DISSERTATION

submitted to the
Combined Faculty of Natural Sciences and Mathematics
of the Ruperto Carola University Heidelberg, Germany

for the degree of
Doctor of Natural Sciences

Presented by
Christy Susan Varghese, Master of Science
born in: Dubai, UAE
Oral examination: 21st of January 2020

Biological Characterization of the Diagnostically Relevant Human Papillomavirus 16 E1C Transcript

Referees: Prof. Dr. Martin Müller

Prof. Dr. Magnus von Knebel Doeberitz

Declarations according to § 8 of the doctoral degree regulations:

I hereby declare that I have written the submitted dissertation myself and in this process have used no other sources or materials than those expressly indicated. I hereby declare that I have not applied to be examined at any other institution, nor have I used the dissertation in any other form at any other institution as an examination paper, nor submitted it to any other faculty as a dissertation.

Heidelberg, 12.09.2019 Christy Susan Varghese

Acknowledgements

First and foremost, I thank God Almighty for giving me the strength and courage to successfully pursue my Ph.D..

I express my fervent gratitude to Dr. Michael Pawlita for giving me the opportunity to conduct my doctoral thesis in his lab. I sincerely thank him for strongly supporting me throughout the course of my thesis, for his strict supervision, for the intense discussions, for the constructive criticisms and for sending me generously to multiple international conferences. Next, I would like to thank Dr. Daniela Höfler for her supervision, for helping me to embark on this project, for all the fruitful discussions and for the continuous positive energy. I also thank both of you for proof-reading this thesis.

I would like to thank Dr. Tim Waterboer for his kind willingness to support me financially in the last year of my doctoral thesis.

I would like to thank Prof. Dr. Martin Müller and Prof. Dr. med. Magnus von Knebel Doeberitz for reviewing this thesis.

I would like to thank all the members of my thesis advisory committee, Prof. Dr. Martin Müller, Prof. Dr. med. Magnus von Knebel Doeberitz and Dr. Massimo Tommasino for the support and encouragement they provided throughout the course of this project.

I take this moment to communicate my profound gratefulness to Dr. Rainer Will and Birgit Kaiser for helping me with the lentiviral transductions, to Dr. Bernd Hessling and Martin Schneider for giving me a hand in mass spectrometry data analysis, to Claudia Tessmer and Natalie Erbe-Hofmann for their support in HPV16 E1C monoclonal antibody generation, to Prof. Dr. Frank Rösl, Dr. Paul F. Lambert, Prof. Dr. Elisabeth Schwarz and Prof. Dr. Martin Müller for their kind gifts of various cell lines and plasmids used in this thesis, to Dr. Martina Niebler for helping me with the proteasome inhibition experiments, to Dr. Daniel Hasche and Rui Cao for teaching me the luciferase assays, to Dr. Maria

Polycarpou-Schwarz for her excellent tips and advices in immunofluorescence assays and finally to Dr. Damir Krunic for his cheerful support during the microscopy sessions.

A special gratitude goes to my wonderful research assistant students, Nikos Avdikos, Yvonne Alt and Maria Boulougouri for their exceptional assistance in the experiments performed in this thesis.

I would like to extend my deep gratitude to all the former and current members of my lab for providing a friendly work environment and for the continuous support that has been given without any hesitation throughout the course of my thesis. I would like to thank in particular Dr. Martina Willhauck-Fleckenstein for the laughs, Dr. Lea Schröder for the good discussions, Andrè Leischwitz for the companionship, Claudia Brandel for teaching me ELISAs and western blots, Birgit Aengeneyndt for lending me a hand in western blots and to my 'Good Samaritan' colleague Dr. Nicole Brenner. I would also like to express my gratefulness to all the other colleagues at DKFZ whose name I have not mentioned here and who have been a pillar of support during my studies.

My heartfelt thanks go to all my friends, in particular, Alex chettan, Pappy chechi, Blessy kutty, Grace kuttan, Sandeep chettan, Cynthi chechi, Unni, Marlo, Alphonsa aunty, Francis, Ramya, Kayden, Joseph annan, Lydia akka, Janice ma, Aaron pa, Phili ma, Nikhil, Sneha, Sheena, Andrew, Dileep, Upasana, Krisztina and my Church family. Thank you for all the fun, flowers, moon, snow and bliss ☺ ☺.

Last but never the least, Thank you Chachen, Amma, Jeeju, Chechi, Ammu and Nivi mon for your unconditional love, support and kindness always. Without you, I would never be where I am today. Love you to Srambical Ushus and back guys! ☺ ☺ ☺

Thank you all!

Contents

List of figures

List of tables

I Summary

Cancer of the uterine cervix (cervical cancer) represents 6.6% of all female cancers with 570,000 new cases as of 2018 and is the fourth most frequent cancer worldwide. Human papillomavirus (HPV), especially type 16 has been recognized as a necessary and causal factor in the pathogenesis of this malignancy. Effective screening for HPV-induced precursor lesions plays a huge role in reducing cervical cancer mortality rate globally and HPV DNA testing has recently become one of the recommended screening tests. However, all the HPV DNA tests available while highly sensitive lack high specificity and positive predictive value for advanced lesions and therefore lead to unnecessary anxiety and overtreatment in many screen-positive women. Consequently, HPV RNA tests have been developed for the intelligible discrimination of low (LSIL) and high grade squamous intraepithelial lesions (HSIL). These tests identified a significant upregulation of E1C transcripts in HSIL and CxCa. Therefore, E1C has been proposed as a diagnostically relevant RNA marker, which is capable of distinguishing LSIL from HSIL.

In this PhD thesis, HPV16 E1C was characterized to define its role in cervical carcinogenesis. Specifically, E1C was overexpressed in four different HPV16-positive cell lines to analyze its effect on different HPV16 transcripts involved in cell cycle progression and transformation. E1^E4 transcripts showed a significant downregulation upon E1C overexpression, whereas E6*I and E7 transcripts remain unchanged. When investigated for the presence of endogenous E1C protein, the HPV16 E1C transcript positive cell lines tested stained negative in western blot. However, transiently overexpressed E1C protein in HEK-293T cells after proteasome inhibition was detectable in western blot and mass spectrometry, which points towards the instability of the protein. HEK-293T cells co-transfected with E1C expression construct and construct expressing a reporter under control of the HPV16 upstream regulatory region (URR) demonstrated URR activation by E1C transcripts even when E1C translation was abolished by mutation hinting that E1C RNA is the biologically active form of E1C. Additional experiments must be performed to study further about the viral and cellular changes driven by E1C RNA and to understand its upregulation in high-grade lesions.

II Zusammenfassung

6.6% aller Karzinome bei Frauen sind Zervixkarzinom. Im Jahr 2018 wurden 570.000 neue Diagnosen gestellt, womit es der vierthäufigste Krebs weltweit ist. Effektives Screening spielt global eine große Rolle für eine Reduktion der Mortalitätsrate und HPV DNA Tests ist einer der empfohlenen Screening tests. Jedoch haben alle bisher verfügbaren HPV DANN Tests schlechte Spezifitäten und schlechte positive prädiktive Werte und führen dadurch zu unnötiger Verängstigung und Überbehandlung von Frauen. Dementsprechend wurden HPV RNA Tests entwickelt, um LSIL und HSIL bzw. CIN Grade zu unterschieden. Ein entwickelter RNA Test basiert auf einer singleplex Sequenz-basierten NASBA Amplifikation und analysiert E6*II/E1^E4 und E1C/L1 Verhältnis und wurde weiterentwickelt zu einer quantitativen Echtzeit-Polymerase-Ketten-Reaktion zur Identifizierung von E6*I/E1^E4-Verhältnis und/oder der E1C Präsenz in HPV16 DANN-positiven Zervixabstrichen. Der ursprüngliche RNA Test hat eine signifikante Hochregulierung von E1C in HSIL und im Zervixkarzinom identifiziert, der weiterentwickelte Test eine fast ausschließliche Präsenz von E1C in ≥CIN3. So wurde E1C als diagnostisch relevanter RNA Marker identifizier, der in der Lage ist, niedrig- und hochgradige Zervixläsionen zu unterscheiden.

In dieser Doktorarbeit wurde HPV16 E1C erstmalig charakterisiert, um dessen Rolle in der zervikalen Karzinogenese zu verstehen. E1C wurde in verschiedenen HPV16-positiven Zelllinien überexprimiert, um einen Effekt auf andere HPV16 Transkripte (E1^E4, E6*I, E7) zu analysieren, die bekannt sind in die Zellzyklusprogression und Transformation involviert zu sein. E1^E4 wurde signifikant herunterreguliert, während E6*I und E7 unverändert blieben nach E1C Überexpression. Jedoch konnte endogenes E1C mittels Western Blot in HPV16-positiven Zelllinien nicht identifiziert werden. Nach der transienten Überexpression von E1C in HEK-293T Zellen konnte E1C erst nach Proteasomeninhibition per Western Blot und Massenspektrometrie nachgewiesen werden, was auf die hohe Instabilität des Proteins hinweist. HEK-293T Zellen wurden neben E1C auch mit HPV 16 URR (vorgelagerte regulatorischer Region) bzw. mit E1C Protein, welches für Translation mutiert wurde, transfiziert. Es wurde eine Aktivierung des URR ausschließlich von E1C identifiziert, was auf eine biologische Aktivität der E1C RNA hinweist. Weitere Experimente müssen in Zukunft durchgeführt werden, um die viralen und zellulären Veränderungen getrieben von E1C RNA zu verstehen und herauszufinden, ob dies Konsequenzen auf die zervikale Karzinogenese hat.

1. Introduction

1.1. Papillomaviruses

Papillomaviruses are DNA viruses that belong to *Papillomaviridae* family, which consists of a diverse group of viruses found in birds, reptiles, marsupials and mammals[1]. More than 150 Human Papillomaviruses (HPVs) have been identified which are divided into five genera Alpha, Beta, Gamma, Mu and Delta[2,222](Figure 1.1.). Although many HPV cause chronic inapparent infections, some can cause serious diseases[3]. Among the five genera, Alpha-papillomaviruses cause mucosal and cutaneous lesions whereas Beta-papillomaviruses cause only cutaneous lesions[4]. Based on their presence as single infections in cervical cancer cases mucosal Alpha-papillomaviruses are further classified into low-risk and high-risk types[4]. Low-risk HPV cause only benign lesions whereas high-risk HPV cause pre-malignant and malignant lesions and immortalize human keratinocytes[4].

1.2. Human Papillomaviruses (HPV)

1.2.1. Virus structure

Papillomaviruses are non-enveloped viruses with an icosahedral capsid of 50-60nm diameter (Figure 1.2.). They comprise of 8 kb genomes which are double-stranded and can encode for eight or nine open reading frames (ORF)[3]. Owing to the use of multiple promoters and complex splicing patterns, the number of proteins encoded by the small HPV genome is much greater[5]. The virus capsid consists of L1 protein, which are arranged into capsomeres. Each capsomere has a beta-jelly core and is made up of 5 L1 molecules. There are a total of 72 capsomeres on the virus capsid and therefore a total of 360 L1 molecules[6,7]. The virus capsid also contains variable number of L2 molecules. Only the N-terminal 120 amino acids (aa) of L2 molecules are exposed on the surface of

the virion[8,9]. During infection, L2 binds to the extracellular matrix and is cleaved by a proprotein convertase, furin[10].

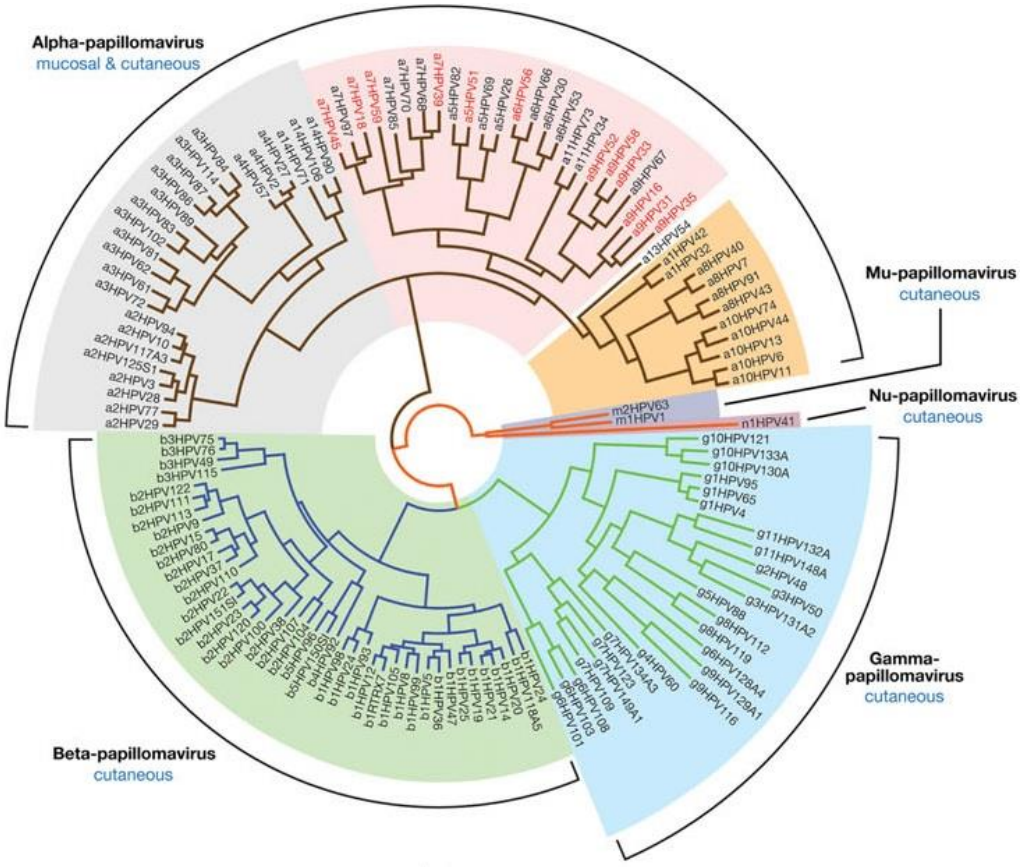


Figure 1.1. **Evolutionary tree of HPV**. Modified from[3]. The HPVs are categorised into five genera with the largest groups belong to the Alpha and the Beta/Gamma genera. The Alpha low-risk cutaneous HPVs are in the gray region, low-risk mucosal in the orange region and high-risk in the pink region. The high-risk types categorised as “human carcinogens” are in red text in the pink region and the other Alpha high-risk types are “probable” or “possible” carcinogens.

1.2.2. Genome organization

The ORF of the HPV genome can be classified to early or late categories[11]. The early ORF consist of E1, E2, E4, E5, E6 and E7 and late ORF consist of L1 and L2 (Figure 1.3.). E1, E2, L1 and L2 genes are well-conserved in papillomaviruses[12]. E1 is involved

in viral DNA replication and amplification and encodes a virus-specific DNA helicase. E2 protein plays important roles in viral transcription, replication and genome partitioning and has binding sites on both viral and cellular genomes[3]. E4 protein helps viruses to escape from epithelial surface[13]. E5 protein is involved in immune evasion and genome amplification[14].

Figure 1.2. **Electron micrograph of HPV particles.** Modified from[86]. A negatively stained transmission electron micrograph showing HPV particles of 55 nm in diameter.

E6 and E7 proteins drive cell cycle entry in all HPV types and allow genome amplification and are therefore called as the oncoproteins. The Upstream Regulatory Region (URR), as the name suggests, consists of regulatory elements such as promoters, origin of replication, transcription factor binding sites etc. and is located between the end of L1 and the start of early region[3]. L1 and L2 aid orchestration of proper genome packaging[6].

1.2.3. HPV life cycle

HPV replication is tightly linked to differentiation of the epithelium it infects[15]. HPV accesses dividing basal epithelial cells through microabrasions or wounds and uses heparin sulphate proteoglycans for cell attachment[16,17] (Figure 1.4.). In some cases, α6-integrin is used as the virus receptor[18,19] and virus entry is aided via either calveolar or clathrin-mediated endocytosis[16,20]. The viral genome is transported to the cell nucleus followed by virus entry. Viral genome is replicated to approximately 50 copies during the division of infected cell and attaches to host chromosomes, thereby gets maintained in the daughter cells upon segregation of the divided cell.

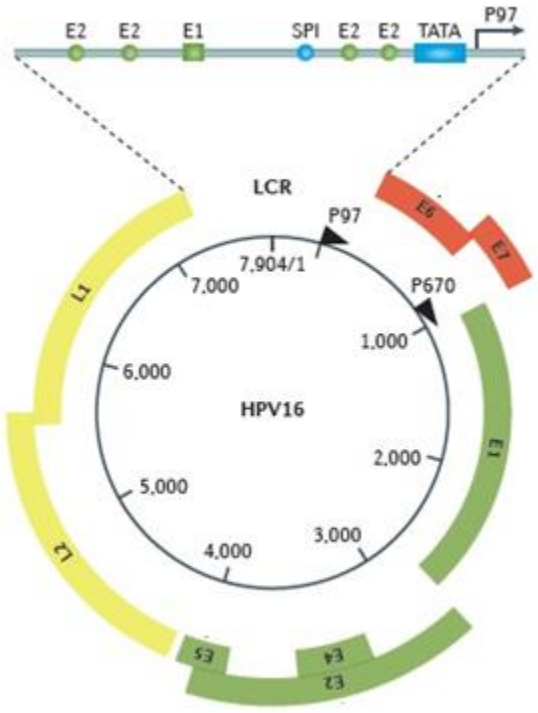


Figure 1.3. **HPV16 genome organization.** Modified from[86]. This is the typical genome organization of all high-risk Alpha HPV types. The two arrowheads represent the early and late promoters, P_{97} and P_{670} respectively. E1, E2, E4 and E5 ORF are shown in green and E6 and E7 ORF in red. The late ORF, L1and L2 are shown in yellow. E1, E2 and cellular SP1 transcription factor binding sites on the long control region (LCR, also known as URR) are shown.

