types, mainly causing cutaneous lesions, with the exception of a few animal types (Figure 3). Patients with the rare disease epidermodysplasia verruciformis are often infected with HPV types from this genus, especially HPV 5 and 8 from species 1 (183). Different studies have found an association of SCC with HPV type within genus beta species 2 (21, 94).

In the classification study by Bernard et al. (published in May 2010), genus Gamma-PV consists of 16 types divided into 10 different species (Figure 3). Since then at least 10 novel types have been deposited in GenBank (Table 1). In addition
Figure 3. Phylogenetic tree based on the L1 sequences from 189 papillomaviruses. Reprinted from Bernard et al., Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments, in Virology 2010; **401**:70-9 with permission from Elsevier.

A phylogenetic study clustered subgenomic sequences of 97 putative HPV types into this genus (90) suggesting a very large diversity of gamma types. Subgenomic sequences used in that study were generated using the FAP59 and FAP64 primers (92), producing an amplicon of approximately 450 base pairs (bp) in L1. Although, most of the HPV types in genus gamma PV are found in cutaneous infections, 5 types are found in various mucosal samples (Table 1). HPV 101, 103 and 108 are all cloned from cervicovaginal cells (48, 179), HPV 112 from a condyloma acuminate (75) and HPV 116 from a rectal swab (152).
The genera Mu- and Nu-PVs contain only three different types (HPV 1, 63 and 41) detected in cutaneous lesions (Figure 3).

Table 1. Characteristics of HPV types within the genus Gamma.

<table>
<thead>
<tr>
<th>Species</th>
<th>Types</th>
<th>Original lesion</th>
<th>Accession nr</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>Cutaneous wart</td>
<td>X70827</td>
<td>Pfister and Gissman, 1978 (192)</td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>Cutaneous wart</td>
<td>X70829</td>
<td>Egawa et al., 2005 (73)</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>Cutaneous wart</td>
<td>AJ620210</td>
<td>Egawa et al., 1993 (74)</td>
</tr>
<tr>
<td>2</td>
<td>48</td>
<td>SCC of the skin</td>
<td>NC_001690</td>
<td>Müller et al., 1989 (172)</td>
</tr>
<tr>
<td></td>
<td>131</td>
<td>Cutaneous wart</td>
<td>GU117631</td>
<td>Nindle et al. unpubl.</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>AK from an EV patient</td>
<td>U31790</td>
<td>Favre et al., 1989 (81)</td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>Plantar cyst</td>
<td>U31792</td>
<td>Matsukura et al., 1992 (165)</td>
</tr>
<tr>
<td>5</td>
<td>88</td>
<td>NA</td>
<td>NC_010329</td>
<td>Egawa et al., unpubl.</td>
</tr>
<tr>
<td>6</td>
<td>101</td>
<td>CIN III</td>
<td>NC_008189</td>
<td>Chen et al., 2007 (48)</td>
</tr>
<tr>
<td></td>
<td>103</td>
<td>Normal cervicovaginal cells</td>
<td>NC_008188</td>
<td>Chen et al., 2007 (48)</td>
</tr>
<tr>
<td></td>
<td>108</td>
<td>Low-grade cervical lesion</td>
<td>NC_012213</td>
<td>Nobre et al., 2009 (179)</td>
</tr>
<tr>
<td>7</td>
<td>109</td>
<td>SCC of the skin</td>
<td>EU541441</td>
<td>Paper III (75)</td>
</tr>
<tr>
<td></td>
<td>123</td>
<td>NA</td>
<td>GQ845445</td>
<td>Chen et al., unpubl.</td>
</tr>
<tr>
<td></td>
<td>134</td>
<td>Cutaneous wart</td>
<td>GU117634</td>
<td>Nindle et al., unpubl.</td>
</tr>
<tr>
<td></td>
<td>149</td>
<td>Cutaneous wart</td>
<td>GU117629</td>
<td>Nindle et al., unpubl.</td>
</tr>
<tr>
<td>8</td>
<td>112</td>
<td>Condyloma accuminata</td>
<td>EU541442</td>
<td>Paper II (75)</td>
</tr>
<tr>
<td></td>
<td>119</td>
<td>NA</td>
<td>GQ845441</td>
<td>Chen et al., unpubl.</td>
</tr>
<tr>
<td>9</td>
<td>116</td>
<td>Rectal swab</td>
<td>FJ804072</td>
<td>Li et al., 2009 (152)</td>
</tr>
<tr>
<td></td>
<td>129</td>
<td>Cutaneous wart</td>
<td>GU233853</td>
<td>Nindle et al. unpubl.</td>
</tr>
<tr>
<td>10</td>
<td>121</td>
<td>NA</td>
<td>GQ845443</td>
<td>Chen et al., unpubl.</td>
</tr>
<tr>
<td></td>
<td>130</td>
<td>Cutaneous wart</td>
<td>GU117630</td>
<td>Nindle et al. unpubl.</td>
</tr>
<tr>
<td></td>
<td>133</td>
<td>Cutaneous wart</td>
<td>GU117633</td>
<td>Nindle et al. unpubl.</td>
</tr>
<tr>
<td>X</td>
<td>128</td>
<td>Cutaneous wart</td>
<td>GU225708</td>
<td>Nindle et al. unpubl.</td>
</tr>
<tr>
<td>Y</td>
<td>127</td>
<td>Healthy skin</td>
<td>HM011570</td>
<td>Schowalter et al., 2010 (216)</td>
</tr>
<tr>
<td></td>
<td>132</td>
<td>Cutaneous wart</td>
<td>GU117632</td>
<td>Nindle et al., unpubl.</td>
</tr>
<tr>
<td></td>
<td>148</td>
<td>Cutaneous wart</td>
<td>GU129016</td>
<td>Nindle et al., unpubl.</td>
</tr>
</tbody>
</table>

*Abbreviations: SCC= squamous cell carcinoma, AK= actinic keratosis, EV= epidermodysplasia verruciformis, CIN III= cervical intraepithelial neoplasia grade III.

**HPV types not included in the phylogenetic tree in Figure 3, was found in GenBank and classification was determined by comparing their L1 sequences to the L1 sequences of other gamma types.

***NA= not available as there is no article describing the origin of the lesion.

****These types have not yet been designated a species.
Genomic organization

All HPV types have a similar genomic organization. It is divided into three regions, a noncoding, an early and a late region (Figure 4). The non-coding region regulates transcription of the open reading frames (ORFs) and the early region, E1, E2, E4, E5, E6 and E7, encodes the proteins required for viral DNA replication and cellular transformation. The capsid proteins, L1 and L2, are encoded by the late region.