Normally, during differentiation epithelial cells stop to divide and terminally differentiate[15]. However, cell division is re-activated, apoptosis is inhibited and differentiation is

abrogated in HPV-infected epithelial cells by the concerted action of the oncoproteins (E6 and E7) and tumor suppressors (p53 and pRb)[21]. Upon division of the infected cell, the daughter cells move to the suprabasal differentiating cell layers. The proliferation of the infected cells is promoted by the restricted expression of E5, E6 and E7. Late viral gene expression, viral DNA replication and formation of capsid proteins such as L1 and L2 are initiated in the infected cells of the suprabasal layers[22]. Assembly and release of the infectious virions take place in the upper layers of the epidermis. It is thought that E4 protein amyloid fibers might play a role in virion release by disrupting the keratin structure of the upper epithelial layers[3] (Figure 1.4.).

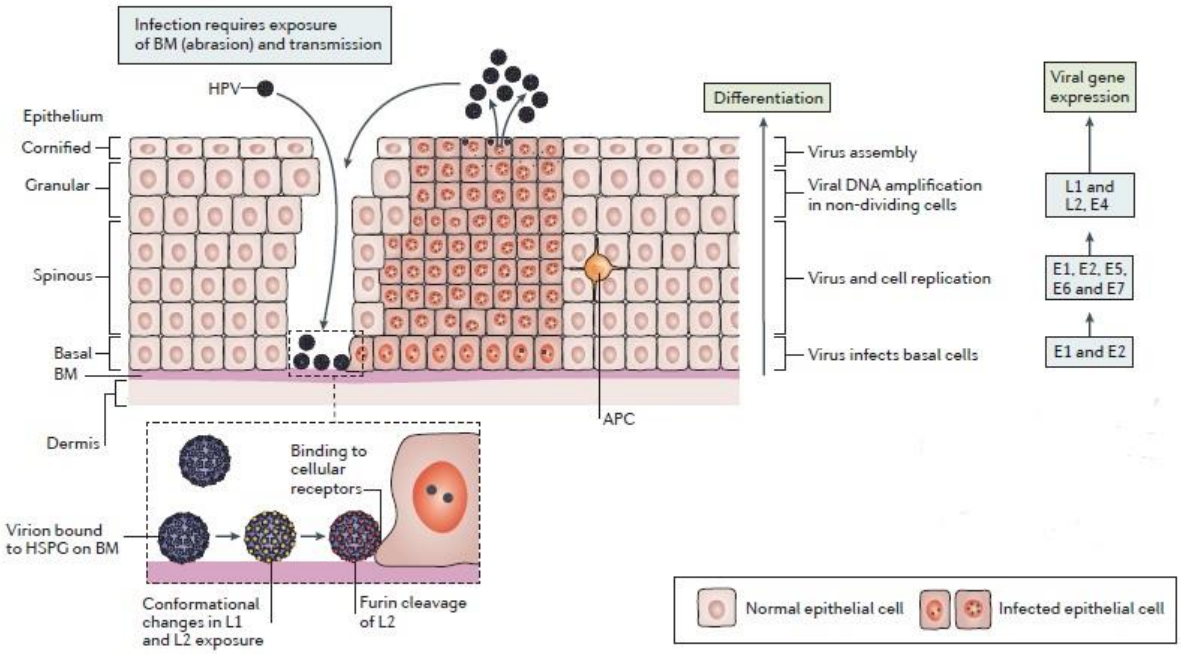


Figure 1.4. **The life cycle of HPV.** Modified from[145]. Abrasion of the basement membrane (BM) gives HPV, access to the basal keratinocytes. During infection, HPV binds to heparin sulphate proteoglycans (HSPGs) on the BM through L1 which leads to conformational changes in L1 and exposes L2. Extracellular furin cleaves amino terminus of L2 and virus entry is initiated. The viral genome replicates to produce ~50 HPV episomes and divides between progeny upon cell division. E6 and E7 stimulates the continued proliferation and E1 and E2 drive viral genome replication. E4, L1 and L2 are expressed in the upper epithelial layers upon differentiation of the infected cells and orchestrate viral genome packaging. With the aid of E4, virions are released by the disintegration of the cytokeratin filaments. No cell death/viraemia /inflammation is caused during the completion of life cycle, thus evading local immune responses. APC – Antigen-Presenting Cell.

1.2.4. HPV gene expression

1.2.4.1. Transcription

HPV transcription is initiated at least from two promoters and is polycistronic. It yields multiple mRNAs with several ORF (Figure 1.5.). For HPV16, the two well characterized early and late viral promoters respectively are P_{97} and P_{670}[15]. P_{97} is located in the URR and is constitutively active throughout the virus replication cycle[23-26].

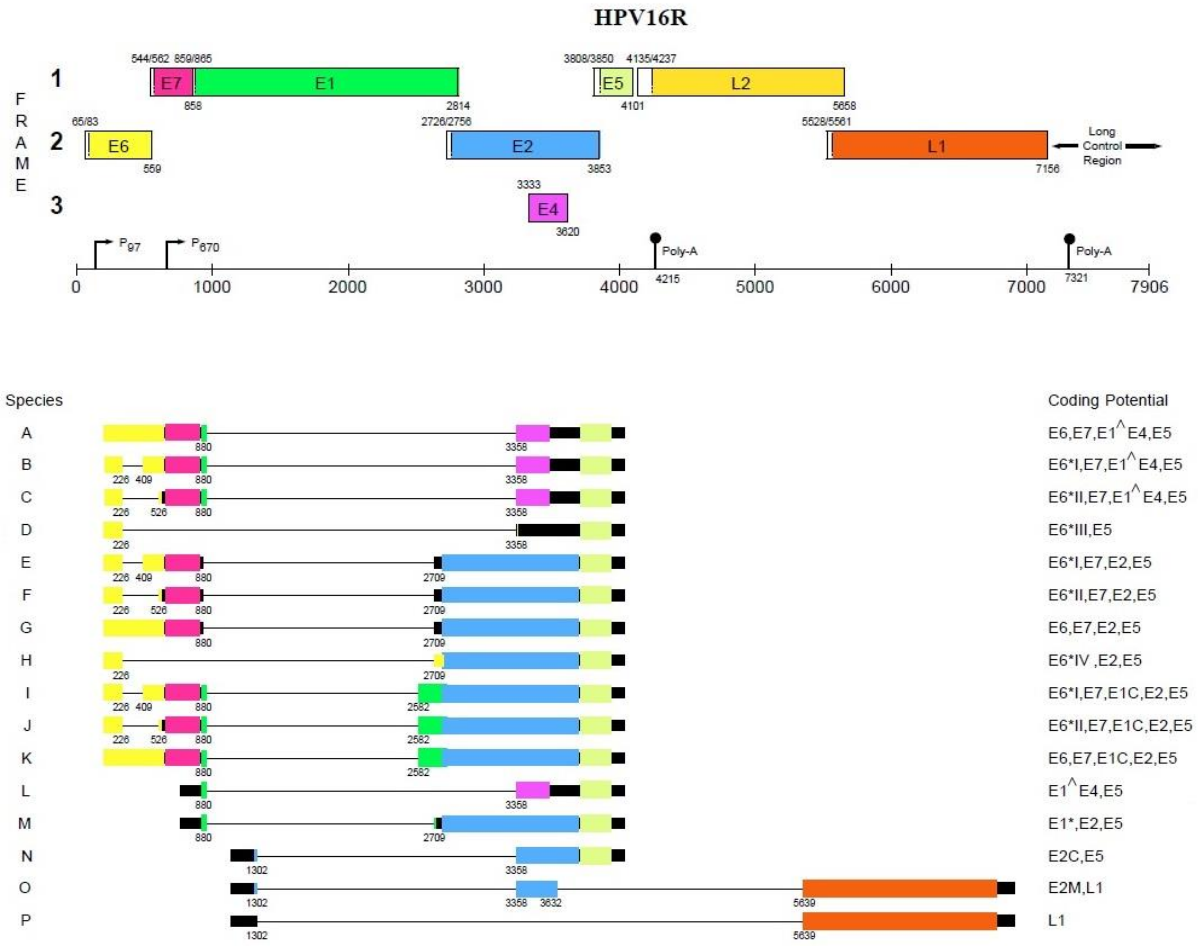


Figure 1.5. **Transcription map of HPV16.** Modified from[204]. Colored rectangles at the top of the figure denote the ORF. The numbers on the rectangles denote the following – first number at the upper left end – start nt of the ORF, second number at the upper left end – first nt of the first start codon (also marked by dotted line inside the rectangle), number at the lower right end - last nt of the last stop codon. Various spliced mRNA species of HPV16 and their coding potential

are shown. Black rectangles are untranslated exons and black lines are intervening introns. Splice donor and acceptor nt positions are indicated below each spliced mRNA species[146,147].

The late differentiation-regulated promoter, P_{670} is located within the E7 ORF[27]. P_{97} is controlled by E2 protein and many cellular transcription factors[28]. There are four E2 binding sites in the promoter, E2BS1, E2BS2, E2BS3 and E2BS4. E2BS1 is located in the distal promoter region, E2BS2 is located upstream to the keratinocytes enhancer and E2BS3 and E2BS4 are near to the TATA box. The P_{97} promoter is transcriptionally repressed by E2 in a normal infection so that E6 and E7 levels remain low and the epithelial cells do not progress to tumor. A combination of E2, TATA box and cellular transcriptional repressor protein, Sp1 seem to be involved in the repression of P_{97} promoter[15]. During epithelial differentiation, the proteins bound to the URR and the strength of their binding are likely to regulate the transcriptional changes required for the virus life cycle completion[29]. In HPV16 many transcription initiation sites have been shown to be regulated by P_{670}[27,30,31]. Studies have also shown that keratinocyte enhancer present in the URR could contribute to P_{670} regulation[15]. Other promoters have also reported to be present in HPV genomes. Like BPV-1, HPV might also have a similar set of promoters in the early region[32]. Presence of a late promoter at the start of the E4 ORF has been verified by several studies[30,33,34].

1.2.4.2. Splicing

Most HPV mRNAs are produced by constitutive and alternative splicing. As transcription progresses, if every intron in the primary RNA transcript is removed, then it is called constitutive splicing. Alternative splicing occurs when exons or introns are skipped. In HPV, the first and second splice events are generated from the splice donor sites in E6 and E4 ORFs, which use different splice acceptor sites in the E6, E7, E4 and L1 ORF[15]. Regulation of splicing occurs through cis-acting splicing signals and RNA/protein complexes knows as spliceosomes[35,36]. Some members of the serine-arginine (SR) –rich protein family regulate viral RNA splicing[37]. SR proteins such as SRp20, SC35 and SF2/ASF are shown to increase in the mid to upper layers of high-risk HPV infected

cervical epithelia[38]. During infection, E2 controls the expression of these SR proteins. During the late phase of virus replication cycle, SF2/ASF is upregulated via E2, which is essential for the expression of L1 and L2[39,40].

1.2.4.3. RNA stability and translation

HPV gene regulation is controlled post-transcriptionally. Expression of L1 and L2 can be regulated through the stability of their mRNAs. In HPV16, a regulatory element is located at the end of L1 ORF which extends to the late 3' untranslated region (UTR)[41-51]. This element blocks L1 and L2 expression in lower epithelial cells by making any transcribed mRNAs unstable[52]. This helps the virus in preventing the cells from triggering any immune response as the capsid proteins are highly immunogenic. The regulatory element of HPV16 has been shown to bind to a cellular RNA stability regulator, HuR. HuR overexpression in undifferentiated HPV-infected epithelial cells increases L1 expression and siRNA depletion of HuR decreases L1 expression, which underlies the significance of cellular factors in regulating viral late gene expression[49,53].

Few studies have shown that virus mRNA translation is regulated for efficient virus replication. A regulatory element which regulates stability of late mRNAs[54] has been characterized for HPV1. This hints to the possibility that mRNA stability and translation could be working hand-in-hand for the regulation of various genes. Also, there could be other HPV regulatory elements that are involved in similar processes[55]. Codon usage also controls translation efficiency. HPV mRNAs use rare codons, perhaps a mechanism for efficient translation in the background of host cell translation[56,57].

1.2.4.4. HPV16 transcripts and proteins

E1 full length (E1fl) or simply E1protein is an ATP-dependent DNA helicase which is essential for replication and amplification of viral genome. It is encoded from E1 ORF and is 600-650 aa long. E1 binds to the ori with the help of E2 protein[202]. A putative E1C protein of 82 aa length[147], first 5 aa from the amino terminus of E1 and the last 77 aa from

the carboxy terminus of E1, could be translated from 880^2582 spliced mRNA (Figure 1.5., species I-K).

E2 proteins are sequence specific DNA binding proteins, which are associated with transcription and replication of the viral genome. It can be encoded from 226^2709 (Figure 1.5., species H), 880^2709 (Figure 1.5., species E-G) and 880^2582 (Figure 1.5., species I-K) spliced mRNA species[64,147]. As a transcription factor, E2 regulates the viral promoter P97 and thus regulates the expression of oncogenes, E6 and E7. E2 activates the transcription of oncogenes at low levels and represses their transcription at high levels[65]. E2C, which is a shorter form of E2 is translated from 1302^3358 mRNA (species N) and probably inhibits the function of fl E2[66]. A putative E2M protein could be expressed from species O transcript.

The E1^E4 protein interacts with proteins of the keratin cytoskeleton and helps in virus assembly and virus release[67]. It is translated from 880^3358 spliced mRNA species (Figure 1.5., species A-C, L)[147].

E5 protein could be involved in potentiation of EGF receptor signalling pathway. Another phenotype, which has been attributed to E5, is assisting cell immortalization and transformation by E6 and E7[203]. The unspliced E2/E5 transcript encodes for E5 protein (Figure 1.5., species E-G, M)[147].

The fl E6 ORF encodes for E6 protein (Figure 1.5., species A, G, K) and E7 is translated from E6*I mRNA[73-75] (Figure 1.5., species B, E, I) or from the fl E6/E7 mRNA[23,76,77,147]. E6 and E7 proteins of high-risk HPVs bind and degrade cellular tumor suppressor proteins such as p53 and pRb respectively and thereby promote transformation.

Either unspliced transcripts or transcripts spliced between nt 1302 and nt 5639 (Figure 1.5., species P)[34] or nt 3632 (Figure 1.5., species O)[147] encode for L1 protein. These transcripts are initiated by the late promoter P_{670}. L2 protein is expressed only from unspliced late transcripts. L1 and L2 proteins are expressed during the late stages of the productive infection. However, detection of L1 and L2 RNA in lower epithelial layers[79,80] indicates that their expression is also regulated post-transcriptionally[81,82].

1.2.4.4.1. Domain structure of E1 and putative E1C proteins

E1 protein consists of three functional segments[58]. A regulatory region at the N-terminus, an origin-binding domain at the centre and an enzymatic domain (also helicase domain) at the C-terminus (Figure 1.6.). The regulatory region is necessary for viral DNA replication and is approximately 200 aa long. It contains various sequence motifs such as nuclear localization signal and nuclear export signal[58]. The origin-binding domain is also known as the DNA-binding domain and is approximately 150 aa long. It is capable of recognizing certain sites of ori[58]. The helicase domain is roughly 300 aa long which has ATPase activity and unwinds viral DNA during replication. It consists of three functional regions: the minimal oligomerization domain, the AAA+ ATP-binding module and the C-terminal brace[58]. Oligomerization domain is involved in the self-assembly of E1 into hexamers during viral DNA replication. AAA+ is an ATPase and therefore as the name suggests, AAA+ ATP-binding module is involved in ATP-binding and hydrolysis. C-terminal brace is a short flexible region. It also helps in assembling E1 to hexamers and its stabilization[58].

Figure 1.6. **Domain structure of BPV1 E1.** Modified from[58]. Light purple box is the N-terminal regulatory region and the nuclear localization signal (NLS) is shown. DNA binding domain (DBD), Oligomerization domain (O), AAA+ ATP-binding module or ATPase domain and Brace are shown in dark purple boxes. The numbers marked below are aa.

As mentioned earlier, the putative E1C protein comprises of 5 aa from the N-terminus and 77 aa from the C-terminus of full length E1 protein. The E1 protein of HPV16 is 649 aa long[204], which means the C-terminus of putative HPV16 E1C protein will have 573 – 649 aa region of E1 protein. Therefore, last few aa of the ATPase domain and brace

region are present in the E1C protein, which could mean that E1C protein can probably hold the function of brace region.

1.3. Cervical Cancer (CxCa)

1.3.1. Global burden of CxCa

CxCa is the fourth most common cancer in women and the seventh most common cancer overall. 528,000 new cases and 266,000 deaths from CxCa had been estimated worldwide in 2012. Around 85% of CxCas occur in the less developed regions, where CxCa accounts for almost 12% of all cancers in females. High-risk regions include eastern Africa, Melanesia, southern Africa and central Africa. Incidence rates are lowest in Australia, New Zealand and western Asia[83].

High-risk HPVs are sexually transmitted and can cause CxCa if they persist. HPV16 is the predominant type causing 60% of the invasive CxCas worldwide followed by HPV18 with 15%[84]. The histology system used for diagnosis is the Cervical Intraepithelial Neoplasia (CIN) scale. Mild dysplasia is categorised to CIN1, moderate dysplasia to CIN2 and severe dysplasia and carcinoma *in situ* to CIN3. Majority of CIN1 (90%) regresses spontaneously. However only a minority of CIN3 progress to cancer in a lifetime[85]. The Lower Anogenital Squamous Terminology (LAST) is used[156] in the United States for diagnosis. According to this system, squamous precursor lesions are grouped into low-grade squamous intraepithelial lesions (LSILs) or high-grade squamous intraepithelial lesions (HSILs), and CIN2 category is placed in either one of them based on $p16^{INK4A}$ staining. $p16^{INK4A}$ is a marker related to cervical cancer progression. However, LAST is unspecific because many HPV infections stain positive for $p16^{INK4A}$ without a precancer[86].

1.3.2. Cofactors for HPV persistence

HPV infections are generally asymptomatic[86]. More than 90% of the infections are undetectable within 5-7 years[88] and persistence past 2 years is not common[87]. The virus

is either cleared or suppressed to a low level by cell-mediated immunity. The three major classes of cofactors that determine the persistence of an acquired cervical HPV infection are viral, host and behavioural variables. Viral variables include genetic differences in HPV. For e.g., infection by HPV16 has a precancer risk of one order of magnitude higher than other carcinogenic types such as HPV51, HPV56 and HPV59[86]. Host variables include host immune response, for e.g., the influence of HIV infection[89]. Behavioural variables include smoking, multiparity and long-term use of hormonal contraceptives[90-92].

The typical transition time from HPV infection to invasive CxCa lasts decades. Apart from age and time, no other strong predictors for invasion have been proven yet. However, certain epidemiological and laboratory studies have shown the role of female steroid hormones as an aetiological cofactor for CxCa[86,91,93-97].

1.3.3. Effects of HPV infection on host cell

CIN1 lesions support the complete life cycle of HPV[98]. The deregulated expression of E6 and E7 from CIN2 onwards promotes cancer progression (Figure 1.7.). E7 proteins stimulate host genome instability via binding and degrading multiple members of pRb protein family whereas E6 proteins bind and efficiently degrade p53 and interrupt the DNA repair pathway[3]. It is not clear what drives the deregulation of the viral episome; however, hormonal changes and/or epigenetic modifications of the infected cell may be the drivers[99-102]. DNA methyltransferases are another targets of E6 and E7 proteins causing hypermethylation of CpG islands of the cellular genome which in turn leads to the possible silencing of cellular tumor suppressor genes[103,104]. Interestingly, a mutational signature involving APOBEC (apolipoprotein B mRNA editing enzyme catalytic subunit) has been reported in HPV associated squamous cell cancers. Increase in E6 and E7 can increase cellular APOBEC levels. It has been suggested that APOBEC-mediated mutations might lead to the generation of virion variants, providing an evolutionary benefit for the virus[105].

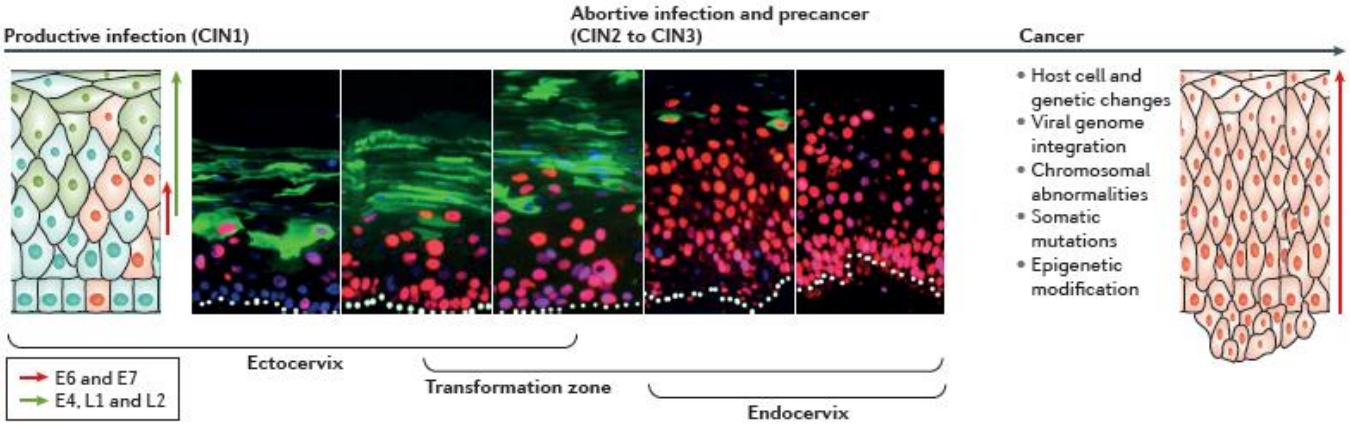


Figure 1.7. **Molecular events during the progression of CxCa.** Modified from[86]. Cell cycle marker minichromosome maintenance protein complex (red staining) is almost not present in the first immunofluorescence image.This is an example of a productive lesion and here, E6 and E7 are essential for genome amplification in the mid-epithelial layers. E4, L1 and L2 expression are mainly in the upper epithelial layers, but can be spreaded (green staining). Such lesions are labelled as CIN1 pathologically. In high-grade lesions (or abortive infections) E6 and E7 are further increased (red staining) and L1 and L2 are restricted (green staining). Such lesions would be labelled as CIN2 or CIN3 pathologically and correspond to precancers.

1.3.4. Host response to high-risk HPV infection

There is no sign of viraemia during HPV infection. That means, neither inflammation nor cell death occurs. In addition, no whole virus is found in the blood. There is no release of pro-inflammatory cytokines during the infection and thereby Langerhans cells are not recruited or activated. HPV16 E6 and E7 downregulate CC-chemokine ligand 20[110] and E-cadherin[111]which are necessary for the function of Langerhans cells. Therefore, antigen-specific responses are also not activated and effector cells, which can target and kill infected epithelial cells, are not recruited[86]. Normally, keratinocytes respond with an antiviral state in response to viral infection[106]. Antiviral state constitutes of the expression of pattern recognition receptors by keratinocytes, which involves initiation of certain signal transduction cascades and activation of cellular transcription factors. These transcription factors are responsible for the generation of interferons and inflammatory cytokines[107].

However, high-risk E6 and E7 directly interfere with the components of these cascades[108] and inhibit the immune response[109].

1.3.5. Diagnosis, Screening and Prevention

The two major cervical cancer prevention strategies based on HPV are primary prevention and secondary prevention. Primary prevention is achieved through prophylactic HPV vaccines and secondary prevention through HPV assays which detect precancerous lesions that can be treated[86].

1.3.5.1. Primary prevention

An effective method of primary prevention to control HPV related cancers is prophylactic vaccination. HPV vaccines which are currently in market comprise of Virus-Like Particles (VLPs) made of L1[112]. VLPs are non-infectious and comprise of the native virus particle with no DNA. The three licensed HPV prophylactic vaccines are Cervarix (GlaxoSmithKline (GSK), London, UK), Gardasil and Gardasil9 (both Merck, Kenilworth, New Jersey, USA). Cervarix is a bivalent vaccine and contains HPV16 and HPV18 antigens[113,114].

Gardasil is a quadrivalent vaccine and provides protection against HPV6, HPV11, HPV16 and HPV18[115,116]. Gardasil9 is a nonavalent vaccine and provides protection against HPV6, HPV11, HPV16, HPV18, HPV31, HPV33, HPV45, HPV52 and HPV58[117,118]. All vaccines are highly efficacious and confer complete protection when administered as per the protocol before virus exposure. The original vaccination administration protocol is three doses within six months[119]. Vaccination programs typically target girls of 9-13 years age and increasingly boys[121]. HPV vaccines had been introduced by 65 countries as per 2016 statistics, which include high- and middle- income countries and also those that participate in GAVI (Global Alliance for Vaccines and Immunization) program[120].

Systemic immunization with HPV vaccines generates 10-1000 times higher serum antibody concentrations than in natural infections[116]. In contrast to individuals having natural infections, those that are vaccinated strongly seroconvert[122]. Both type-specific and cross-neutralizing antibodies are generated, however concentrations of cross-neutralizing antibodies are 10-100 times lower than type specific ones[123,124].

1.3.5.2. Screening

Screening helps in detecting treatable precancers and early cancers. There are three major cervical cancer screening methods: 1) cytology 2) detection of DNA or RNA of high-risk HPVs 3) cytology-HPV co-testing. In cytology screening, pathologists evaluate cells obtained from cervical sampling through a microscope. However, it is less sensitive than HPV testing or co-testing. Therefore, in order to achieve good sensitivity, shorter intervals between cytology screens are required[125,126]. HPV testing is capable of detecting precancers earlier and thereby can reduce the numbers of cancers in follow-up[125,127]. Therefore, longer intervals between screens can be considered. Compared to HPV DNA testing, RNA testing is a bit more specific. However, reassurance after a RNA test negative case still needs to be proven[128].

1.3.5.2.1. Triage of cervical screening positive women

HPV-based screening is highly sensitive in detecting viral infection but cannot distinguish early stages of infection from advanced precancerous lesions, i.e. the native predictive value is high but the positive predictive values is rather low. Therefore, additional triage tests are needed to decide whom among the screen-positive individuals more intense examination by colposcopy (cervical examination) and biopsy[129]. During colposcopy, the outer surface of the cervix is observed with a magnifying lens and the detection of lesions is helped by staining with acetic acid. Colposcopy is used to guide biopsy for obtaining lesion tissue for final diagnosis by histology.

There are two triage tests available – microscopic or molecular. The microscopic triage option for HPV-positive women is cervical cytology. Recent studies suggest that in HPV-positive women cytology is more sensitive for detecting precancer than in a screening situation, although with a possible loss in specificity[130-132].

The dual staining based on p16 and Ki67 is one of the molecular triage tests used in cervical screening. p16 is the cellular protein upregulated in transforming HPV infections and Ki67 is a marker for cellular proliferation. As per studies, women who stain positive for both should be referred to colposcopy and those who stain negative for both require only monitoring[133-135]. Another example of molecular triage is HPV genotyping, which is in use in the United States. Here, both HPV16 and HPV18 are genotyped[136]. Other molecular triage tests include checking for methylation of CpG sites in viral (E2, L1 and L2) and host genes[137-141].

1.3.5.3. Treatment and Management

Histopathology is the accepted standard for deciding how the triage test-positive women should be treated. For e.g., although only some CIN3 lesions are invasive, all CIN3s are treated as they are considered as a surrogate of precancer[142]. By contrast, CIN1 lesions are considered as HPV infections and are not treated. CIN2 represents a heterogenous mixture of CIN1 and CIN3 and a large proportion of CIN2 regresses. Logically, therefore, clinicians decide not to treat CIN2 lesions[143].

The different treatment strategies available for cervical precancers are excision, ablation and possibly immunotherapy. The gold standard excision procedures include 1) using a cold knife 2) loop electrosurgical excision 3) loop excision of transformation zone. However, deep excisions could result in future childbirth complications. Ablation technologies include laser, cold coagulation and cryotherapy which are less invasive. However, ablation procedure might not be adequate for deep and large precancers[86]. Immunotherapy is still in experimental stage[144].

The major options for treating invasive CxCas are surgery, external beam radiation, brachytherapy and cisplatin-based chemotherapy. Ultimately, the biggest goal is to provide vaccination, screening and early treatment to fight CxCa. Maximum coverage for HPV vaccines should be achieved as part of the ultimate preventive strategy for CxCa.

1.4. HPV16 E1C

1.4.1. HPV16 transcript pattern based triaging

1.4.1.1. Diagnosing CxCa and high-grade precursors by HPV16 transcription patterns using NASBA-Luminex assay

Unlike HPV DNA testing, RNA detection for screening identifies transcriptionally active viruses. However, E6/E7 transcript-based tests also in respect to specific detection of advanced lesions result in many false-positive results. This leads to overtreatment, unnecessary costs and anxiety for women concerned[148-152]. Therefore Schmitt and colleagues[152] have developed an assay identifying novel HPV16 transcripts or RNA patterns for the intelligible grading of LSIL (low-grade squamous intraepithelial lesion) and HSIL (high-grade squamous intraepithelial lesion), which could be used as an improved screening tool for CxCa precursor screening programs. The study identifies 880^2582 transcript (E1C transcript) as one of the useful biomarkers amongst others.

They employed competitive nucleic acid sequence-based amplification (NASBA) and hybridization to oligonucleotide probes coupled to Luminex beads[153] to quantify different HPV16 transcripts. They analyzed 10 spliced HPV16 RNA sequences (226^409 (E6*I),226^526 (E6*II), 226^3358 (E6*III), 226^2709 (E6*IV),880^2582, 880^2709, 880^3358 (E1^E4), 1302^3358,1302^5639, 3632^5639), 5 fl sequences(E6 fl, E7 fl, E1 fl, E5 fl, L1 fl), cellular $p16^{INK4A}$, and two housekeeping transcripts (Ubiquitin C (UbC) and U1A). The sensitivity of the NASBA assay was a detection limit of 25 to 2,500 copies for all transcripts except 25,000 copies for $p16^{INK4A}$ transcript. Regarding specificity, no cross-reactivity was observed with highly HPV16-related HPV types 31 and 33[152]. The study was conducted in 79 cervical swabs selected from Reims HPV Primary Screening

Cohort Study[154,155] and consisted of 25 NIL/M (no intraepithelial lesion or malignancy), 23 LSIL, 24 HSIL and 7 CxCa cases.

The study describes two combinations of two viral transcripts to distinguish between LSIL and HSIL. The first combination comprises of 880^2582 transcript normalized to 3632^5639 transcript. The second combination comprises of 226^526 (E6*II) transcript normalized to 880^3358 (E1^E4) transcript. The second combination can be analyzed when the transcripts of the first combination are negative or their ratio is below the threshold.

This study has observed that 880^2582 transcript is significantly increased in HSIL and CxCa patients (Figure.1.8.). This transcript potentially encodes E1C, but also E2. The upregulation of this transcript could be due to a) E1C can transactivate virus URR[63] b) suppression of E2 translation after translation termination of E1C, because both ORFs overlap. A repression of E2 is required to repress the URR and thereby repression of E6 and E7 expression is acquired in cells harbouring HPV16 episomes. Low E2 expression is achieved in most cancer cases by integration of HPV16 genome in the E2 ORF region, which deregulates E6 and E7 expression and thereby leading the cells to transformation. However, in HSIL and CxCa cases where HPV16 genomes are present as episomes, E2 must still be repressed. It could be that the E1C encoding transcript may counteract the E2 protein and promotes E6 and E7 upregulation. This could hint an integration-independent mechanism for cancer progression in the cervical epithelial cells where HPV16 genomes are present as episomes[152].

1.4.1.2. HPV16 RNA patterns as triage marker in cervical cancer precursor screening

Detection of HPV16 RNA patterns assay by NASBA –Luminex[152] assay is complex as well as labour-, time- and cost-intensive[85]. Therefore, Höfler and colleagues developed a second generation HPV16 RNA patterns assay using reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) technology which is comparatively less complex than NASBA-Luminex assay[85]. This is the second study that showed E1C transcript as a

useful biomarker for distinguishing the severity of cervical lesions. E1C transcript was found to be present almost exclusively in high-grade lesions (Figure 1.9.).

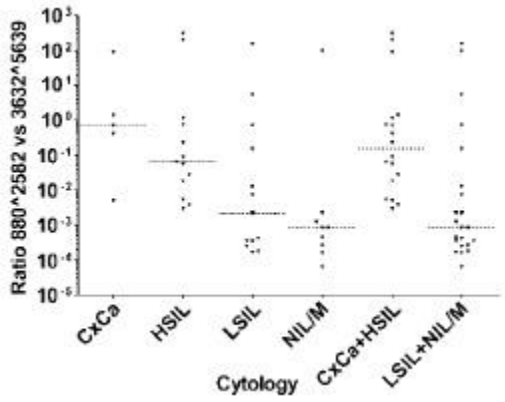


Figure 1.8. **Ratio of viral transcript pairs.** Modified from[152]. Ratio of 880^2582 to 3632^5639 transcripts in different cytology grades and the combined groups of CxCa and HSIL (CxCa+HSIL) and LSIL and NIL/M (LSIL+NIL/M) is shown. Dotted lines represent median values.

This assay uses two RT-qPCRs for the detection of three spliced HPV16 transcripts E6*I (226^409), E1^E4 (880^3358) and E1C (880^2582) and ubiquitin C (cellular housekeeping gene). One RT-qPCR is a triplex reaction quantifying E6*I, E1^E4 and ubC transcripts and the second RT-qPCR quantifies E1C transcript[85]. This assay analyzed a total of 158 singly HPV16 DNA-positive cervical cell samples obtained from a German referral colposcopy clinic and a Mongolian HPV prevalence study and consisted of 36 NIL/M, 11 CIN0, 11 CIN1, 41 CIN2, 41 CIN3 and 18 CxCa. ≥ CIN3 included Histological CIN3 and CxCa samples and ≤ CIN1 included cytological NIL/M and histological CIN0 and CIN1 samples. CIN2 was considered as intermediate because of the highest inter-observer variations reported among the pathologists[85].

Out of the 113 E6*I positive samples analyzed in this study, E6*I was significantly upregulated in ≥ CIN3. Out of the 95 E1^E4 positive samples, E1^E4 showed an increase in CIN3, but decreased in CxCa samples, and the differences were not statistically significant. Out of the 34 E1C positive samples, the mean E1C copies per PCR were in

CIN2 (40), CIN3 (16) and CxCa (9). The reduction of E1C copies in CxCa was statistically significant when compared to CIN2.