![Figure 4. Schematic picture of the genomic organization of HPV.](image)

_E1_

Only two viral proteins are needed for efficient replication, E1 and E2, as the host cell is providing all other proteins and factors. E1 is the largest of the PV coded proteins and the only one with enzymatic activities; it is an adenosine triphosphate (ATP) dependent DNA helicase (265). The E1 protein binds to the origin of replication as a dihexamer, unwinds the DNA, recruits DNA polymerase α/primase and the single stranded DNA binding protein replication protein A to initiate replication (52, 115, 153, 218). E1 is required both for initiation and elongation of viral DNA synthesis (157).
**E2**

E2 is a multifunctional DNA binding protein that regulates both replication and transcription. It is divided into three distinct regions; a DNA binding domain in the carboxyl terminal end, transcriptional activity in the amino terminal end and in between a flexible hinge region (103). It binds to the DNA as a dimer at E2 protein binding sites in the replication origin and enhances replication by interaction with E1 (68, 147, 166). E2 comes in many different forms depending on alternative ribonucleic acid (RNA) splicing and alternative promoter usage. The full length protein is required for the interaction with E1 (5). The shorter forms compete with the full length protein to regulate transcription and replication (50, 157, 240). This modulation is abrogated in most cervical cancers as the HPV genome is integrated into the human genome at sites that disrupts the E2 gene (237).

E2 also has a role in viral DNA attachment to chromosomes, ensuring episomal maintenance within replicating cells. Segregation of viral episomes depends on tethering to viral chromosomes (150, 225), an event mediated by the cellular protein bromodomain-containing protein 4 (267). It has also been shown that, for at least some HPV types, the E2 protein associates with spindle fibers rather than the chromosome during mitosis (251).

**E4**

The most abundant HPV protein is the fusion protein E1^E4 and different functions has been suggested; e.g. promoting genome amplification and S phase maintenance (263, 264) but results have not been consistent (80), suggesting that HPV type-specific differences exist.

**E5**

The E5 protein is the main transforming protein in bovine PVs (64, 212). For human PVs, the E5 gene is expressed in high-risk HPV types, but many low-risk genital types and cutaneous types lack an E5 ORF or a translation start codon for E5 (98, 211). E5 may play a role in HPV infection (46, 65, 83), but is frequently missing in cervical cancer, due to integration, (167) and is probably not necessary for maintenance of the transformed phenotype.
**E6 and E7**

E6 and E7 are the major oncogenes in humans and they are regularly expressed in HPV associated lesions and cancers. Both genes code for growth-stimulating proteins and especially in high-risk HPV types this can lead to malignant growth.

The most well-known mechanism of the high risk E6 protein is its ability to bind to and degrade the p53 tumor suppressor protein (210, 259). The E6 protein of low risk types do not share this property, but some ability to repress p53 dependent transcription regulatory functions have been shown (246). The E6 proteins of high risk types all have a PDZ binding motif (245). The PDZ motifs are found in many different proteins and the biological consequence of the interaction with E6 is currently investigated. The oncogenic HPV E6 proteins disrupt cellular tight junctions through the degradation of MAGI-1 (143), a protein possessing PDZ motifs.

Interestingly there exists three HPV types, HPV 101, 103 and 108, lacking an E6 gene (48, 179). They were all isolated from cervicovaginal cells, despite being phylogenetically clustered with HPV types normally infecting the skin in the genus Gamma-PVs.

The function of E7 is to maintain viral replication by promoting S-phase entry in otherwise non dividing differentiating cells in the epithelia (108). The tumor suppressor protein pRb, together with relative p130, inhibit cell cycle progression until a cell is ready to divide. Interaction of E7 with pRb disrupts the growth-suppressive pRb-E2F complex, thereby promoting S-phase replication (108). The ability to bind and degrade pRb differs between the HPV types, with the high risk types being more efficient (97). The E7 protein from both high and low risk types bind and degrade p130 thereby inducing viral replication (100).

**L1 and L2**

The two late genes, L1 and L2, code for the structural proteins. The L1 ORF is the most conserved region of HPVs and as mentioned previously used for classification (58). Five units of L1 form the pentameric capsomers of which 72 constitute most of the capsid, while one L2 molecule is found beneath each of the pentamers (41). The L1 proteins can self-assemble into virus-like particles (140), which is the basis for the HPV vaccine (12). L2 is not required for capsid formation, but may facilitate the assembly (87, 114). L2 has also been proposed to
have a role in e.g. the capsid assembly by introducing HPV DNA to the virus particles (268) and to facilitate penetration of the viral genome from endosomes (136).

**Non coding region**

The noncoding region, also called the upstream regulatory region or the long control region is the least conserved region among PVs (58). The non coding region contains the replication origin as well as binding sites to the regulatory protein E2 and also to various cellular transcriptional regulatory factors (20, 44, 107).

**The papillomavirus life cycle**

PVs infect epithelial cells in the skin or the mucosa and they are dependent on the epithelial differentiation for viral replication (Figure 5). Little is known about the initial steps of infection and it has long been thought that the entry occurs in the basal epithelial cells at sites of microlesions (67), but a recent study showed that HPV pseudovirions bind to the basement membrane and become cell-associated at a later time-point (139, 203). The receptor necessary for uptake of the virion is not

**Figure 5.** Overview of the papillomavirus life cycle. After infection, the early proteins E6 and E7 are produced to enable S-phase entry. In the higher epithelial layers production of proteins required for replication is increased. In the upper layers of the epithelium the late capsid proteins L1 and L2 are expressed and the viral capsid is released. Adapted from Doorbar, The papillomavirus life cycle in Journal of Clinical Virology 32S (2005) S7–S15 with permission from Elsevier.
known, but the widely expressed heparin sulphate proteoglycan mediates the initial attachment (104, 135, 139) probably also involving a co-receptor (219). Internalization is not completely understood, and both the clathrin and caveolar mediated endocytic pathways have been described to be involved (38, 54). Inside the cell, the virus starts to replicate: first in a non-productive way with on average one replication round per cell cycle with the viral genome maintained as an episome (125). When an uninfected cell differentiates it leaves the cell cycle, losing the ability to replicate (241). In order to continue replication in differentiating cells the virus stimulates G1 to S-phase progression utilising the transforming proteins E6 and E7 (108). During differentiation the expression of E6 and E7 is upregulated (67), viral DNA is amplified at a high copy number and in the higher epithelial layers the viral capsid proteins, L1 and L2, are synthesized and finally the virions are assembled and released (67, 187).