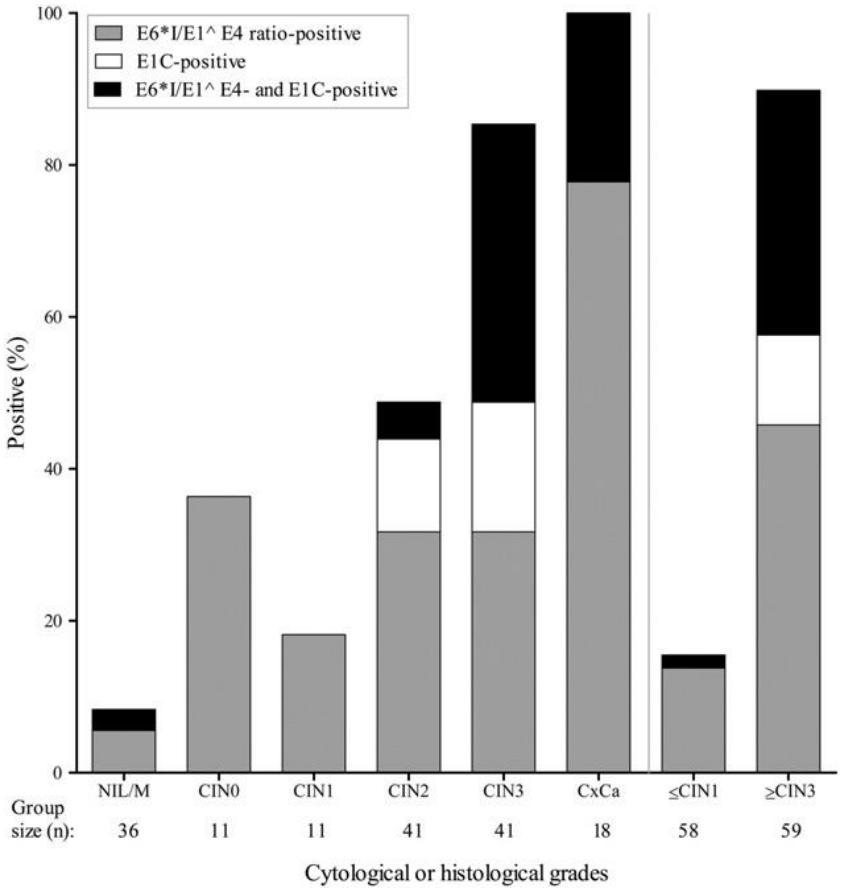


Figure 1.9. **HPV16 RNA patterns in 158 cervical cell samples.** Modified from[85]. X-axis shows the different grades of cervical lesions. Colour code – grey bars represent E6*I/E1^E4 ratio-positive samples, white bars represent E1C-positive samples and black bars represent E6*I/E1^E4 ratio- and E1C-positive samples. Number of samples in NIL/M = 36, CIN0 = 11, CIN1 = 11, CIN2 = 41, CIN3 = 41, CxCa = 18, ≤CIN1 = 58 (combination of NIL/M, CIN0 and CIN1) and ≥CIN3 = 59 (combination of CIN3 and CxCa).

1.4.2. HPV16 E1C activates LCR

Alloul and Sherman[63] have identified that cDNAs consisting of E2 and upstream ORFs of a (880^2708), a' (880^2581) and d (226^2708) mRNAs translate E2 protein at different efficiencies (Figure 1.10.). They checked whether the differential E2 protein translation from a-, a'- and d- type mRNAs affect LCR activity. They also investigated whether the co-translation of proteins by these mRNAs has an effect on LCR regulation by E2[63].

The transcriptional activities of the polycistronic cDNAs were measured by calcium phosphate transient transfections in Cf2Th (Canine foetal thymus) cells using a chloramphenicol acetyltransferase (CAT) reporter gene associated to HPV16 LCR. Details of polycistronic cDNAs containing E2 ORF and upstream ORFs of a-, a'- and d-type mRNAs are described elsewhere[157].

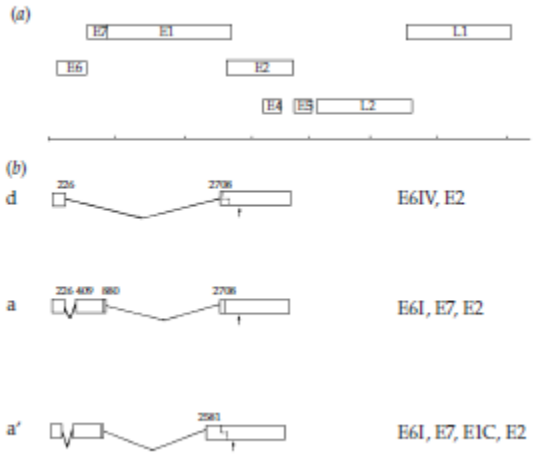


Figure 1.10. **Structure of HPV16 genome and cDNAs containing E2 ORF.** Modified from[63]**.** a) Schematic diagram of HPV16 genome depicting the ORFs. b) Structures of d-, a- and a'- type mRNA species are shown together with their coding potential on the right. Splice donor and acceptor nt positions are marked.

In this study, they have been observed that E2 ORF-containing monocistronic cDNA negatively regulated LCR on a dose-dependent fashion and was E2 protein mediated. They have also observed that E2 ORF-containing polycistronic cDNAs repress HPV16 LCR and the repression was mainly due to the E2 ORF function.

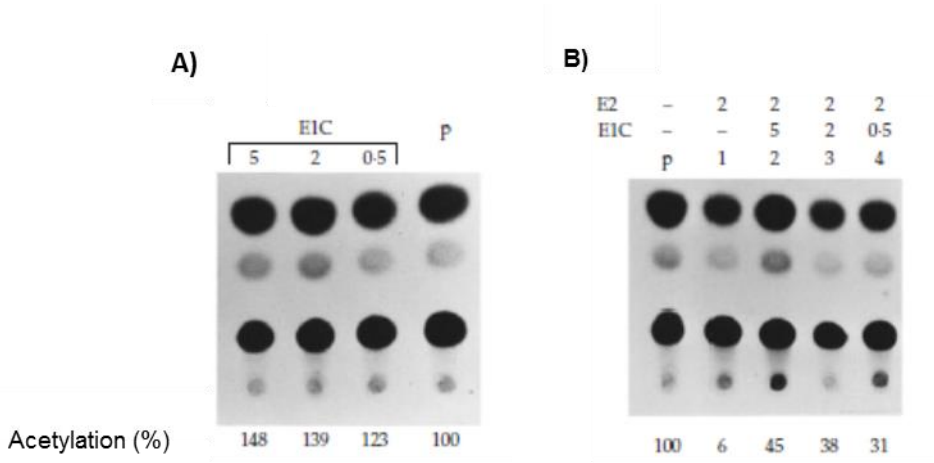


Figure 1.11.**Transcription-modulatory activities of E1C.** Modified from[63]. A) E1C transfections were carried out in Cf2Th cells with 0.5, 2 and 5 µgs as indicated on the top. B) Co-transfections of varying amounts (µgs) of E1C and constant amounts of LCR-CAT plasmid (2 µg) and E2 plasmid (2 µg) were carried out as indicated on the top. Control cells were transfected with pJS55 (backbone vector of E1C) and is indicated as p. The normalized values of acetylation percentages are shown at the bottom.

To study the transcription regulatory effects of individual 5' ORFs, vectors carrying each of the individual ORF were constructed and performed the CAT assay. They have found that E6 splice variant E6IV from d-type mRNA repressed LCR. From the a-type mRNA, E6I weakly activated LCR and E7 did not affect the basal activity of LCR. From the a'-type mRNA, E1C activated LCR[63].

When E1C expression vector was transfected in increasing amounts in Cf2Th cells, the LCR showed dose dependent activation (Figure 1.11. (A)). Also, co-transfection experiments of varying amounts of E1C and E2 expression vectors showed a reduction of E2-mediated LCR repression (Figure 1.11. (B)). The repression of LCR by a'-type mRNA could be due to E2 protein, which is dominant when compared to the activation by E1C and E6I proteins. Alloul and Sherman could not detect E1C protein upon a'-type mRNA translation in their translation studies[157] and speculated that the inability to detect the protein could be due to its low expression levels or instability.

1.5. Aim of my thesis

Papillomaviruses are ubiquitous DNA viruses, which can cause benign flat warts (papillomas) and occasionally, cancer. The high-risk HPVs of Alpha genus are sexually transmitted and can cause cervical cancer, one of the most common cancers in women. Of the high-risk HPVs, HPV16 is the predominant type causing 60% of the cervical cancers and 85% of the non-cervical cancers. Elaborate and intensive research had taken place in the field since the discovery of the link between HPV16 and cervical cancer by Prof. Dr. Harald zur Hausen in1983, which led to the development of prophylactic vaccines for the prevention of cervical cancer. Extensive molecular studies helped to gain information about the viral proteins, which plays a role in cervical keratinocyte transformation, therefore opening up several therapeutic study approaches. As part of the screening program, several diagnostic assays have been developed for the accurate stratification of different grades of cervical cancer precursors to avoid unnecessary anxiety and overtreatment in women. The diagnostic assay developed by Schmitt and colleagues, based on NASBA-Luminex technology and the one developed by Höfler and colleagues, based on RT-qPCR technology are two good examples. In both these diagnostic assays, HPV16 E1C transcript was found to be almost exclusively present in high-grade cervical lesions. This transcript can putatively encode a protein of 82 aa length.

The aim of my thesis was to characterize the diagnostically relevant HPV16 E1C protein. The specific questions I addressed were:

1) What are the effects of overexpression of HPV16 E1C on virus and host?
2) Does HPV16 E1C transcript encode a protein?
3) Does HPV16 E1C transcript or HPV16 E1C** protein activate the URR of HPV16 genome?

** - HPV16 E1C is designated as E1C from hereafter in the ‘Results’ section.

2. Materials

2.1. Cell lines

All the cell lines used in the current study were tested for authenticity and contaminations by multiplex cell line authentication and multiplex cell contamination tests by the company Multiplexion (Felipe Castro 2013 and Markus Schmitt 2009). The table below provides the summary and description of the cell lines used in this study together with the sources indicated.

Table 2.1. **Description of cell lines used in this study.**

Name of the cell line	Description	Source
CaSki	HPV16-positive CxCa	Dr. Michael Pawlita lab[223]
SiHa	HPV16-positive CxCa	
MRI-H186	HPV16-positive CxCa	
MRI-H196	HPV16-positive CxCa	
HPK-1A	Primary human foreskin keratinocytes transfected with HPV16 genome	
C33A	HPV16-negative CxCa	Prof. Dr. Elisabeth Schwarz lab[223]
W12-epi* (clone 20850)	Low grade cervical lesion – HPV16 genome is present as episome	Prof.Dr. Paul. F. Lambert lab[163]
W12 –int** (clone 20831)	Low grade cervical lesion – HPV16 genome is integrated	
NIH-3T3	Mouse fibroblasts	Prof.Dr. Frank Rösl lab
HEK-293T	Human embryonic kidney cells which are adenovirus-transformed and are stably expressing SV40[a] large T-antigen	Prof. Dr. Martin Müller lab
HeLa	HPV18-positive CxCa cell line	Prof. Dr. Martin Müller lab

2.2. Plasmids

Name of the Plasmid**	Description
pGEX-GST-Tag	Glutathione S-Transferase (GST) –Tag fusion protein[a]
pGEX-GST-E1C-Tag	GST–E1C–Tag fusion protein
p1-His-GFP	Histidine (His) – Green Fluorescent Protein (GFP) fusion protein
p1-His-GFP-E1C	His-GFP-E1C fusionprotein
p-ABCB5	Truncated ATP-Binding Cassette sub-family B member 5 (ABCB5) protein with 6x-His tag
p2	eukaryotic expression vector with an EF alpha promoter, GFP sequence driven by SV40 promoter and zeocin resistance gene
p2-E1C	E1C protein
pCDNA3.1(+)	Eukaryotic expression vector with CMV promoter
pCDNA3.1(+)-CASIMO	CASIMO protein with a Flag tag at the C-terminus[192]
LV	lentiviral empty vector with CMV immediate early promoter, IRES sequence and puromycin resistance gene
LV-E1C	E1C protein
p-HPV16-URR-FLuc[d]	Firefly Luciferase (FLuc) enzyme driven by HPV16 URR
p-TATAbox-RLuc	encodes Renilla Luciferase (RLuc) enzyme driven by TATA box
p-EV	eukaryotic expression vector with CMV promoter and ampicillin resistance gene
p-HPV16-E2	HPV16 E2 protein
p-HPV16-E1C	HPV16 E1C protein
p-HPV16-E1C-RNA mut[b]	random RNA sequence with the same GC content as that of E1C RNA
p-HPV16-E1C-Pro mut[c]	mutated E1C sequence with first 3 codons replaced by TGA, TAG and TAA respectively

**- vector maps and DNA sequence information are given in the supplementary information 9.2. and 9.3. respectively.

a – Tag peptide consists of the carboxy-terminal undecapeptide (KPPTPPPEPET) of simian virus 40 large T antigen.

b – 'RNA mut' stands for mutated RNA sequence

c - 'Pro' in 'Pro mut' stands for Protein and 'mut' for mutated, which means that 'E1C-Pro mut' sequence cannot encode for E1C protein

d – HPV16 URR sequence is 7007-102 bp from the HPV16 sequence[225], Firefly Luciferase sequence is taken from *Photinis pyralis*[225-227].

All plasmids containing spliced HPV16 transcripts were kindly provided by Dr. Markus Schmitt (Alumnus of Dr. Michael Pawlita lab at DKFZ, Heidelberg, Germany) and are described in detail in the doctoral thesis of Dr. Daniela Höfler[147] and also given in the supplementary information (Table 9.1.).

2.3. Primers and probes

The primers and probe used for the quantification of HPV16 E7 transcript in reverse transcriptase quantitative PCR is given below. Those used for the quantification of E6*I, E1^E4, E1C and ubC were designed by Dr. Daniela Höfler[85,147] and is protected by patent rights and therefore not disclosed. The primers used for sequencing p-HPV16-URR-FLuc are M13 forward and reverse primers and were taken from primer repository of Eurofins Genomics, Ebersberg, Germany. All the primers and probes were ordered from Sigma Aldrich, St. Louis, USA and TIB MOLBIOL, Berlin, Germany respectively.

Name	Forward primer (5'-3')	Reverse primer (5'-3')	Amplicon (bp)	Probe sequence
HPV16 E7	TTTGCAACCAGAGACAACTGAT	TGTTGCAAGTGTGACTCTACGCT	153	AGAGCCCATTACAATATTGTA

2.4. Antibodies

Antibody	Host species	Application (Dilution)	Company (City, Country)
E1C mAb	mouse	ELISA (1:5) WB (1:5), IF (1:3)	Genomics and Proteomics core facility, Monoclonal antibodies, DKFZ (Heidelberg, Germany)
Goat anti-Mouse pox	mouse	WB (1:10,000)	Dianova GmbH (Hamburg, Germany)
6x-His mAb	mouse	Hybridoma clone supernatant	Genomics and Proteomics core facility, Monoclonal antibodies, DKFZ (Heidelberg, Germany)
Polyubiquitin mAb (D9D5)	rabbit	WB (1:1000)	Cell Signaling Technology (Massachusetts, USA)
Flag mAb	mouse	WB (1:500)	Sigma Aldrich (St. Louis, USA)
Tag mAb*	mouse	ELISA (1:50)	Dr. Michael Pawlita lab, DKFZ (Heidelberg, Germany)
Alexa 594	mouse	WB (1:500)	Thermo Fisher Scientific (Eugene, USA)

mAb – monoclonal antibody, WB – Western blot, IF – Immunofluorescence

*- Tag mAb used was the cell culture supernatant from KT3 hybridoma cell line[219,220].