HPV and mucosal infections

Genital infections

The most important malignancy associated with HPV is cervical cancer, representing one of the most common cancers in women worldwide (86), with more than half a million new cases every year and a yearly mortality of more than 250,000 deaths. The prevalence of HPV in cervical tumors is nearly 100% (254), and the most common high risk HPV types are HPV 16 followed by 18, 31, 33 and 35 (51, 173). An infection with a sexually transmitted HPV in the genital area is highly prevalent among women, especially adolescents (57), but most infections are transient (79, 195). Persistence of infection (124, 213) and high viral load (171, 214, 266) are risk factors for malignant conversion. Other co-factors are smoking (19), multi-parity (174), use of hormonal contraceptives (227) and other sexually transmitted diseases (226, 228).

Condyloma acuminata (genital warts) is a very common sexually transmitted disease caused by low-risk HPV types, mainly HPV 6 and 11 (112). Condyloma is a benign lesion that may regress spontaneously or after treatment (71).

HPV also causes several other genital diseases, notably vulvar, vaginal, anal and penile cancer (4).
Oral infections

Although consumption of alcohol and tobacco are risk factors for oral and pharyngeal cancers (189), there is an increased incidence in young people with no smoking and drinking history (176). Epidemiological studies suggest a strong association of HPV infection and oral cancer development, with HPV 16 being the most prevalent type (144). Globally, HPV is estimated to attribute to 3% of oral cancers and 12% of oropharynx (189).

Recurrent respiratory papillomatosis (RRP) is a rare disease with an estimated incidence of 2 per 100 000 in adults and 4 per 100 000 in children (109). RRP is primarily caused by HPV 6 and 11 (234). Although benign, there is a significant morbidity due to repeated treatment caused by recurrent lesions (109). Extension of the growths into the lower airways indicates a poorer prognosis.

HPV and cutaneous infections

Infections of HPV on the skin are very common (12, 17, 32) and acquisition appears to occur already shortly after birth (18, 110, 126, 256). A large spectrum of cutaneous HPV is commonly detected both on healthy skin (16, 17), in plucked eyebrow samples (39, 55, 194) as well as in different skin lesions (193).

Warts

Common manifestations of cutaneous HPV are skin warts, a benign skin disorder, considered to be no more than a cosmetic nuisance (130). The warts can occur on almost every location on the skin and generally, they resolve spontaneously. Cutaneous HPV types that are frequently associated with warts are type 1, 2, 3, 4, 7, 10, 41 and 57 (105, 130, 184).

Psoriasis

HPV DNA, especially from HPV 5 and 8, can be detected in around 90% of psoriatic lesions (82, 257). There is a significantly higher prevalence of antibodies to HPV 5 and 8 in patients with psoriasis compared to healthy donors (82, 233). A role of HPV in the development of psoriasis is uncertain. A recent study proposed that the presence of the virus is due to the immunosuppression induced in patients receiving phototherapy (209).
Epidermodysplasia verruciformis

The association of HPV with skin cancer was as previously mentioned first demonstrated in patients with epidermodysplasia verruciformis (EV) (129). Patients suffering from the rare hereditary disease EV develop skin lesion in early infancy. The EV lesions are refractory to conventional wart treatment and by the fourth decade, more than half of the patients develop precancerous lesions and invasive NMSC, especially SCC (reviewed in 131, 183). HPV types within the genus Beta PV are commonly found in EV lesions, with HPV 5 being the most prevalent type (182). HPV 5 is also the type that has the strongest association with the malignant conversion to SCC in these patients (185).

Non-melanoma skin cancer

Many studies have investigated the presence of HPV in NMSCs with prevalence rates up to 90 % and generally a little higher in immunosuppressed patients compared to immunocompetent individuals (reviewed in 118, 177). Many cutaneous HPV types are also common on healthy skin (17, 22, 39). Detection of DNA in tumors does not necessarily mean an infection, as it may merely be a viral contamination of the skin surface. Forslund with colleges showed that cleansing of the skin with tape before taking punch biopsies clearly reduces the proportion of HPV positive samples (95): An HPV prevalence of 69% in swab samples taken on top of the lesions from SCC, BCC, AK and seborrhoeic keratosis (SK) was reduced to 12% in the cleansed biopsies (94). The viral load of known HPV types is typically very low in NMSCs (91, 105, 120, 184, 186, 258), normally with less than 1 copy per cell indicating that they are not involved in the growth of the lesion. Genital cancers require continued presence of HPV for continued growth of the cells and they have at least 1 copy per cell of the causative HPV type (171, 214).

Nevertheless, recent studies have found an association of Beta-PV species 2 DNA in biopsies from SCC of the skin compared to adjacent healthy skin (odds ratio (OR), 4.0; 95% confidence interval (CI), 1.3–12.0) (21) and OR, 4.40; 95% CI, 1.92–10.06) (94)) in immunocompetent individuals. An association of Beta-PV and has also been found in the Netherlands (OR, 2.8; 95% CI 1.3–5.8) using hair bulbs from plucked eye brows for detection of HPV (26). Hair bulbs are thought to be a reservoir of PV (39), and eyebrow hairs harbor persistent beta PV (55, 194). Persistence of HPV has also been observed in forehead swab samples (121).
Studies analyzing prevalence of antibodies to HPV have also found an association of Beta-PVs with SCC (26, 84, 138, 239, 255). In two of the studies, an association of HPV and BCC was investigated, but with contradictory results (84, 138). One study also included Gamma-PVs and found an association with SCC (255). There are also serological studies showing no association (13, 43), implicating that additional investigations are required to clarify the possible role of HPV in NMSC.

Identification of novel human papillomavirus types

In the 1970:ies, the identification of novel PV types was a difficult process as conventional cell-culture systems do not allow reproduction of PVs (58). The first HPV types were purified using cesium-chloride gradient centrifugation followed by extraction of DNA, restriction enzyme digestion and gel electrophoresis (105, 106). Closely related types could be detected by Southern blot hybridization (186). These techniques require large amounts of viral DNA. During the last decades there has been a rapid increase in the number of putative HPV types identified by the use of polymerase chain reaction (PCR) and degenerate primers within the conserved L1 ORF (29, 56, 92, 111, 119, 220). At least 150 HPV types have been completely sequenced (30) and new types are continuously found (47, 48, 179, 252, 253). Unfortunately, this technique is limited to detection of viruses with similarity of the L1 gene to previously characterized HPV types and only a small part of the genome is obtained. Despite amplification with PCR cloning and sequencing of complete genomes can be troublesome by the low amounts of virus available in many samples.
Staphylococcus Aureus

Generalities

*S. aureus* is a gram-positive, nonmotile, catalase-positive coccus. The name *staphylococcus*, from the Greek *staphylē* (=a bunch of grapes), was suggested by Sir Alexander Ogston (180) more than 100 years ago. *S. aureus* appears as golden-yellow colonies when grown on blood agar plates, thereby the name aureus (gold in Latin).

The bacterium may occur as a commensal on human skin and in the nose. The carriage rate varies depending on body site, with the highest prevalence of approximately 40% in the nose (141, 262). The prevalence on skin varies between 10 to 20% (141). There are three different types of carriers; persistent carriers, persistent non carriers and intermittent carriers (262). Approximately 10 to 20% are consistently *S. aureus* negative in swabs from the nose, another 10 to 20% are positive at every testing. The intermittent carriers can be positive for several weeks and then negative for comparable periods.

The genome of *S. aureus* consists of a single circular chromosome of 2.8 Mbp and the majority of *S. aureus* strains also carry one or more plasmids ranging from 1 to 60 kbp (170).

Pathogenesis and virulence factors

*S. aureus* is normally not able to cause an infection in an immunocompetent person unless normal barriers have been breached by surgery or a splinter wound (248). Once inside the blood stream access and adherence to host tissue is mediated by surface receptors or adhesions. *S. aureus* produces many different surface proteins involved in cell binding e.g fibronectin-binding proteins (134) and collagen adhesions (243). *S. aureus* also produces different toxins facilitating colonization, one example being α-toxin also called α-hemolysin. α-toxin is a pore-forming toxin that induces a wide array of cellular events in the infected epithelial cells and also in neighboring cells by diffusion of toxins (128, 223). This includes activation of nuclear factor kappa B (NF-kB) and up regulation of various inflammatory cytokines (69, 205). Other virulence factors are involved in the immune response to the bacterium, e.g. protein A (89, 191). Protein A facilitates the survival of *S. aureus* by inhibiting opsonization and thereby phagocytosis.
Staphylococccus aureus associated diseases

As mentioned previously many humans are carriers of *S. aureus* and disease is not developed until *S. aureus* is spread into the blood stream and to other tissues and organs. The most common infections caused by *S. aureus* are different skin infections such as impetigo and carbuncles (70). More severe complications include septic arthritis, and staphylococcal endocarditis (infection of the heart valves) and pneumonia (248). During surgery or dialysis, carriage of *S. aureus* has been identified as a risk factor for the development of infections (141). Other complications are food poisoning and toxic shock syndrome in women (248). A tobacco tar-resistant strain of *S. aureus* has also been suggested to have a carcinogenic potential in the buccal cavity (96).
Methods for detection of potential pathogens

Identification of previously unrecognized pathogenic agents is of great medical interest. PCR using “general” or “degenerate” primers, has been extensively used to discover HPV types, but this method limits the discovery to pathogens with partly conserved DNA. Representational differences analysis (RDA) (156) require no prior knowledge of the nucleic acid sequence. RDA is based on a sequence-independent PCR amplification with subtractive hybridization and is used to identify sequences that is present in cases but not in controls. Both human herpes virus type 8 (45) and torque teno virus (178) were discovered by RDA. Sequence-independent, single primer amplification (SISPA) is another technique enabling nucleic acids of unknown sequence to be amplified, by ligation of primers to blunt end DNA (202). SISPA has been used to discover e.g. hepatitis G virus (155) and two bovine parvovirus’s (8).

Among the techniques that have been used in this thesis two are worth attention, namely multiple displacement amplification (MDA) and high throughput sequencing.

Multiple displacement amplification

The MDA technique is based on rolling circle amplification (88) and has been developed to amplify all DNA in a sample from a very small amount of starting material (60). MDA was initially used to amplify circular DNA (60) but has been extended to work for linear DNA as well (59, 149). The method makes use of a strand displacing DNA polymerase originating from bacteriophage phi29 (33, 99). The strand displacing activity of the enzyme gives it advantages compared to other polymerases, because when the polymerase reaches a primer it displaces it and continues to synthesize DNA (Figure 6). This activity makes the processivity high, the average product length in every primer-extension event being >10 kb (59). The enzyme is very stable and synthesis can carry on for many hours (60).
**Figure 6.** Overview of the multiple displacement amplification reaction: (A) Double stranded circular DNA. (B) Denatured single stranded DNA with annealed primers (−). Phi29 (●) binds to the primers and (C) amplification can start. (D) When Phi29 reaches a downstream primer strand displacement occurs and new primers can anneal to the displaced product. (E) The end product is double stranded repeated copies of the genome. Adapted from Rector et al., A sequence-independent strategy for detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle amplification in Journal of Virology 2004, 78(10):4993-8 with permission from American Society for Microbiology.

**Applications of MDA**

The MDA technique has been used for several different applications, e.g. whole genome amplification of human DNA (25, 59). The ability to increase the quantity of genomic DNA can eliminate technical problems in situations where only limited amounts of DNA is available, e.g. in forensic studies (101), even though the starting DNA has to be of good quality (24). Archival plasma and serum samples for genetic epidemiological studies can also be successfully amplified using MDA (224). MDA has been applied to amplify circular viral genomes and several novel PV types (199, 200, 236, 247, 250) as well as two polyoma viruses (133) have been detected.

**High-throughput sequencing**

The high demand on low-cost and fast sequencing has driven the development of high-throughput sequencing forward and there are now several alternatives available e.g. Applied Biosystems SOLiD system (2), Illumina Genome Analyzer (3) and Genome sequencer (GS) FLX developed by Roche (1). Their sequence methodologies differ and have various advantages. Both SOLID and Genome Analyzer have relatively short read lengths, 75 and 150 bp respectively, but have high throughputs of up to 20-30 Gb per day for SOLID and 6 Gb for Genome Analyzer. The GS FLX has the longest read length with an average of 200 to 300
bp, and for the Titanium update of GS FLX 400 to 600 bp can be achieved. Even though the throughout is lower for GS FLX than for the other technologies, approximately 1 billion bases per day, the long read length makes this technology more suitable for de novo sequencing. The GX FLX has been used in this thesis and is described in more detail below.