2.5. Enzymes

Enzyme	Company (City, Country)
EcoRV - HF	New England Biolabs (Massachusetts, USA)
SmaI	New England Biolabs (Massachusetts, USA)
T4 DNA ligase	New England Biolabs (Massachusetts, USA)
T7 RNA polymerase	Fermentas (St. Leon-Rot, Germany)

2.6. General buffers and solutions

Buffer	Ingredients
Blocking buffer (Immunocytochemistry)	PBS+0.1% Triton X-100 2% Horse serum
Blocking buffer (ELISA)	2 mg/mL casein in PBS-Tween®
10% Blocking milk (Western blot)	5 g milk powder 50 mL PBS-Tween
Coating buffer (ELISA)	4 parts $NaHCO_3$ 1 part Na_2CO_3 Adjust pH to 9.6
Coating medium (ELISA)	10 µL Glutathion-casein 10 mL coating buffer
DNA-loading buffer	0.02% Xylene Cyanol 40% (w/v) Sucrose in bidest. H2O Storage at -20ºC
50x Electrophoresis-buffer (TAE)	2 M Tris, pH 7.8 0.25 M NaAc (water free) 0.05 M EDTA Add 8L bidest. H2O Storage at RT
50% Glycerol	25 mL glycerol (100%) 25 mL H2O Autoclave
LB-Agar	LB medium 1.5% (w/v) Bacto-Agar Autoclave, storage at 4ºC
LB medium	10 g (1% (w/v)) Bacto-Tryptone 5 g (0.5% (w/v)) Bacto-yeast extract 10 g (1% (w/v)) NaCl Add 1L bidest. H2O Adjust pH to 7.5 with 5M NaOH Autoclave, storage at 4ºC
3.7% Paraformaldehyde	18.5 g paraformaldehyde 500 mL 1X PBS
PEI	Add 0.16 g to 400 mL water Adjust pH to <2 Stir the solution for 3h Adjust pH to 6.9 - 7.1 Make upto 500 mL with water Sterile filter and storage at -20°C
1X Phosphate Buffered Saline (PBS)	124 mM NaCl 22 mM Na2HPO4

	10 mM KH2PO4 Adjust to pH 7.4 Autoclave, storage at RT
PBS+30 mM Glycine	1.1 g Glycine 500 mL 1X PBS
PBS+0.1% Triton X-100	500 µL Triton X-100 499.5 mL 1X PBS
PBS-Tween®	0.5 mL Tween 20® 1 L 1X PBS
10X Running buffer (Western blot)	151.5 g Tris 720 g Glycine 50 g SDS Make it to 1X by the addition of deionized water
Substrate buffer	0.1 M CH_3COONa (sodium acetate) Adjust pH to 6.0
Substrate solution (ELISA)	100 µL tetramethylbenzidine 2 µL H_2O_2 10 mL substrate buffer
Stop solution (ELISA)	1M H_2SO_4
T4 DNA ligase buffer (10X)	New England Biolabs (Massachusetts, USA)
Transfer buffer (Western blot)	4.84 g Tris 2.92 g Glycine Adjust pH to 9.2 with NaOH Add 200 mL Methanol Make it up to 1L with deionized water

2.7. Reagents

Reagent	**Company (City, Country)**
Agarose	GIBCO Life Technologies (Paisley ,Scotland)
Ampicillin	Roche (Mannheim, Germany)
Bacto-Agar	Becton Dickinson, Sparks (MD, USA)
Bovine Serum Albumin (BSA)	PAA Laboratories GmbH (Linz, Austria)
Bradford reagent	Carl Roth (Karlsruhe, Germany)
DMEM – low glucose	Sigma Aldrich (Steinheim am Albuch, Germany)
DMEM – high glucose	Sigma Aldrich (Steinheim am Albuch, Germany)
1 kb DNA ladder	New England Biolabs (Massachusetts, USA)
EDTA	Acros Organics (Geel, Belgium)
Ethanol (Absolute)	Sigma Aldrich (Steinheim am Albuch, Germany)
Fetal Bovine Serum (FBS)	Gibco, Thermo Fisher Scientific (Eugene, USA)
Fluorescence mounting medium (Dako)	Agilent Technologies (California, USA)
Glycerol (100%), water free	Carl Roth (Karlsruhe, Germany)
Ham's F-12 nutrient mixture	GIBCO Life Technologies (Paisley ,Scotland)
Horse serum (IF)	GIBCO Life Technologies (Paisley ,Scotland)
Hygromycin B	Roche (Mannheim, Germany)
Isopropanol	Sigma Aldrich (Steinheim am Albuch, Germany)
ß-Mercaptoethanol	AppliChem GmbH (Darmstadt, Germany)
Methanol	Sigma Aldrich (Steinheim am Albuch, Germany)
Milk powder	GERBU Biotechnik GmbH (Heidelberg, Germany)
Mitomycin C	Sigma Aldrich (Steinheim am Albuch, Germany)

MS2 RNA	Roche (Mannheim, Germany)
Paraformaldehyde	Carl Roth (Karlsruhe, Germany)
Polyethylenimine (PEI)	Polysciences GmbH (Hirschberg an der Bergstraße, Germany)
Penicillin/Streptomycin	Gibco, Thermo Fisher Scientific (Eugene, USA)
PEQ green	
Protease inhibitor (cOmplete™, EDTA-free Protease Inhibitor Cocktail)	Roche (Mannheim, Germany)
Puromycindihydrochloride	Santa Cruz Biotechnology (Dallas, USA)
RNA ladder “low range”	New England Biolabs (Frankfurt am Main, Germany)
RNase away spray	Molecular Bioproducts (San Diego, USA)
Sodium dodecyl sulfate (SDS)	
Sulphuric acid (96%)	Carl Roth (Karlsruhe, Germany)
TE buffer, RNase-free, 1X (10 mMTris-HCl, 1mM EDTA, pH 8.0)	Acros Organics (Geel, Belgium)
Tris-base	Sigma Aldrich (Steinheim am Albuch, Germany)
Triton-X-100	Carl Roth (Karlsruhe, Germany)
Trypsin-EDTA (0.25%)	Sigma Aldrich (Steinheim am Albuch, Germany)
Tween® 20	Sigma Aldrich (Steinheim am Albuch, Germany)
Water, DNase/RNase-free	Invitrogen (Carlsbad, USA)

2.8. Kits

Kit	Company (City, Country)
Dual-Luciferase® Reporter Assay Kit	Promega (Madison, USA)
ECL Select™ Western Blotting Detection Reagent	GE Healthcare (Chicago, USA)
LightCycler® 480 Probes Master	Roche (Mannheim, Germany)
LightCycler®480 RNA Master Hydrolysis Probes	Roche (Mannheim, Germany)
QIAshredders	QIAGEN (Hilden, Germany)
QIAquick Gel Extraction Kit	Qiagen (Hilden, Germany)
QIAGEN Plasmid Midi Kit	Qiagen (Hilden, Germany)
QIAGEN Plasmid Mini Kit	Qiagen (Hilden, Germany)
RNeasy Mini Kit	QIAGEN (Hilden, Germany)

2.9. Consumables

Consumable	Company (City, Country)
Bottle top filter (500 mL)	Corning Incorporated (Corning, USA)
Cell culture dish (10 cm)	Corning Incorporated (Corning, USA)
Cell culture flasks (T25, T75, T150)	Techno Plastic Products (TPP) AG, (Trasadingen, Switzerland)
Cell culture (Falcon®) plates (6-well, 24-well, 96-well)	Corning (Wiesbaden, Germany)
Cryogenic vials (Nalgene®)	Thermo Scientific (Eugene, USA)
Disposable syringe (2 mL, 5 mL, 10 mL, 50 mL)	Terumo (Tokyo, Japan)

Electroporation cuvette	Steinbrenner Laborsysteme GmbH (Wiesenbach, Germany)
Examination gloves (Microflex®)	Ansell Health Care (Brussels, Belgium)
Falcon tubes (15 mL and 50 mL)	Greiner Bio-one (Frickenhausen, Germany)
Filter tips (TipOne®) (10 µL, 20 µL, 200 µL, 1000 µL)	STARLAB (Hamburg, Germany)
LightCycler® 480 Multiwell plate 96, white	Roche Applied Science (Mannheim, Germany)
LightCycler®480 Sealing foil	Roche Applied Science (Mannheim, Germany)
Low retention tubes (Nucleic acid extraction and storage)	Kisker Biotech GmbH & Co. KG (Steinfurt, Germany)
Microplates (96-well, Chimney well, white, lumitrac) (Dual-luciferase reporter assay)	Greiner Bio-one (Frickenhausen, Germany)
Microscopic slides	Thermo Scientific (Eugene, USA)
Neubauer counting chamber	BRAND GmbH + Co KG (Wertheim, Germany)
Nitrocellulose membrane	GE Healthcare (Chicago, USA)
Nunc Polysorp 96-well plates	Thermo Scientific (Eugene, USA)
Parafilm	Bemis Company (Wisconsin, USA)
Reaction tubes (1.5 mL and 2 mL)	Eppendorf (Hamburg, Germany)
Reagent reservoirs (50 mL)	Corning (Wiesbaden, Germany)
Waste plastic bags	nerbe plus GmbH (Winsen, Germany)

2.10. Instruments

Instrument	**Company (City, Country)**
Agarose gel electrophoresis chamber, gel tray and combs	Renner (Dannstadt, Germany)
Cell culture hood	The Baker Company (Sanford, USA)
Cell observer SD microscope	Zeiss (Jena, Germany)
4° centrifuge	Eppendorf (Hamburg, Germany)
Centrifuge (Biofuge, cell culture)	Heraeus (Hanau, Germany)
Centrifuge (Megafuge 1.0 R, cell culture)	Heraeus (Hanau, Germany)
Centrifuge (Sorvall RC 6 Plus)	Thermo Scientific (Eugene, USA)
Chemiluminescence western blot detection system (Chemocam Imager)	Intas Science Imaging Instruments GmbH (Göttingen, Germany)
DNA gel cutting instrument	Herolab GmbH laboratory equipment (Wiesloch, Germany)
Eclipse Ti fluorescence inverted microscope	Nikon (Minato, Japan)
Gel documentation system (Gel Doc EZ)	Bio-Rad Laboratories (Hercules, USA)
Gene Pulser Electroporation device	Bio-Rad Laboratories (Hercules, USA)
Ice machine (automatic ice maschines AF30)	Scotsman (Milan, Italy)
37ºC Incubator (Cell culture)	LabotectLabor-Technik-Göttingen GmbH (Rosdorf, Germany)
Leica DM IL Inverted laboratory microscope	Leica microsystems (Wetzlar, Germany)
LSM 710 confocal microscope	Zeiss (Jena, Germany)
Luminometer	Biotek (Vermont, USA)
Microwave oven	Mikromat AEG (Nuernberg, Germany)
Multichannel pipette (20-200 µL)	BRAND GmbH + Co KG (Wertheim, Germany)
Multiskan™ FC microplate photometer	Thermo Scientific (Eugene, USA)
Nanodrop™ 1000 spectrophotometer	Thermo Scientific (Wilmington, USA)
Neubauer chamber	Carl Roth (Karlsruhe, Germany)

pH meter	inoLab® (Weilheim, Germany)
Pipetboy	Integra Biosciences (Fernwald, Germany)
Pipettes (2 µL, 10 µL, 20 µL, 200 µL, 1000 µL)	Gilson International France (Villiers-le-Bel, France)
Power supply (Agarose gel electrophoresis and SDS-PAGE)	Consort (Turnhout, Belgium)
QIAgility	QIAGEN (Hilden, Germany)
Skyline digital rocking shaker DRS-12 (at 4ºC)	ELMI (Riga, Latvia)
Stuart™ see-saw rocker	Fisher Scientific (Loughborough, UK)
Shaker at 37ºC	INFORS AG (Bottmingen, Switzerland)
Shaker at RT	Heidolph (Schwabach, Germany)
Thermomixer	Eppendorf (Hamburg, Germany)
Vortexer (Vortex Genie)	Bender & Hobein GmbH Laboratory Chemicals (Bruchsal, Germany)
Water bath (Cell culture)	JULABO GmbH (Seelbach, Germany)

2.11. Computer programs and online softwares

Program/Software	Company (City,Country) or Website
BLAST NCBI	http://www.ncbi.nlm.nih.gov/BLAST/
EMBL-EBI Clustal Omega	https://www.ebi.ac.uk/Tools/msa/clustalo/
GraphPad Prism®6	GraphPad Software, Inc. (La Jolla, USA)
Nucleotide NCBI	http://www.ncbi.nlm.nih.gov/nuccore/
Primer 3 web	http://primer3.ut.ee/
Zen microscope software	Zeiss (Jena, Germany)

2.12. Services

Service	Website
DKFZ Genomics and Proteomics core facility – Light Microscopy	https://www.dkfz.de/en/bildgebung/index.html
DKFZ Genomics and Proteomics core facility – Antibodies	https://www.dkfz.de/gpcf/antibodies
DKFZ Genomics and Proteomics core facility – Cellular Tools/Vector & Clone Repository	https://www.dkfz.de/gpcf/isogenic-cell-lines
DKFZ Genomics and Proteomics core facility – Mass Spectrometry based Protein Analysis	https://www.dkfz.de/gpcf/mass-spectrometry
Multiplexion GmBH	http://www.multiplexion.de/
Eurofins Genomics	https://www.eurofinsgenomics.eu/

3. Methods

3.1. Cell culture

3.1.1 Maintenance of cell lines

CaSki, SiHa, MRI-H186, MRI-H196, HPK-1A, C33A and HEK-293T cells were cultured in DMEM containing low glucose supplemented with 10% FBS and 1% penicillin/streptomycin. NIH-3T3 cells were grown in DMEM containing high glucose supplemented with 10% FBS and 1% penicillin/streptomycin. All cell lines were cultured in a 37ºC incubator with 5% CO_2.

Confluent grown cells were splitted 1:4 or 1:5. For this, the growth medium was removed from the cell culture flask and 1X PBS was added to wash the cells and remove the residual growth medium from the flask. To harvest the adherent cells, Trypsin-EDTA was added and the cells were incubated at 37ºC from 2-10 minutes depending on the cell line. Afterwards, pre-warmed growth medium was added to the cells to stop trypsinisation and the cells were harvested in a Falcon tube. Finally, the required volume of cells was transferred back to the flask and appropriate amount of pre-warmed growth medium was added depending on the cell culture flask. The cell culture flask was then incubated at 37ºC incubator with 5% CO_2 until further use.

3.1.2. Culturing of W12 cells

The protocol for culturing of W12 cells was adopted from Dr. Paul F. Lambert laboratory, McArdle Laboratory for Cancer Research, University of Wisconsin-Madison. All the materials required for the preparation various cell culture reagents for W12 cells[221] are given in the supplementary information 9.4..

W12-epi and W12-int cells cells were grown on mitomycin C treated NIH-3T3 cells, which acted as feeder cells. NIH-3T3 cells were grown in 3T3 cell medium in a T-150 flask and were fed two times per week. For the preparation of feeder layer, the growth medium was aspirated from a confluent T-150 flask and 200 µL 50X mitomycin C mixed in 10 mL 3T3 cell medium was added to the cells. The cells were incubated 37ºC incubator with 5%

CO_2 for 2h. Afterwards, 3T3 cell medium and mitomycin C were aspirated from the flask and the cells were washed with 1X PBS two times. To detach the cells from the bottom, 2 mL 0.25% Trypsin-EDTA solution was added to the flask and incubated for 1 min at RT. $1/4^{th}$ of the cells were seeded to a new T-150 flask and were incubated at least 2 h to 24 h in W12 medium before the W12 cells were seeded onto them

For the thawing of W12 cells, $1x10^6$ cells were thawed onto prepared NIH-3T3 feeder cells in W12 medium in a T-150 flask. The W12 cells were fed every two days and small clones of cells were observed after 1 week of culturing. A confluent T-150 flask of W12 cells were obtained after 2 weeks of culturing.

For passaging W12 cells, the medium was aspirated from the confluent flask. To remove NIH-3T3 feeder cells, 5 mL 0.02% EDTA solution was added and placed in the 37°C incubator for 10 min. The flask was tapped on the sides to remove the attached feeder cells. 0.02% EDTA solution was removed and the cells were washed with 1X PBS to completely remove the feeder cells. To detach the W12 cells, 2 mL 0.25% Trypsin-EDTA solution was added to the flask and incubated at 37°C for 10 min. The cells were re-suspended in W12 medium by pipetting up and down to obtain a single cell suspension. After counting, around $1x10^6$ cells were seeded onto a new T-150 flask.

3.1.3. Freezing and thawing of cell lines

To freeze cells, the culture medium was removed and cells were rinsed once with 1x PBS. Then Trypsin-EDTA was added to detach the cells from the culture flask. After incubation at 37°C, trypsinisation is stopped by the addition of growth medium and the cells were centrifuged at 1500 rpm for 5 min to remove trypsin containing supernatant. The cells were resuspended in 1 ml freezing medium and immediately transferred to cold cryotubes and were placed in freezing box (The freezing box should be cold in order to maximize the number of viable cells which go to freezing). This facilitates a constant freezing rate of -1°C/ min and is stored in -80°C overnight and can be shifted from the freezing box to another storage box.

The cells frozen at -80°C in a cryotube are thawed by quickly transferring it in warm growth medium in a 15 mL Falcon tube to dilute the toxic DMSO of the freezing medium. The cells were then centrifuged at 1500 rpm for 5 min. The supernatant was removed thereby removing the toxic DMSO and the cell pellet is dissolved in fresh warm growth medium. The cells were then seeded to appropriate culture dishes.

3.1.4. Determination of cell number

In order to count the cells before seeding for various experiments, manual cell counting by Neubauer chamber was employed. After harvesting the cells from the cell culture flask, 10 µL of cells was taken in a 1.5 mL Eppendorf tube and mixed with 90 µL of Trypan blue to make a 1:10 dilution of the cells. Then roughly 10 µL of the cell dilution was pipetted onto the counting grid of the Neubauer chamber. The cells distributed on the counting grid were counted with the help of an optical microscope and cell counter. The cells present in the 4 corner squares on the counting grid were counted for calculation. The total number of cells harvested from the cell culture flask is calculated using the below formula.

Total no: of cells = (total cells counted from 4 squares / 4) x 10^4 x 10 x total volume in which cells are suspended

where 10^4 is the multiplication factor of the chamber and 10 is the dilution factor of the cells in Trypan blue.

3.2. Generation of E1C monoclonal antibody (mAb)

A monoclonal Ab specific for the fusion peptide generating by E1C splicing was generated using hybridoma technology in collaboration with the Antibodies unit of the Genomics and Proteomics core facility at DKFZ. In hybridoma technology, lymphocytes isolated from the mouse immunized with the antigen of interest are fused with myeloma cells to generate hybridomas. These hybridomas are then cultured and clones secreting antibody of desired specificity are selected and expanded.

For the generation of E1C mAbs, an E1C peptide was designed for mouse immunization (see Results 4.2.1.1.).Three rounds of immunization were performed in four mice, 2 BALB/c and 2 C57black6N with an interval of 7-9 days. 20 µg of E1C peptide was injected intramuscularly on both the hind legs of each mouse per immunization. After the immunizations, the sera of all the mice together with pre-serum from one mouse were tested for the presence of E1C Abs in an ELISA with E1C peptide as antigen and in western blot using E.coli lysate overexpressing GST-E1C and HeLa cellular lysate overexpressing GFP-E1C fusion proteins. The mouse in which the serum contained E1C Abs was selected to perform fusion. The B and T lymphocytes were isolated from the lymph nodes of the popliteal fossa (knee pit) of the mouse and were fused with Sp2/0 cells, which is a B cell myeloma cell line from BALB/c mouse. Polyethylene glycol was used for the fusion of the cells and afterwards the fused cells were plated on a 96-well plate and incubated at 37ºC and 5% CO2 for 4 – 7 days. Supernatants of growing clones on the plate were screened by ELISA and western blot to identify the E1C Ab producing clone. E1C Ab positive clones obtained from the screening were cloned for two more rounds and further screened to finally obtain E1C Abs from a stable single sub-clone. Antibody was purified from about 1L of cell culture supernatant using Protein A Sepharose® beads.