High-throughput sequencing using GS FLX

The GS FLX method is based on emulsion-based amplification and pyrosequencing. The hallmark of the technology is the PicoTiterPlate™, which allows a single instrument to produce millions of nucleotide bases per run (161).

Genomic DNA is isolated and fragmented by nebulisation, ligated to adaptors and separated into single strands. The sheared DNA strands are then blunt-ended to allow ligation of adaptors and fragments are captured on beads under conditions that favour one fragment per bead. The beads are captured within the droplets of a PCR-reaction-mixture-in-oil emulsion and PCR amplification occurs within each droplet, resulting in beads each carrying ten million copies of a unique DNA template. Emulsion is broken, DNA strands are denatured, and beads carrying single-stranded DNA clones are deposited into wells of a fibre-optic slide, the PicoTiterPlate™, containing hundreds of thousands of wells, just deep and wide enough for one DNA-containing bead. Smaller beads carrying immobilized enzymes required for pyrophosphate sequencing are deposited in each well. The DNA is sequenced using a technology known as pyrosequencing.

Figure 7. Overview of high-throughput sequencing using GS FLX: (A) Fragmentation of DNA and ligation of adaptors. (B) Emulsion-based PCR amplification. (C) PicoTiterplate with one DNA-containing bead per well. (D) Addition of smaller eads with enzymes required for sequencing. Adapted from Margulies et al., Genome sequencing in microfabricated high-density picolitre reactors in Nature 2005, 415;376-380 with permission from nature Publishing group.
Applications of GS FLX

The GS FLX has been used in a variety of applications, e.g. sequencing of the human genome in less than two months (260), exploring the structural variations in humans by mapping DNA from two individuals against a reference genome (142) and sequencing of the complete genome for several large bacteria (23, 62).

In addition, this high-throughput sequencing technology has been used for several metagenomic studies (15, 53, 151, 188, 204, 231, 249). As the process does not require cloning or preamplification before sample preparation, previously unknown and unculturable organisms can be easily detected. Studies have shown an enormous diversity of microbes in the marine environment (15, 231). Most sequences found in these environmental samples were not present in the current GenBank database, predicting that the composition of microbial communities is greater than the estimation of a few thousand distinct microbes per litre of seawater. The technique has also been used to investigate the microbiome of soil (151, 204).

Using metagenomic analysis with the Genome Sequencer FLX system, researchers can quickly discover known and unknown organisms in outbreaks of infectious diseases, which could have major implications in health care. E.g. three women in Australia who received organ transplants from a single donor died of unknown causes shortly after the transplantation (188). RNA was evaluated from the transplanted organs and high-throughput sequencing identified a novel Arenavirus as the cause. Further, a novel Ebola virus was linked to an outbreak of hemorrhagic fever in Uganda (249). As the virus diverged substantially from other known ebolaviruses the study have important implications for design of effective diagnostics and vaccines. Another important discovery is that of Israeli acute paralysis virus, which is strongly correlated to colony collapse disorder (CCD) in honeybees (53). CCD is responsible for the loss of the adult bee population in colonies (181) and a decrease in healthy honey bee colonies can have severe impact in agricultural commodities that depend on insect pollination. Metagenomic sequences from healthy colonies compared to CCD hives detected candidate pathogens (53). Screening of samples collected from various sites during three years then found the association with Israeli acute paralysis virus.
Summary of papers

Aims

Paper I

**Cutaneous human papillomavirus 88: remarkable differences in viral load.**
To analyze skin biopsies with paired controls for the presence and viral load of HPV 88 that was found at extreme high viral loads in a patient infected with HIV.

Paper II

**Staphylococcus aureus and squamous cell carcinoma of the skin.**
To investigate if squamous cell carcinomas of the skin contain as yet unidentified HPV types or other microorganisms. As most of the detected sequences were from *Staphylococcus aureus*, different skin lesions and controls were investigated for its presence.

Paper III

**Three novel papillomaviruses (HPV 109, HPV 112 and HPV 114) and their presence in cutaneous and mucosal samples.**
To characterize the complete genomes of three novel HPV types, HPV 109, 112 and 114 and analyze for presence and viral load in skin and genital samples.

Paper IV

**High-throughput sequencing reveals diversity of Human Papillomaviruses in cutaneous lesions.**
To use high-throughput sequencing to determine the presence of known and previously unknown HPV types in skin lesions preamplified by degenerate PCR.
Materials and methods

Human papillomavirus-modified multiple displacement amplification

An in-house multiple displacement amplification (MDA) (see the section “Multiple displacement amplification”), that preferentially should amplify human papillomaviruses (HPVs), was developed. In addition to random hexamer primers, primers generic to any HPV were included. Design of the HPV primers was based on an alignment of 72 HPV types, belonging to the genera Alpha-, Beta-, Gamma-, Mu- and Nu-PVs, from the HPV database (Los Alamos, 1997). Two undecamers were chosen in the relatively conserved regions L1 and E1. The random hexamer primers and the “HPV-general” undecamer primers were thiophosphate-modified to be protected from degradation of the 3′–5′ exonuclease proofreading activity of the ϕ29 DNA polymerase.

Real-time PCR

In Paper I to III a single- or triplex real time PCR was designed to analyze presence and viral load of different HPV types. Primers and probes were designed, to only amplify the HPV type of interest. A standard curve was included in each PCR and at least 5 copies were detected in each approved run. To calculate viral load per cell, human DNA was analyzed with real time PCR using primers and probe for the β-globin gene.

Paper I

HPV 88 was isolated, by the use of the in-house HPV-generic MDA, from a human immunodeficiency virus (HIV)-infected man with several extensive squamous cell carcinomas (SCCs) on fingers and on one toe. The MDA product was digested with Bam HI, separated using gel electrophoresis and cloned. HPV 88 was also visualized without preamplification on an ethidium bromide stained gel. The complete genome was sequenced using primer walking.

Patients used for screening of HPV 88 were from three different studies. Skin biopsies were collected from 362 immunocompetent patients attending a hospitalized based case control study in Swedish and Austrian hospitals (94); including 84 SCCs, 147 basal cell carcinomas (BCCs), 58 actinic keratosis (AKs) and 73 seborrhoeic keratosis (SKs). To ensure that viruses detected are from the
lesion and not from surface contamination the skin was stripped by using a tape that was attached and removed five times, followed by the same procedure with a new tape. After tape stripping a biopsy was taken from both the lesion and from adjacent healthy skin. The second series of samples were immunosuppressed patients from Austria (AK, n=5, SCC, n=21). The last patient group of both immunosuppressed (n= 38) and immunocompromised (n=21) were attending a clinic in Australia. They were diagnosed with SCC (n= 21), BCC (n=22) and AK (n= 16). All samples were extracted and DNA quality checked by real-time PCR for the β-globin gene.