3.3. Transient transfection of HEK-293T cells

The overexpression of E1C to detect the E1C protein and the overexpression of various plasmids to study their effects on HPV16 URR activity by luciferase assay were achieved by transient transfection using polyethylenimine (PEI). Appropriate number of HEK-293T cells was seeded 24h prior to transfection. The cells need to be 70% confluent at the time of transfection. The following table summarizes the amount of each component of the transfection mix that was used in the case of a 24 well plate or a 10 cm dish transfection.

Plate/Dish	No: of cells seeded for transfection	Water (µL)	Plasmid DNA	DMEM	PEI (µL)	DMEM+5% FBS
24 well plate	75000	8.3	600 ng	125 µL	1.7	500 µL
10 cm dish	$1.5x10^6$	62	10 µg	1.5 mL	31	8.5 mL

For the preparation of transfection mix, water, plasmid DNA, DMEM and PEI were added to a 1.5 mL Eppendorf tube (24 well plate transfection) or a 15 mL Falcon tube (10 cm dish transfection) in the same order. The transfection mix was vortexed for 10s and incubated at RT for 10 min. Afterwards, DMEM+5% FBS medium was added carefully to the transfection mix and pipetted gently one time up and down for the even mixing of the PEI-DNA complex. The growth medium from the cells was removed and the PEI-DNA complex was added. The plate/ dish was gently moved back and forth and sideways for the even distribution of PEI-DNA complex in the wells/ dish. The transfection mix was removed from the cells 12-14 h post transfection and pre-warmed growth medium was added. The cells were analyzed 48h post transfection.

3.4. Stable transfection of W12-epi, W12-int, CaSki and SiHa cells

In order to study the effects of E1C overexpression on HPV16 transcript levels, stable transduction of HPV16-containing cell lines by lentiviral vector particles transfection was employed. The production of lentiviral particles and lentiviral transduction were performed in co-operation with Cellular Tools/Vector & Clone Repository unit of DKFZ Genomics and Proteomics core facility.

For the cloning of E1C gene to a lentiviral vector, E1C gene flanked by attL recombination sites was cloned into a pMX plasmid (Thermo Fischer Scientific). Using Gateway recombination technology, the E1C construct was shuttled in a lentiviral expression vector fusing an internal ribosomal binding site (IRES) at the C-terminus. The pMX plasmid contains a kanamycin resistance marker and lentiviral expression vector consists of a puromycin resistance marker.

For the generation of lentiviral particles[201], the lentiviral expression constructs (LV and LV-E1C) were co-transfected with 2nd generation viral packaging plasmids VSV.G (Addgene #14888) and psPAX2 (Addgene #12260) in HEK293FT cells (Thermo Fischer Scientific, Germany). The supernatant containing virus was removed 48 h post transfection and cleared by centrifugation at 500 g for 5 min. To remove cell debris, the supernatant was passed through a 0.45 µm filter

In order to study the effects of E1C overexpression on HPV16 transcripts, stable transfection by lentiviral transduction was employed. Appropriate no: of W12-epi, W12-int, CaSki and SiHa cells was seeded to a 6-well plate, 24 h prior to transduction. An additional well of cells was also seeded which will not be transduced and which acted as the 'control well' later for antibiotic selection of transduced cells. The cells need to be 70% confluent at the time of transduction. 24 h post seeding, cells were transduced with lentiviral particles with the aid of 10 µg/mL polybrene (Merck, Germany). Virus containing medium was replaced with puromycin selection medium 24h post transduction. The cells were subjected to puromycin selection until the cells in the 'control well' are all dead. Afterwards, the cells were harvested for gene expression analysis by RT-qPCR.

3.4.1. Dose response kill curve for puromycin selection

A puromycin kill curve was performed to determine the minimum concentration required to kill non-transduced W12-epi, W12-int, CaSki and SiHa cells. Cells were seeded in 96-well plate to be 50% confluent. Growth medium was removed and selection medium (growth medium + puromycin) was added. A range of puromycin concentrations from 0.5 – 10 µg/mL were used. The negative control was cells supplemented with growth medium without any puromycin. The cells were monitored under microscope daily for three days after puromycin addition and observed the percentage of surviving cells. The minimum puromycin concentration used for each cell line was the lowest concentration that killed 100% of cells in three days. The lowest puromycin concentrations determined for all the cell lines were – W12-epi (700 ng/mL), W12-int (800 ng/mL), CaSki (500 ng/mL) and SiHa (700 ng/mL).

3.5. Proteasome inhibition of cells by MG132

To enhance detection of overexpressed E1C protein in HEK-293T cells and endogenous E1C protein in CaSki and SiHa cells, proteasome inhibition by MG132 was employed.

3.5.1. MG132 tolerance test of cell lines

Appropriate no: of HEK-293T, CaSki and SiHa cells were seeded onto a 24-well plate so that the cells will be 70% confluent 24 h later. Different MG132 concentrations ranging from 0 µM – 20 µM were applied on the cells 36 h post seeding, to determine the maximum MG132 concentration tolerated by each cell line. The cells were monitored under the microscope for cell death or morphology differences every 2 h for 10 h. All the cell lines were showed to tolerate 10 µM MG132 concentration.

3.5.2. Proteasome inhibition by MG132

HEK-293T cells were seeded onto 10 cm dishes so that the cells will be 70% confluent at the time of transfection. The cells were transfected with p2 and p2-E1C to two 10 cm dishes each by PEI transfection. The transfection medium was changed 12 h post transfection. 42 h post transfection, one pair of p2- and p2-E1C-transfected dishes were treated with DMSO and the second pair were treated with 10 µM MG132 for 6 h. Afterwards, the cells were analyzed for E1C protein detection by western blot and immunofluorescence.

CaSki and SiHa cells were also processed the same way except that they were not transfected with p2 or p2-E1C, because the aim was to detect the endogenous E1C protein in these cell lines after proteasome inhibition by MG132.

3.6. Dual-luciferase reporter assay (DLRA)

In order to study the effect of E1C on HPV16 URR, DLRA was employed. Reporter systems are widely used in molecular biology to study gene expression, promoter activity, transcription factors etc.. DLRA employs the simultaneous expression and measurement of two reporters in one assay and thereby improves assay accuracy. Out of the two reporters, the ‘experimental’ reporter is associated with the different experimental conditions and the ‘control’ reporter is used for normalization of the experiment. This normalization helps in minimizing experimental variabilities caused by transfection

efficiency or differences in pipetting volumes[224]. The optimal seeding cell number for HEK-293T cells to be 70% confluent in a well of a 24-well plate 24 h post seeding was determined to be 75,000 cells. 500 – 600 ng of total plasmid DNA was determined to be the optimal total DNA transfected to the cells to avoid DNA overdose toxicity. p-HPV16-URR-FLuc, p-TATAbox-RLuc, p-EV, p-HPV16-E2, p-HPV16-E1C, p-HPV16-E1C-RNA mut and p-HPV16-E1C-Pro mut were the plasmids transfected by PEI transfection. Here, p-HPV16-URR-FLuc was the experimental reporter and p-TATAbox-RLuc was the control reporter. Post 48 h transfection, growth medium was removed and the cells were washed with 1X PBS. 100 µL 1X passive lysis buffer was added to the cells and the plate was kept on a rocker with gentle rocking to mediate cell lysis. After 20 min of lysis, the cells were observed under the light microscope to confirm complete dissolution of the cell structure. Afterwards, the cells were harvested and stored at -80ºC till the luciferase activity measurement.

3.6.1. Measurement of luciferase activity

The samples were taken out from -80ºC and were thawed at room temperature. After proper mixing, 20 µL of the samples were added to the wells of 96-well microplates. Firefly luciferase and renilla luciferase substrates were prepared using the Dual-luciferase® reporter assay kit according to the manufacturer's instructions. Luciferase activities were measured in a luminometer using Gen5^{TM} software. To measure the luciferase activities, 50 µL of firefly luciferase substrate was added and waited 30s before the measurement of firefly luciferase activity. Afterwards, 50 µL of renilla luciferase substrate was added and waited 30s before the measurement of renilla luciferase activity.

3.7. Molecular biological methods

3.7.1. Agarose gel electrophoresis

Agarose Gel Electrophoresis is a method to separate a mixed population of DNA in a matrix of agarose. The DNA fragments are separated by size. An electric field is applied

to move the negatively charged DNA molecules in the agarose gel matrix. Smaller DNA molecules run faster leaving behind larger DNA molecules. The concentration of agarose prepared depends upon the size of the desired DNA fragment and varies in between 0.5%-1.5%. The concentration of the agarose is inversely proportional to the size of the DNA fragment to be separated. Agarose was prepared by boiling it in 1X TAE buffer. After agarose was completely dissolved in TAE buffer, PEQ green was added which is a DNA intercalating agent and help in the visualization of DNA upon UV illumination. After proper mixing, the melted agarose was poured onto gel casting trays previously set with combs. After the gel was set, it was slowly transferred to the electrophoretic tanks filled with 1X TAE. DNA samples were mixed with 6X loading dye and pipetted into the gel slots. A 1kb ladder was added at the extreme ends of the gel as size standard. Separation of DNA fragments was made possible by applying 100V and the gel was analyzed later under UV light.

3.7.2. DNA purification

The correct DNA band identified by agarose gel electrophoresis was excised under UV light after taking right protection by wearing a helmet and nitrile gloves. The DNA present in the excised bands was purified using QIAquick Gel Extraction Kit according to the manufacturer's instructions.

3.7.3. Restriction digestion

In order to clone the desired DNA fragment into plasmid vector, both vector and inserting DNA fragment should have same restriction sites. For this purpose, both the vector and the insert should be enzymatically digested. Therefore, DNA (vector or insert), restriction enzyme, enzyme buffer and ddH_2O were mixed and digested for 2 h at 37^oC. Afterwards, the digested mixture was subjected to agarose gel electrophoresis for analyzing the accuracy of digestion. The correct products were excised and purified as described earlier. The concentration of the purified samples were measured using NanoDrop™ 1000 spectrophotometer.

3.7.4. Ligation

To integrate the digested insert into the digested vector, which were enzymatically digested with the same restriction enzymes, ligation method was used. For optimal ligation reaction, a 1:3 vector: insert ratio was used. The formula which was used to calculate the amount of insert to be used is as follows:

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{molar ratio of } \frac{\text{insert}}{\text{vector}}$$

In a total volume of 20 µl, 1 µl of T4 DNA ligase, 2 µl of ligation buffer and respective amounts of vector and insert were used. The ligation was performed at room temperature for 10 min. Afterwards, the ligation mixture was heat inactivated at 65°C for 10 min. Then the ligation mixture was kept on ice and used 5 µL for transformation into electrocompetent E*. coli* BL21 cells. This ligation protocol was based on the New England Biolabs kit and manual. Two ligation mixes were prepared – control ligation mix and experimental ligation mix. To the control ligation mix, only the purified digested vector was added, and to the experimental ligation mix, both vector and insert were added.

3.7.5. Transformation

Electrocompetent *E.coli* BL21 were thawed for 2-3 min on ice. The prepared ligation mixture/ diluted midiprep or miniprep (with final amount 100 ng) was added to the bacteria, mixed quickly and transferred to precooled Trafo-cuvette. Afterwards the transformation was performed with electroporation device. Soon after the electroporation of the plasmid inside the bacterial cells, 200 µL LB medium was added to the Trafo-cuvette, mixed quickly and added to a 1.5 mL Eppendorf tube. Then the bacteria was incubated at 37°C for 1 h with shaking at 1400rpm in a thermomixer. Afterwards, the bacterial suspension was spread on LB agar plates containing ampicillin antibiotic as the plasmids possess ampicillin resistance gene. The plates were incubated overnight at 37ºC.

3.7.6. Plasmid DNA purification

If the transformation was successful, i.e. if the experiment plates had many colonies and control plates had null or very few colonies, then the purity of cloned plasmids was checked by performing minipreps followed by sequencing. The colonies were picked from the experimental plates and put in 2 mL LB medium containing ampicillin antibiotic and cultured overnight at 37ºC in a shaking incubator. Minipreps were performed the following day using the QIAGEN Plasmid Mini Kit according to the manufacturer's manual. If cloning was successful, then midipreps were prepared. Transformation was again carried out with the positive clone and one clone was picked and put in 5 mL LB ampicillin medium containing ampicillin antibiotic and cultured overnight at 37ºC in a shaking incubator. Midipreps were performed the following day using QIAGEN Plasmid Midi Kit according to the manufacturer's manual.

3.7.7. Sequencing

After doing minipreps, the DNA samples were sent for sequencing to Eurofins Genomics, Ebersberg, Germany. The primers for sequencing were either selected from the primer database of Eurofins Genomics or were ordered from them. Sequencing results were analyzed using Clustal Omega multiple sequence alignment program.

3.8. Quantification of gene expression

3.8.1. RNA extraction

Cells harvested were stored at -80ºC till RNA was extracted. Total RNA extraction was performed using the RNeasy Mini Kit (QIAGEN) including DNaseI treatment according to the manufacturer's instructions. RNA was eluted in 30 µL RNase-free water and the concentration was measured using NanoDrop™ 1000 spectrophotometer at 260 nm. Extraction controls were also added during RNA extraction to check for RNA carry-over contamination.

3.8.2. Reverse transcriptase quantitative PCR (RT-qPCR)

RT-qPCR was performed to determine the expression levels of E6*I, E1^E4, E1C, E7 and ubC transcripts in CaSki, SiHa, W12-epi and W12-int cell lines after the overexpression of E1C. Roche cobas® z 480 analyzer system and LightCycler® 480 RNA Master Hydrolysis Probes were used for this purpose. Primers and probes used for amplification are described in Materials 2.3.. E6*I, E1^E4 and ubC transcripts were measured in a triplex RT-qPCR and E1C and E7 transcripts were measured in a singleplex RT-qPCR separately. Table below describes the different steps and the conditions used for the singleplex and triplex RT-qPCRs.

	Temperature (°C)		Time		No: of cycles	
PCR step	**Singleplex**	**Triplex**	**Singleplex**	**Triplex**	**Singleplex**	**Triplex**
Reverse transcription	63	63	5 min	10 min	1	1
Denaturation	95	95	30 s	30 s	1	1
Amplification	95,65,72*	95,60,72*	10 s, 40 s, 1 s*	10 s, 30 s, 1 s*	50	50
Cooling	40	40	10 s	10 s	1	1

*- amplification temperatures and times correspond to each other in the given order.

3.8.3. Standard curves and quantification of transcripts

A dilution series of 10^6 to 10^1 copies of *in vitro* transcripts of E6*I, E1^E4 and ubC transcripts together and E1C transcript separately were prepared in the background of 20 ng/µL MS2 RNA. Also, a dilution series of 10^6 to 10^1 copies of full-length HPV16 DNA were prepared in 50 ng/µL human placenta DNA, for the quantification of E7 transcript levels. Dilution series of each transcript was amplified and the standard curves were plotted with the copy numbers of transcripts and the corresponding crossing point (Cp) values obtained. The absolute copy numbers of transcripts in unknown samples were calculated from the standard curves. During the RT-qPCR, apart from standards and samples, water controls and extraction controls were also applied in duplicates to check for PCR and RNA contaminations respectively.

3.9. Protein analyses

3.9.1. Preparation of cell lysates

Cell lysates were prepared for detection and quantification of proteins by performing SDS-PAGE and Western blot. The cells were harvested and washed two times with 1X PBS. Afterwards, the cells were lysed by adding 100 µL CSK I lysis buffer and incubated on ice for 20 min. The cells were then centrifuged at 13,000 rpm in a table top centrifuge at 4ºC for 10 min. The supernatant, which is the cell lysate was collected in fresh 1.5 mL eppendorf tubes and stored at -80ºC until further use.

3.9.2. Bradford Assay

Bradford assay was performed in Nunc Polysorp 96-well plates to measure the protein concentration in cell lysates. For this, the Bradford reagent was diluted 1:4 in autoclaved bidest H_2O. A 1:8 BSA dilution series ranging from 0.25 mg/mL – 8 mg/mL was used as the standard. 1 µL of BSA standard and sample each were added to the wells of the 96-well plate in duplicates. Afterwards, 200 µL of diluted Bradford reagent was added using a multichannel pipette and mixed thoroughly by pipetting up and down 10 times without the formation of bubbles. Then the plate was incubated for 5 min at RT. The protein concentration of the cell lysates was measured at 595 nm in a Multiskan^{TM} FC microplate photometer using SkanIt software.

3.9.3. SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

In order to separate the denatured proteins based on their molecular weight, SDS-PAGE was performed. The cell lysates stored at -80ºC were thawed on ice. Appropriate volume of thawed cell lysate was mixed with water and 4X SDS so that the final concentration of the SDS becomes 1X. The cell lysate was then incubated at 95ºC for 5 min to disrupt any protein-protein interactions. Afterwards the cell lysates was loaded onto lab-made gels

and run in 1X running buffer at 200V until the bromophenol blue reached the bottom of the gel. The separating and stacking gels were prepared in the lab and the concentration of the acrylamide in the separating gel was dependent on the size of the protein of interest. Prestained Protein Ladder, Broad Range (10-230 kDa) (New England Biolabs) or Spectra™ Multicolor Low Range Protein Ladder (Thermo Fisher Scientific) were used as protein markers to estimate the size of the separated proteins.

3.9.4. Western blot (WB)

For the detection of proteins by specific antibodies, the separated proteins were blotted onto nitrocellulose membranes. A sandwich was prepared inside a cassette, before kept inside the transfer block. The components of the sandwich were dipped in transfer buffer before being placed on the cassette. The order of the sandwich components placed inside the cassette were as follows – Sponge ->Whatman filter paper -> nitrocellulose membrane -> gel on which the proteins were separated ->Whatman filter paper -> Sponge. Remove the trapped air bubbles in the sandwich by rolling a glass pipette on top. The proteins were then blotted onto the membrane at 250 mA for 35 min. After the blotting, the nitrocellulose membrane was blocked in 10% blocking milk for 1h. The membrane was washed with PBS-Tween three times for 10 min afterwards. The membrane was then incubated in primary Ab overnight at 4°C. Unbound primary Ab was removed by washing with PBS-Tween three times for 10 min. The membrane was then incubated in secondary Ab coupled to Horse Radish Peroxidase for 2h at RT. Unbound secondary Ab was removed by washing with PBS-Tween three times for 10 min. For the detection of proteins, Amersham™ ECL Select™ Western Blotting Detection Reagent (GE Healthcare) was used. Equal volumes of ECL Select ™ Luminol solution and ECL Select ™ Peroxide solution were mixed and applied onto the membrane. The membrane was incubated for 1 min at RT and developed by INTAS chemocam imager suitable for chemiluminescence imaging.

3.9.5. Immunocytochemistry

For the visualization and localization of HPV16 E1C protein, immunofluorescence (IF) staining was utilized. Cover slips were placed inside 10 cm dishes at first, followed by seeding of $1x10^6$ HEK-293T cells. 24 h post seeding, HEK-293T cells were transfected by PEI reagent as described previously. 48 h post transfection, the growth medium was removed from the cells and the cells were fixed by adding 3.7% paraformaldehyde for 10 min at RT. Afterwards, the cells were washed with 1X PBS three times for 5 min each. Each cover slip was placed in a well of a 6-well plate and performed the following steps except for the staining steps with gentle horizontal oscillation in a see-saw rocker. The cells were washed with PBS+30 mM Glycine for 5min at RT to quench paraformaldehyde. The cells were then permeabilized with PBS+0.1% Triton X-100 for 10 min at RT. Afterwards, blocking buffer was added to the cells and incubated at RT for 45 min. The cells were then washed three times with 1X PBS for 5 min each. A parafilm sheet was kept inside an aluminium foil covered box and 30 µL drop of E1C Ab diluted in blocking buffer was placed on the parafilm sheet. Depending on the number of cover slips to be stained, several 30 µL drops were placed. The cover slips were placed on the E1C Ab drops with cells touching the Ab solution and incubated for 2 h at RT in dark. The cover slips were then placed back in the wells of 6-well plate and washed three times for 5 min each with PBS+0.1% Triton X-100. For the secondary Ab staining and washing afterwards, same procedure like E1C Ab staining was followed. The secondary Ab used was Alexa Fluor™ 594. The cells were again washed with 1X PBS for 5 min. Afterwards, the cells were rinsed in deionized water and dried by placing on a Whatman filter paper. The cover slips were then mounted on glass slides using fluorescence mounting medium and dried at RT overnight and stored at 4°C from next day onwards. For the visualization of stained E1C protein, confocal microscopy using Zeiss LSM 710 ConfoCor 3 was employed and the data was analyzed using Zen (Black edition) software from Zeiss, Germany.

3.9.6. Mass spectrometry (MS)

Proteins have been loaded on SDS-gel, which ran only a short distance of 0.5 cm. After Commassie staining the total sample was cut out unfractionated and digested with Trypsin as described by Shevchenko et al. (Shevchenko, Andrej, et al. "In-gel digestion for mass spectrometric characterization of proteins and proteomes." Nature protocols 1.6 (2006): 2856.) carried out on the DigestProMSi robotic system (INTAVIS Bioanalytical Instruments AG).

Samples digested with trypsin were loaded on a cartridge trap column, packed with Acclaim PepMap300 C18, 5µm, 300Å wide pore (Thermo Fisher Scientific) and separated in a 180 min gradient from 3% to 40% ACN on a nanoEase MZ Peptide analytical column (300Å, 1.7 µm, 75 µm x 200 mm, Waters) and a UltiMate 3000 UHPLC system. Furthermore, eluting peptides have been analyzed by an online coupled Q-Exactive-HF-X mass spectrometer (Thermo Fisher Scientific) running in a data depend acquisition mode where one full scan was followed by up to 12 MSMS scans of eluting peptides.

Data analysis was carried out by MaxQuant (version 1.6.0.16, Tyanova, Stefka, TikiraTemu, and Juergen Cox. "The MaxQuant computational platform for mass spectrometry-based shotgun proteomics." Nature protocols 11.12 (2016): 2301) using an organism specific database extracted from Uniprot.org under default settings. Identification FDR cutoffs were 0.01 on peptide level and 0.01 on protein level. Match between runs option was enabled to transfer peptide identifications across raw files based on accurate retention time and m/z.

Quantification was done using a label free quantification approach based on the MaxLFQ algorithm (see Cox J.et al., Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. Mol Cell Proteomics, 2014, 13, pp 2513-2526). A minimum of 2 quantified peptides per protein was required for protein quantification.

Data have been further processed by in-house compiled R-scripts to plot data and the Perseus software package (version 1.6.2.5) using default settings for filtering, imputation of missing values and statistical analysis (Tyanova, Stefka, and Juergen Cox. "Perseus:

A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research." Cancer Systems Biology. Humana Press, New York, NY, 2018. 133-148.).

3.9.7. Enzyme-linked immunosorbent assay (ELISA)

96-well Nunc Polysorp plates were coated with coating medium (100 µL/well) and incubated overnight at 4°C the day before performing ELISA. On the following day, the coating medium was discarded from the plate wells thoroughly and incubated with blocking buffer (180 µL/well) for 1 h at 37°C. Afterwards, the blocking buffer was thoroughly removed bacterial lysate dilution series (2 µg/µL lysates in blocking buffer; 100 µL/well) containing desired overexpressed protein (GST-Tag or GST-E1C-Tag) was added to the plate wells. The plate was then incubated for 1 h at RT on a shaker. The bacterial lysate was then removed and washed three times with PBS-Tween® (200 µL/well). Afterwards, desired primary Ab diluted in blocking buffer, anti-Tag (1:50 dilution) or anti-E1C (1:5 dilution) (100 µL/well) was added and incubated for 1 h at RT in the dark. Then the plate was washed for three times with PBS-Tween® (200 µL/well) after removing the primary Ab. Then goat anti-mouse pox secondary Ab diluted in blocking buffer (1:10,000 dilution, 100 µL/well) was added and incubated for 1 h at RT in the dark. Then the plate was washed for three times with PBS-Tween® (200 µL/well) after removing the secondary Ab. Afterwards, 3,3',5,5'-tetramethylbenzidine chromogenic substrate solution (100 µL/well) was added to the plate wells and the peroxidase enzyme substrate reaction was stopped after 2-8 min by adding stop solution (50 µL/well). The absorbance was measured at 450 nm in the Multiskan™ FC microplate photometer.

GST-E1C-Tag ELISA (Results 4.2.1.2.) used in the screening of E1C Ab producing hybridoma clones was performed according to the above mentioned protocol and is also described Sehr et al., 2001[228]. The anti-IgG ELISA performed for determining the subtype of E1C mAb was a sandwich ELISA. Here, the ELISA plates were coated with anti-IgG antibodies (Jackson ImmunoResearch Laboratories, Inc., Cambridgeshire, UK) as the capture antibody followed by the addition of hybridoma clone supernatants. Anti-IgG

antibodies conjugated with horseradish peroxidase were used as the secondary antibodies (Jackson ImmunoResearch Laboratories, Inc., Cambridgeshire, UK).

4. Results

4.1. Effects of E1C overexpression

4.1.1. Effects of E1C overexpression on HPV16 transcripts

For testing the effect of E1C overexpression on HPV16 transcripts, E6*I, E1^E4 and E7 transcripts were quantified in four different HPV16-positive cell lines i.e. W12-epi, W12-int, CaSki and SiHa. E1C was overexpressed using LV-E1C lentiviral vector as described in Methods 3.4.. Corresponding control cells were transduced with LV lentiviral vector. After the selection of transduced cells by puromycin treatment, they were harvested and RNA was extracted. The extracted RNA from LV- and LV-E1C- transduced cells were subjected to a triplex RT-qPCR for the quantification of E6*I, E1^E4 and the housekeeping transcript ubC transcripts and to two singleplex RT-qPCRs for the quantification of E1C and E7 transcripts. The PCR programming conditions for both triplex and singleplex RT-qPCRs are explained in Methods 3.8.2.. The data obtained after RT-qPCRs were analyzed using LightCycler® 480 software and the quantification of all the transcripts were performed from the standard curves generated using the dilution series of corresponding *in vitro* transcripts.

A total of eight lentiviral transductions were performed for E1C overexpression in W12-epi and W12-int cell lines and five transductions in CaSki and SiHa. Transductions number 1 and 5 of W12-epi were excluded from data analysis, because of a mistake performed in the puromycin selection protocol and because the LV-E1C transduction did not work respectively. The details of no: of cells harvested post puromycin selection, total amount of RNA harvested, total no: of RT-qPCRs performed on each sample are given in supplementary information 9.5.. Tables 4.1. to 4.8. provide details about the copies per PCR obtained for E6*I, E1^E4, E7 and ubC for all the four cell lines in all the LV and LV-E1C transductions. They also summarize the ratio of copies per PCR for each transcript obtained in LV-E1C transduction relative to LV transduction (LV-E1C/LV ratio). E6*I E1^E4 and ubC copies per PCR were observed to be higher in transductions 6 – 8 for

W12-epi, 5 – 8 for W12-int, 2 – 5 for CaSki and SiHa, whereas E7 copies per PCR were more or less equivalent across transductions in all the cell lines (Tables 4.1. to 4.8.). However, the variation in LV-E1C/LV ratios was observed to be much less. E1C RNA levels post transduction were in the range of 4 – 12 copies/cell. No E1C RNA copies were found in LV-transduced cells.

Because RNA levels varied between transductions, median LV-E1C/LV ratios were calculated and values outside of 2x median and 0.5x median were defined as outliers and excluded from statistical analysis. The ubC LV-E1C/LV ratios for all the transductions in all the cell lines were inside the range of 0.5x median and 2x median except for the third transduction in W12-epi (Table 4.2.).

Table 4.1. **E6*I and E1^E4 copies per PCR for LV and LV-E1C transductions in W12-epi.**
LV-E1C/LV denotes the ratio of corresponding transcripts in LV-E1C- to LV- transduced cells.
The values in red indicate outliers.

Transduction	E6*I			E1^E4		
	LV	LV-E1C	LV-E1C/LV	LV	LV-E1C	LV-E1C/LV
2	75,000	72,500	0.97	2,025,000	1,472,500	0.73
3	9,600	52,800	5.50	68,000	561,600	8.26
4	17,600	41,600	2.36	164,000	518,400	3.16
6	2,277,500	3,930,000	1.73	1,805,000	903,750	0.50
7	868,750	1,013,750	1.17	1,815,000	1,568,750	0.86
8	893,750	1,103,750	1.23	1,775,000	1,265,000	0.71

The statistical analysis was performed on the LV-E1C/LV ratios of E6*I, E1^E4 and E7 transcript levels from all the transductions in all the cell lines by paired t test and two-

tailed p values were calculated. In W12-epi, E6*I (p = 0.12) and E7 (p = 0.30) levels in LV-E1C-transduced cells were not significantly different compared to LV-transduced cells (Figure 4.1. (A) and (C)). However, E1^E4 levels were significantly reduced (p = 0.03) with a mean of 30% (Figure 4.1. (B)).

In W12-int, E6*I (p = 0.75) and E7 (p = 0.33) levels in LV-E1C-transduced cells were not significantly different (Figure 4.1. (A) and (C)). E1^E4 levels were significantly reduced (p = 0.02) with a mean of 20% (Figure 4.1. (B)) when the value from third transduction was not considered and with a mean of 10% with no statistical significance (p = 0.25) when the value from third transduction was considered.

Table 4.2. **E7 and ubC copies per PCR for LV and LV-E1C transductions in W12-epi.**
LV-E1C/LV denotes the ratio of corresponding transcripts in LV-E1C- to LV- transduced cells.
The values in red indicate outliers.

Transduction	E7			ubC		
	LV	LV-E1C	LV-E1C/LV	LV	LV-E1C	LV-E1C/LV
2	52,500	37,500	0.71	878,750	1,282,500	1.46
3	800	84,000	105.00	18,400	72,800	3.96
4	24,000	60,800	2.53	32,800	80,800	2.46
6	55,000	80,000	1.45	1,011,250	846,250	0.84
7	73,750	70,000	0.95	2,287,500	2,301,250	1.01
8	68,750	86,250	1.25	1,942,500	1,973,750	1.02

In CaSki, E6*I (p = 0.26) and E7 (p = 0.38) levels in LV-E1C-transduced cells were not significantly different (Figure 4.1. (A) and (C)). However, E1^E4 levels were significantly reduced (p = 0.02) with a mean of 20% (Figure 4.1. (B)) when the value from fifth transduction was not considered and with a mean of 10% with no statistical significance (p = 0.43) when the value from fifth transduction was considered. In SiHa, E6*I (p = 0.18)

and E7 ($p = 0.55$) levels in LV-E1C-transduced cells were not significantly different (Figure 4.1. (A) and (C)).

Table 4.3. **E6*I and E1^E4 copies per PCR for LV and LV-E1C transductions in W12-int.**
LV-E1C/LV denotes the ratio of corresponding transcripts in LV-E1C- to LV- transduced cells.
The values in red indicate outliers.

Transduction	E6*I			E1^E4		
	LV	LV-E1C	LV-E1C/LV	LV	LV-E1C	LV-E1C/LV
1	318,750	555,000	1.74	3,500,000	6,816,250	1.95
2	233,750	243,750	1.04	5,008,750	4,653,750	0.93
3	342,500	430,000	1.26	4,495,000	5,802,500	1.29
4	497,500	322,500	0.65	7,037,500	4,335,000	0.62
5	21,771,250	20,708,750	0.95	30,366,250	23,783,750	0.78
6	21,847,500	17,770,000	0.81	25,866,250	22,502,500	0.87
7	4,662,500	4,055,000	0.87	17,171,250	13,812,500	0.80
8	5,723,750	1,270,000	0.22	22,282,500	3,538,750	0.16

Table 4.4. **E7 and ubC copies per PCR for LV and LV-E1C transductions in W12-int.**
LV-E1C/LV denotes the ratio of corresponding transcripts in LV-E1C- to LV- transduced cells.
The values in red indicate outliers.

Transduction	E7			ubC		
	LV	LV-E1C	LV-E1C/LV	LV	LV-E1C	LV-E1C/LV
1	107,500	275,000	2.56	472,500	335,000	0.71
2	90,000	95,000	1.06	2,036,250	1,998,750	0.98
3	327,500	382,500	1.17	572,500	417,500	0.73
4	515,000	325,000	0.63	802,500	402,500	0.50
5	212,500	160,000	0.75	1,953,750	1,748,750	0.90
6	217,500	232,500	1.07	2,067,500	1,925,000	0.93
7	331,250	237,500	0.72	8,827,500	9,287,500	1.05
8	363,750	95,000	0.26	7,580,000	7,525,000	0.99

Table 4.5. **E6*I and E1^E4 copies per PCR for LV and LV-E1C transductions in CaSki.**
LV-E1C/LV denotes the ratio of corresponding transcripts in LV-E1C- to LV- transduced cells.
The values in red indicate outliers.

Transduction	E6*I			E1^E4		
	LV	LV-E1C	LV-E1C/LV	LV	LV-E1C	LV-E1C/LV
1	470,000	377,500	0.80	7,380,000	5,845,000	0.79
2	2,780,000	2,662,500	0.96	4,770,000	4,096,250	0.86
3	372,500	1,830,000	4.91	508,750	2,917,500	5.73
4	1,060,000	857,500	0.81	5,218,750	3,962,500	0.76
5	886,250	947,500	1.07	3,915,000	4,756,250	1.21

Table 4.6 E7 **and ubC copies per PCR for LV and LV-E1C transductions in CaSki.**
LV-E1C/LV denotes the ratio of corresponding transcripts in LV-E1C- to LV- transduced cells.

Transduction	E7			ubC		
	LV	LV-E1C	LV-E1C/LV	LV	LV-E1C	LV-E1C/LV
1	547,500	505,000	0.92	2,645,000	2,547,500	0.96
2	55,000	70,000	1.27	189,368,750	167,791,250	0.89
3	12,500	25,000	2	138,728,750	151,351,250	1.09
4	150,000	120,000	0.80	55,616,250	57,818,750	1.04
5	117,500	125,000	1.06	53,875,000	68,370,000	1.27

Table 4.7. **E6*I and E1^E4 copies per PCR for LV and LV-E1C transductions in SiHa.**
LV-E1C/LV denotes the ratio of corresponding transcripts in LV-E1C- to LV- transduced cells.

Transduction	E6*I			E1^E4		
	LV	LV-E1C	LV-E1C/LV	LV	LV-E1C	LV-E1C/LV
1	195,000	290,000	1.49	E1^E4 transcript not present in SiHa cell line		
2	6,081,250	5,648,750	0.93			
3	5,246,250	6,982,500	1.33			
4	1,728,750	1,741,250	1.01			
5	1,673,750	1,827,500	1.09			

Table 4.8. **E7 and ubC copies per PCR for LV and LV-E1C transductions in SiHa.**
LV-E1C/LV denotes the ratio of corresponding transcript in LV-E1C- to LV- transduced cells.