Prevalence and copy number of HPV 88 was investigated in eight of the fingers from the HIV patient and in the three different studies using real time PCR. As the lesions also were shown to be positive for HPV 26 (116), viral loads for HPV 26 were investigated.

**Paper II**

Patient samples from the same hospitalized based case control study as in Paper I were used also in Paper II with a few modifications: 82 SCCs, 142 BCCs, 57 AKs and 72 SKs were included. In addition to biopsies from the lesions and healthy skin, swab samples from the top of the lesions and from healthy skin were also used. Swab samples from the top of the lesions were collected using a cotton-tipped swab drawn back and forth over the lesion. From healthy adjacent skin the swab was drawn back and forth 15 times within an area of 5 x 5 cm. The swab was thereafter suspended in 1 ml saline.

The HPV-generic MDA was performed on 83 SCC biopsies. MDA treated samples were digested with Hinc II and analyzed by agarose gel. DNA fragments were excised, cloned and sequenced. To increase the amount of DNA in the biopsy samples, MDA using only random hexamer primers was performed. Ordinary PCR with primers against the nuc gene was used to analyze for presence of *S. aureus* in biopsies and swab samples. SYBR Green was added and a dissociation curve was carried out in a Gene Amp 5700 SDS. Samples containing a detectable product of the same melting point as the positive control *S. aureus* were confirmed by gel electrophoresis.
Odds ratio (OR) and 95% confidence interval (CI) was calculated using LogXact, version 8, and the statistical software R, version 2.7.2.

Paper III

Three novel HPV types were cloned and sequenced in this paper:

HPV 109 was isolated from an SCC of the skin using MDA with primers generic to HPV. The MDA product was digested using Hinc II and run on a gel. A band of approximately 4 kb was excised, purified and cloned. The remainder of the genome was amplified using long PCR with primers designed within the MDA fragment. The complete genome was sequenced using primer walking. HPV 112 was isolated from a condyloma acuminata brush sample using the HPV modified MDA followed by amplification with FAP-PCR. The third type, HPV 114 was found in a CIN I lesion using a modified HPV general priming system, MGP (230), followed by typing with Luminex (215). Luminex is a bead-based multiplex genotyping method that includes type-specific probes and a universal probe. In the case of the HPV 114 index sample, only the universal probe but none of the probes for known HPV types was positive and the amplimer was sequenced. Specific primers were designed for both HPV 112 and 114 and complete genomes were amplified with long PCR. PCR products were separated on E-gel iBase Power System and cloned. The genomes were sequenced using transposons where the transposon EZ-Tn5 <TET-1> was randomly inserted into the clones and the clones were subsequently sequenced bidirectional using primers in the transposon.

Skin samples included were biopsies from different lesions and paired healthy skin from the hospitalized-based study described in Paper I: 52 SCCs, 118 BCCs, 52 AKs and 47 SKs. Genital samples from four different studies were included: From Mozambique 312 cervical cancer biopsies and 271 brush samples from controls were analyzed (175). A second series of samples came from Latvia; 431 brush samples from cervical cancer and 234 brush samples from population based controls were included (222). Also included were 1581 samples from women with atypical cells of undetermined significance (ASCUS) or cervical intraepithelial neoplasia (CIN) I attending the Swedish cervical screening program in Stockholm (63). Finally 27 brush samples from patients with condyloma acuminata were included (242). These samples were previously found to be negative for HPV with PCR using GP5/6+ primers. In total 2856 samples from the genital area and 538
cutaneous samples were analyzed. All patient samples were extracted and checked for presence of human DNA using β-globin PCR.

A triplex real-time PCR with minor groove binding probes was used to investigate presence and viral load of the three viruses in the patient panels.

A phylogenetic tree using MEGA 3.1 was based on an alignment of the L1 sequences of HPV 109, 112 and 114 and representative relatives.

Paper IV

Also in this paper the cutaneous samples from the study from Sweden and Austria described in Paper I was analyzed in addition to keratoacantoma (KA) biopsies from both immunosuppressed and immunocompromised patients that were collected in Norway (93).

The samples were amplified (one by one) with the general primer pair FAP, and then mixed into three pools: 1) fresh frozen biopsies from 37 SCC lesions and 36 AK lesions, 2) fresh frozen biopsies from 92 KA lesions and 3) swab samples from the top of lesion from 86 SCCs and 92 AKs. The samples were run on a gel, purified and thereafter sent for sequencing on the GS FLX platform at KTH, Stockholm, Sweden (see the section “High-throughput sequencing using GS FLX”).

Sequences from the GS FLX platform were filtered to remove human DNA and remaining sequences were assembled and compared to GenBank to classify them as non-HPV or HPV-related.

The program Mr Bayes was used to construct a phylogenetic tree of putative new HPV types with relatives.
Results and discussion

Paper I

The complete genome of HPV 88 was 7 326 bp with a genomic organization typical for PVs. HPV 88 belongs to species 5 within the genus Gamma, with the closest relative being HPV 60 with a similarity of 61%.

All eight SCCs of the HIV infected male were positive for HPV 88, with the highest viral load in the left hand fingers. The SCC that HPV 88 was isolated from contained an exceptionally high viral load of $1.3 \times 10^6$ copies per cell, but also the other fingers of the left hand had high copy numbers varying from approximately 30 to 16 000 copies per cell. The four tumors of the right hand were also positive for HPV 88, but with lower copy number ranging from 0.1 to 1.6 copies per cell. By contrast, HPV 26 were found at high copy numbers on the right hand SCCs (56 to 44 000 copies per cell) and at lower copy number on the left hand SCCs (0.4 to 1.2 copies per cell). Comparing to cervical cancers, which normally carry at least 1 viral genome per cell (171, 214), the HPV viral load of known types in NMSCs are with a few exceptions very low (91, 105, 120, 184, 186, 258). Screening of 809 skin samples, both immunocompetent and immunosuppressed, detected only seven HPV 88 positive specimens. Five of these (SCC =4 and AK =1) were from immunosuppressed patients visiting the same Austrian clinic as the index patient. This result could indicate an association of HPV 88 with SCC in immunosuppressed patients, but as we also had a series of HPV 88 negative immunosuppressed patients from Australia it appears that HPV 88 infection is not generally associated with immunosuppression or with SCC. A possible explanation for the HPV 88 positive patients in the Austrian clinic is that the index patient could have contaminated the waiting room. The same explanation may also be applied to his own hands; HPV 88 of the left hand could have contaminated the right hand and vice versa for HPV 26. The high viral loads of HPV 88 in the SCCs of the left hand and of HPV 26 in the SCCs of right hand suggest that these viruses could be the causal agents for these SCCs.

HPV 88 was cloned using HPV generic MDA, but with the knowledge of its high copy number and the fact that it was possible to visualize HPV 88 directly on an ethidium stained gel without prior amplification, cloning without any amplification would probably have been possible.
Eighty-three SCC biopsies were analyzed using the HPV generic MDA for presence of HPV types and other microorganisms. Strong bands were visualized in 22 of the 83 biopsies, and 28 bands were cloned and sequenced. Half of the sequences (14 of 28) contained sequences from the human genome, another 12 were found to be different plasmids belonging to *S. aureus*. Only one sequence was a new HPV type, and the last one matched with *Escherichia coli*.

As so many of the sequences belonged to *S. aureus* we decided to investigate different skin lesions and swab samples for its presence. The highest prevalence of *S. aureus* was in biopsies from SCC lesions, 29.3%, compared to 1.4% in SK, 12.3% in AK, 7.7% in BCC and 5.7% in biopsies from healthy skin. Presence of *S. aureus* DNA in biopsies was strongly associated with SCC (OR, 6.23; 95% CI, 1.47 – 4.83) when using biopsies from healthy skin as the reference. There is a possibility that the bacterium could merely adhere to protruding growth of the skin so the same analysis was made using the benign protruding growth SK as the reference. These calculations gave an even stronger association of *S. aureus* DNA with SCC (OR, 23.84; 95% CI, 3.69-1004) and also a weak association for AK (OR, 10.01; 95% CI, 1.37-∞). As AK is considered a precursor to SCC (36, 163), this finding is of interest as it suggests an increased colonization of *S. aureus* early in the carcinogenic process. Also when considering the swab samples, the highest prevalence of *S. aureus* was on top of the SCC lesions (31.7%) (OR, 2.67; 95% CI, 1.47-4.83). *S. aureus* is a commensal with reported prevalences on healthy skin varying from 10 to 20% (262) which was confirmed in this study with 15% positive swab samples taken on healthy skin. Some humans appear to be persistent carriers of *S. aureus* and thus we wanted to investigate if SCC subjects could be more susceptible to *S. aureus* infection by comparing the *S. aureus* positivity of SCC lesions to biopsies from healthy skin from the same patient with the result that *S. aureus* was more frequently detected in SCC biopsies (OR, 7.26; 95% CI, 1.38-76.55) than in the matched healthy skin biopsies. Consequently, the association cannot be explained by genetic or disease-induced overall susceptibility to bacterial colonization.

The complete genomes of three novel HPV types were characterized; HPV 109-7346 bp, HPV 112 – 7227 bp and HPV 114 – 8069 bp. All three types had a
genomic organization typical for HPV types except that HPV 114 lacks an E5 ORF which is normally present in genital types. Phylogenetic analysis suggests that HPV 112 constitutes a new species in genus Gamma-PVs and it was most closely related to HPV 65 with 64% similarity. HPV 114 belongs to species alpha-3, with 84% sequence similarity to HPV 84. The classification of HPV 109 is difficult as the bootstrap values in the phylogenetic analysis indicate low reliability of the tree. It clusters with the recently identified HPV types lacking an E6 ORF: HPV 101, 103 and 108, but the closest relative based on similarity in the L1 gene is the Gamma-type HPV 4, with a 65% identity in the L1 gene.

The prevalences of HPV 109 and 112 were very low. HPV 109 was positive in only three specimens; the index SCC, an SK and as a co-infection with HPV 114 in a CIN I lesion. The viral loads in the SK and CIN lesion were low (4 x 10^2 and 5 x 10^6 respectively) as for most cutaneous samples (91, 105, 120, 184, 186, 258), but the index SCC had a relatively high viral load of 11 copies per cell. HPV 112 is a rare virus and was only positive in the index condyloma, despite 2856 samples from the genital area and 538 cutaneous samples analyzed. The high viral load of 842 copies per cell in the index condyloma indicates that its presence in this lesion is probably most likely not a contamination.

HPV 114 was found in 48 of the 2856 genital samples with a viral load ranging from 5 x 10^4 to 240 copies per cell (mean = 11 copies per cell). The highest prevalence of HPV 114, 2.7% (42/1581), was among women with ASCUS and CIN I lesions (ASCUS, 1.8% (10/554) and CIN I, 3.1% (32/1027)). HPV 114 existed as a co-infection in 37 of 42 samples, with up to five different HPV types. The MGP-Luminex missed only one of the HPV 114 single infections indicating adequate sensitivity of detecting unknown HPV types as long as the universal probe can hybridize to the sequence.

**Paper IV**

In total 19 436 reads were obtained from the three pools. The non human sequences (n=17 022) were assembled to 3898 reads of which 2196 reads, 56%, were identified as HPV related. An explanation for the non-HPV sequences found could be that the FAP primers are degenerated and also contains two inosines. In total 43 known HPV types and 67 previously described putative HPV types were detected. Putative types are subgenomic sequences of HPVs, e.g. FA- isolates
amplified using the FAP primers. Most of the HPV types were found in the pool of SCC- and AK-swab samples, 35 types and 48 putative types, compared to 29 types and 19 putative types in the pool of SCC- and AK-biopsies and 26 types and 38 putative types in the pool of KA-biopsies. The majority of the HPV-sequences found were from the genera Beta-PVs (n= 52) and Gamma-PVs (n= 55) which contain the HPV types that mostly are found in cutaneous lesions (58). However three types, HPV 3, 16 and 77, from species alpha were found as well. HPV types within the genus Alpha-PVs are normally found in the mucosa, but HPV types within alpha species 2 and 4, where HPV 3 and 77 are found, more frequently cause cutaneous than mucosal lesions (58). HPV 16 and other mucosal types have also occasionally been found in skin infections (6, 127, 145).

The SCC-, AK- and KA- biopsies had previously been tested using conventional cloning and sequencing after PCR, but still additional sequences from 60 novel putative new types were detected with the use of high-throughput sequencing. The read length for the GS FLX is 200-300 bp and consequently a complete FA-fragment cannot be sequenced from one read. For 11 types more than 400 bp was obtained and for the remaining 49 a sequence from either the 5´ - (n= 26) or the 3´ -end (n=23) was obtained. Most likely many of them belong to the same type but as the length of the partial sequences varied from 84 to 258 bp there were no or too short overlaps to make larger contigs. Most of the novel putative new types were found in this pool (4 with > 400 bp and 30 with a shorter sequence). In the pool with SCC-and AK-biopsies 4 novel putative types with more than 400 bp and 15 additional shorter sequences were detected and in the KA-pool 4 sequences larger than 400bp and 11 with a shorter sequence were found.

A phylogenetic tree based on 34 sequences (>200 bp from the 5´-end) clustered 23 of the novel putative types within the genus Gamma-PVs and 11 within the genus Beta-PVs.
Concluding remarks and future perspectives

The large spectrum of different HPV types found on human skin continuously grows and in this thesis the complete genomes of three novel types, HPV 109, 112 and 114, were characterized. In addition, sequences from at least 37 novel putative types were detected. Many known HPV types have been discovered using degenerate PCR-systems, which require some similarity to previously known types. In paper I to III we tried to overcome this problem by using MDA with HPV generic primers. The HPV generic primers were designed in regions relatively conserved in 72 HPV types from genera Alpha-, Beta-, Gamma-, Mu- and Nu-PVs, and with the addition of random hexamer primers the MDA should allow amplification of any HPV type present at a sufficient amount. In paper I we identified HPV 88 in digital SCCs of an HIV infected man using this HPV modified MDA. HPV 88 is only distantly related to known HPV types and would be difficult to amplify using general HPV detection systems. A disadvantage with the MDA is that the virus probably has to be present at a relatively high copy number to be able to visualize on a gel after amplification. Both HPV 88 and HPV 109, cloned after MDA amplification, had high viral loads.

The characterization of HPV 114 expands the genus Alpha-PV and this infection was not uncommon in a large material of genital lesions. HPV 88 and 112 were classified into the genus Gamma-PV and later a phylogenetic study also clustered HPV 109 within the same genus (30). Genus Gamma-PV appears to be the genus most rapidly growing with at least 26 completely characterized types, including HPV 88, 109 and 112 from paper I and III, 97 putative types, clustered into the genus Gamma-PV in a phylogenetic study (90), and at least 23 of the novel putative HPV sequences found in paper IV. Many of the partially sequenced novel putative types were detected in samples previously tested for HPV, indicating a diversity of HPV greater than what has been revealed with conventional methods.

For future studies we believe that unbiased MDA amplification, avoiding preamplification by PCR, followed by high-throughput sequencing with the GS FLX system could increase the rate at which new HPV types are discovered. With the new high-throughput sequencing techniques it would be possible to sequence a sufficient number of cases and controls separately to provide sufficiently large amounts of epidemiological data to allow rapid establishment of whether some of these new viruses are associated with a major human disease such as NMSC.
S. aureus is a commensal on human skin and in the nose, but is also present in many skin infections. The association of S. aureus DNA with SCC of the skin found in paper II is stronger than what has been found for HPV. However, the study design cannot determine the causality of the association. It would be of interest to investigate if the high prevalence of S. aureus in SCC is due to the ulcerating growth of the SCC or whether the bacterium could be involved in the development of cancer. A possible mechanism for S. aureus to contribute to the cancer development is by production of a chronic inflammation, as in the case of *Helicobacter Pylori* causing gastric cancer and Hepatitis B and C causing liver cancer. A well-known tumor promoter, produced in Staphylococcal inflammation, is nuclear factor-κB. S. aureus has also been suggested to have a carcinogenic potential in the buccal cavity.

Taken together, the work in this thesis has expanded our knowledge of the wide genomic diversity of human papillomaviruses on the skin and has discovered that SCCs of the skin are also associated with S. aureus. Future studies aiming to elucidate a possible infectious etiology of SCC will obviously need to include an unbiased metagenomic sequencing of both viral and bacterial genomes.
Acknowledgements

I wish to express my sincere gratitude to everyone who has been involved in this work and to family and friends who have supported me during these years.

My supervisors,

Joakim Dillner, for showing me the wonderful world of research that I love and for making me feel inspired after almost every Friday meeting.

Ola Forslund, for always taking time for discussions, helping me with laboratory work as well as writing issues.

My co-authors for fruitful collaborations.

Special thanks to:

All the past and present members of our group, especially: Agustin, Anna, Annika, Carina, for helping me with this and that, knowing so much about almost everything, Cecilia, Christina, for all help with laboratory work, Davit, co-author and sequence guru, Hanna, Helena, co-author and for all help with ultracentrifugation, Herman, Janka, Kia, for being helpful in the lab, Kristin, Malin, Natasa, for being such a good travel companion, Vesna and Zoltan. You all make life at work fun, with cake fridays/thursdays and interesting discussions at lunchtime.

My students, Carly, Caroline, Carl and Martin.

Maria Anderberg, for all fun times, in and outside work, for all help with e.g. excel-problems.
Sophia Harlid, my new roommate, for many discussions both regarding science and the life outside.

Helena Person and Anna OlofssonFranzoia, for all administrative help and for always being so kind.

Margareta, Anki, Nasida and Lars-Göran for always being so helpful with preparing laboratory stuff and ordering supplies.

Members of the “tjejmiddag” Dr Tina Hazard, for helping when I came to the lab knowing nothing and for all fun we have had during the years, both at the lab, conferences and outside work. Jenny and Margareta for all delicious food and fun conversations.

My friends scattered all over Sweden (and Germany) for all fun times.

My parents, Lotta och Kjell, for supporting and believing in me and for always being there for me.

My sister and best friend, Malin, for all fun we have together and for being the best god-mother for Isak.

My little family,
   Isak, my wonderful son, for making me so happy every day.
   Mats, the love of my life, for making me believe in myself, without you I would not be where I am today!

This work was supported by the Swedish Cancer Society and the Swedish Research Council, the European Union FP5 project VIRASKIN and the Swedish Center for Metagenomic Sequencing.
References

51


metagenomic survey of microbes in honey bee colony collapse disorder. Science 318:283-7.


54


56


58


166. McBride, A. A., J. C. Byrne, and P. M. Howley. 1989. E2 polypeptides encoded by bovine papillomavirus type 1 form dimers through the common