Transduction	E7			ubC		
	LV	LV-E1C	LV-E1C/LV	LV	LV-E1C	LV-E1C/LV
1	400,000	365,000	0.91	615,000	845,000	1.37
2	37,500	80,000	2.13	27,322,500	33,992,500	1.24
3	100,000	70,000	0.70	25,396,250	36,045,000	1.42
4	300,000	312,500	1.04	17,096,250	12,240,000	0.72
5	265,000	295,000	1.11	11,448,750	11,682,500	1.02

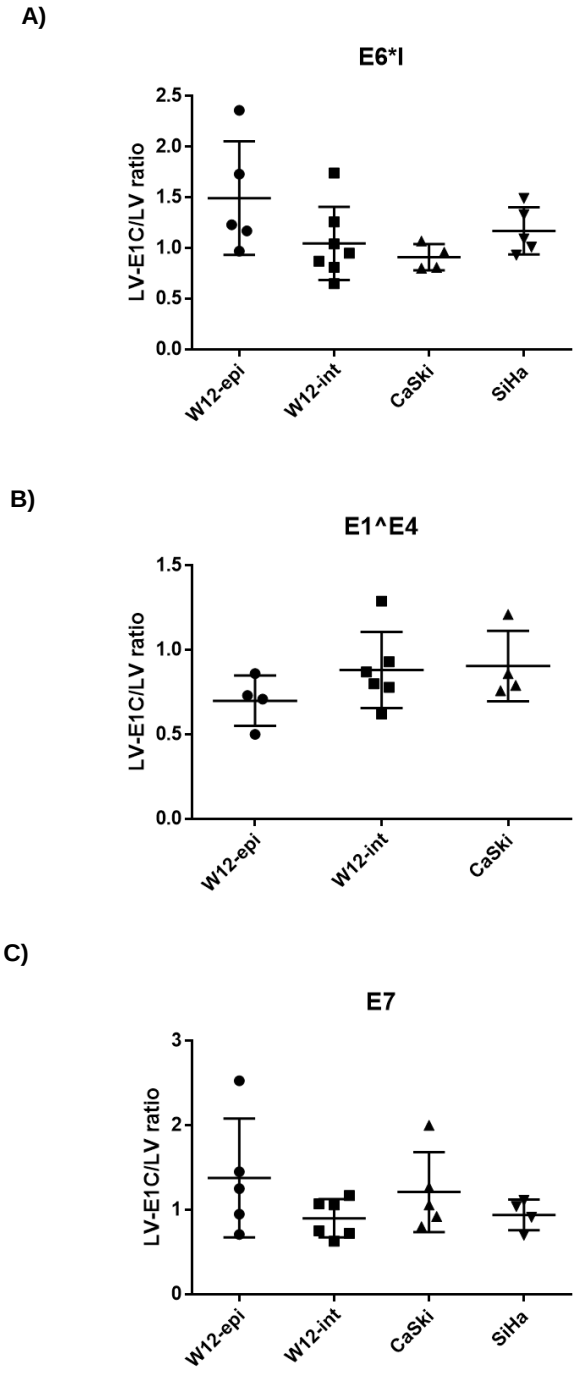


Figure 4.1. **Ratios of HPV16 transcript levels in LV-E1C to LV transductions from different cell lines.** LV-E1C/LV ratios are displayed with outliers excluded as defined in tables 4.1. - 4.8.. A) and C) show E6*I and E7 transcripts in W12-epi, W12-int, CaSki and SiHa cells respectively. B) shows E1^E4 transcript in W12-epi, W12-int and CaSki respectively. Mean together with standard deviation is shown for all the data points.

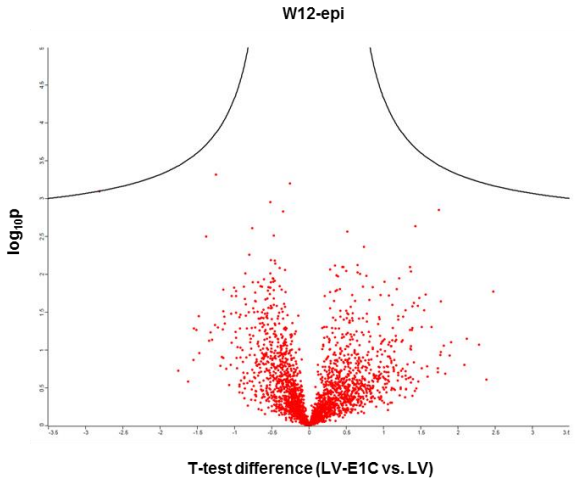
4.1.2. Effects of E1C overexpression on host or cellular proteome

In order to identify the cellular proteome changes in E1C overexpressed cells, MS was employed. For this purpose, LV-E1C-transduced W12-epi, W12-int, CaSki and SiHa cells together with their corresponding LV-transduced cells as controls were utilized. For W12-epi and W12-int, cells from third, fourth, seventh and eighth lentiviral transductions and for CaSki and SiHa, cells from first, fourth and fifth lentiviral transductions were used (details of transductions are explained in Results 4.1.).

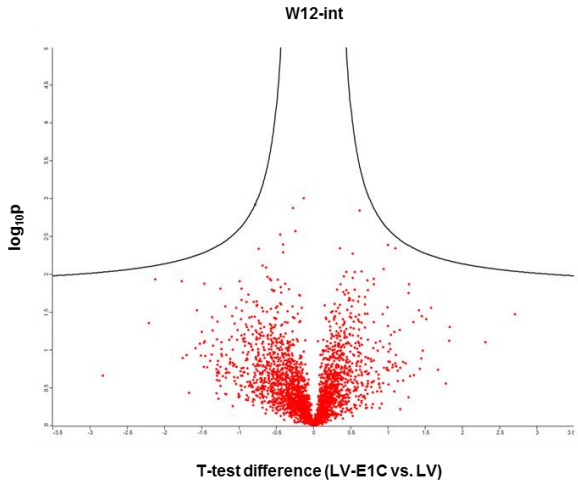
In total, 77883 peptides and 4899 proteins were identified by MS. The Perseus software was used for the statistical data analysis[218]. To check for statistically significant changes in the protein quantities in LV-E1C-transduced cells, Two-sample t-tests with $\log_2$ () transformed LFQ values were used after replacing missing values based on normal distribution. A permutation based false discovery rate threshold of 0.05 was used. S0, which gives the minimal fold change, was set to 0.1. This means that although a protein shows a good p-value, it would not be significant if the S0 is below 0.1.

The results from the Two-sample t-tests are displayed as volcano plots (Figure 4.2.). X-axis represents the t-test difference, which is equivalent to building a ratio between protein expressions in LV-E1C- to LV- transduced cells. Y-axis shows the p values, which determine the significance of the t-test differences for each protein. If the p value for a given protein was below 0.05, then it is considered as statistically significantly changing. S0 value helped to exclude the very significant hits, which showed almost no or zero t-test difference. All red data points below the black lines are non-significant. When the Two-sample t-tests were analyzed, no statistically significant changes in the protein quantities were observed in LV-E1C-transduced W12-epi, W12-int and SiHa cells compared to corresponding LV-transduced cells (Figure 4.2. (A), (B) and (D)). However, in LV-E1C-transduced CaSki cells, PDIA3 protein was found to be significantly changing when compared to LV-transduced cells (Figure 4.2. (C)). PDIA3 protein showed a t-test difference of 1.05 which is the $\log_2$ expression and $2^{1.05}$ gives the fold change, which means PDIA3 protein was two fold upregulated in LV-E1C-transduced CaSki cells. When Two-sample t-tests were performed for all the cell lines together, i.e. LV-E1C-transduced versus LV-transduced, no statistically significant hits have been found (Figure 4.2. (E)).

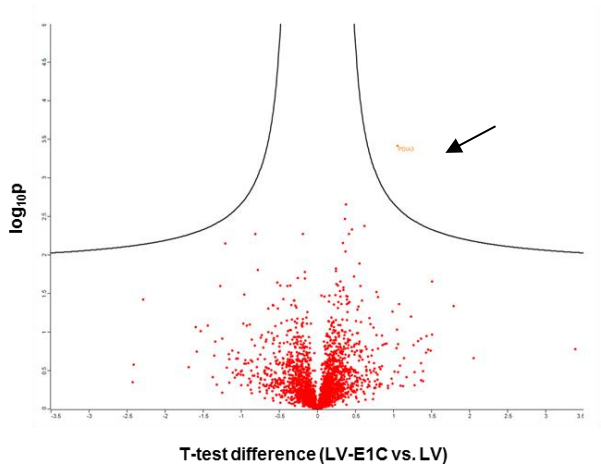
A)


W12-epi
log10p
T-test difference (LV-E1C vs. LV)

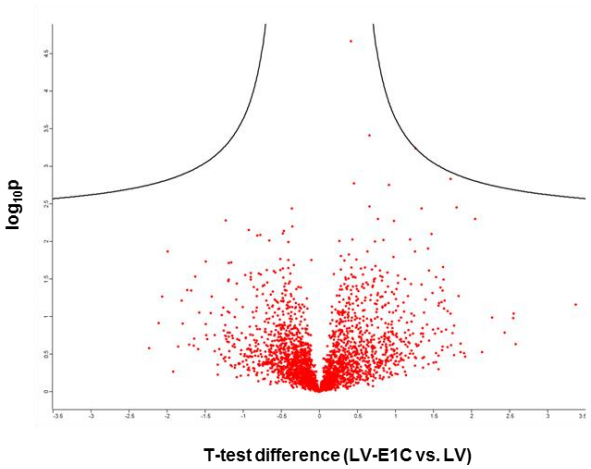
B)


W12-int
log10p
T-test difference (LV-E1C vs. LV)

C)


CaSki
log10p
T-test difference (LV-E1C vs. LV)

D)


SiHa
log10p
T-test difference (LV-E1C vs. LV)

E)

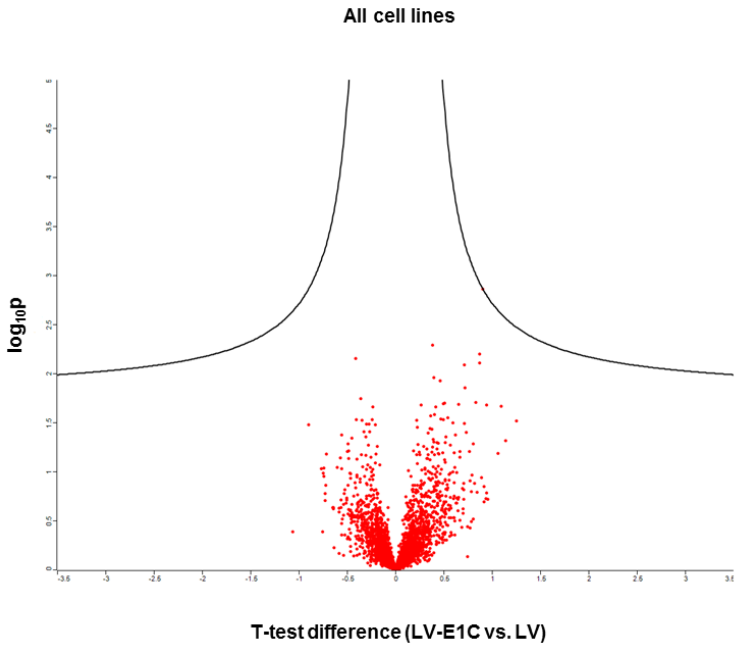


Figure 4.2. **Proteomic changes in LV-E1C- vs. LV-transduced cells represented using volcano plots.** X-axis shows the two sample t-test difference of protein quantities between LV-E1C- and LV- transduced cells and Y-axis shows the $\log_{10}$p values of the t-test differences of W12-epi (A), W12-int (B), CaSki (C), SiHa (D) and for all the cell lines together (E).

4.2. Detection of E1C protein

4.2.1. Generation of E1C mAb

A murine mAb specific for the detection of E1C protein was generated using hybridoma technology (see Methods 3.2.). The procedure involved design of antigen i.e. an E1C specific peptide, mouse immunization, production of hybridomas from the lymphocytes of mouse with the highest Ab titer and screening and cloning of hybridomas.

4.2.1.1. Design of an E1C peptide for mouse immunization

A short peptide of 10 aa length (E1C sp) was designed for immunizing mice to generate E1C mAb (Figure 4.4. (B)). E1C sp spanned the splice junction-derived peptide fusion of the putative E1C protein (Figure 4.4. (A)), and thus confers E1C specificity against E1fl protein. The first 5 aa of the putative E1C protein are from the amino terminus of E1 and the last 77 aa from the carboxy terminus of E1 (Figure 4.3.). To elicit significant antibody responses, 2 mg of chemically synthesized E1C sp was conjugated via a C-terminally added cysteine to keyhole limpet hemocyanin (KLH), a large carrier protein. KLH is a copper-containing glycoprotein, with a molecular weight of $3.9x10^5$ Da excluding the glycosylation. Because molluscs are phylogenetically distant from mammals, KLH is less likely to immunologically cross-react with human or mouse proteins. A total of 5 mg E1C sp was synthesized by Peptide Speciality Laboratories GmbH, Heidelberg, (www.peptid.de), out of which 2 mg was coupled with KLH and 3 mg was uncoupled.

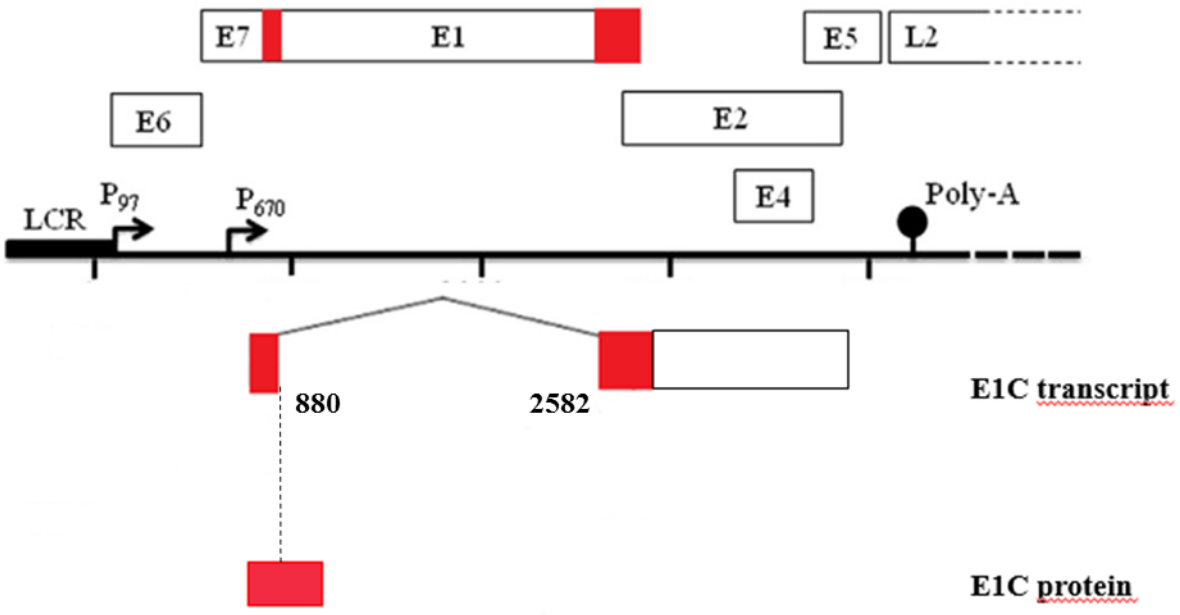


Figure 4.3. **Schematic diagram of E1C transcript and protein.** The diagram shows the HPV16 genome and the coding regions of the E1C transcript and protein. The regions marked in red in the E1 open reading frame are present in the E1C transcript. The splice junction 880^2582 of the E1C transcript is also shown. The regions marked in red in the E1C transcript encode the putative E1C protein (HPV16 genome edited from Höfler et al., 2015).

A.

5’- ATGGCTGCCTGCAGATTCTAGGTGGCCTTATTTACATAATAGATTGGTG
GTGTTTACATTTCCTAATGAGTTTCCATTTGACGAAAACGGAAATCCAGTGTA
TGAGCTTAATGATAAGAACTGGAAATCCTTTTTCTCAAGGACGTGGTCCAGA
TTAAGTTTGCACGAGGACGAGGACAAGGAAAACGATGGAGACTCTTTGCCA
ACGTTTAAATGTGTGTCAGGACAAAATACTAACACATTATGA – 3’

B.

MADPADSRWPYLHNRLVVFTFPNEFPFDENGNPVYELNDKNWKSFFSRTWSR
LSLHEDEDKENDGDSLPTFKCVSGQNTNTL

Figure 4.4. **E1C cDNA and protein sequences.** A) cDNA sequence of E1C showing the start (ATG, highlighted in green) and stop (TGA, highlighted in yellow) codons and the splice junction ‘GA’ (highlighted in red) embedded in the blue sequence coding for the 10 aa long E1C sp used for immunization. B) Sequence of the putative E1C protein showing the splice junction-derived peptide fusion ‘AD’ in red and the 10 aa long E1C sp used for mouse immunization in blue.

4.2.1.2. Mouse immunization, fusion and cloning

The sera from the E1C sp-immunized mice were tested for the presence of E1C Abs. For this purpose, lysates from *E. coli* overexpressing GST-E1C and GST-E1fl fusion proteins and from HeLa cells overexpressing GFP-E1C fusion protein were applied to gel electrophoresis and transferred to the WB membranes. E.coli lysate overexpressing GST protein and HeLa cellular lysate acted as corresponding negative controls. Expected bands of 35 and 36 kDa corresponding to GST-E1C and GFP-E1C fusion proteins respectively were visible in the blots stained with immune sera from C57black6N, but not from BALB/c mice (Figure 4.5. (A)). An additional prominent 25-28 kDa band was observed in the GST-E1C lane, probably a product of N- or C-terminal fusion protein degradation or premature translation stop shortly after the E1C specific peptide sequence (Figure 4.5. (A)). No bands were visible in the control lanes without E1C peptide sequence

and also not in the GST-E1fl protein lanes as well, confirming the E1C-specificity of the antibodies. GST protein and HeLa cellular lysate lanes. In addition, no bands around 70 kDa were visible in the GST-E1fl protein lanes as well, confirming the specificity of E1C antibodies produced. The blot stained with pre-serum taken before immunization served as a negative control for sera staining and no bands were visible except for one around 35 kDa (Figure 4.5. (A)) which is unexpected and remains unexplained.

E1C-immunized mouse C57black6N-1 was chosen for fusion. A total of 370 primary hybridoma clones were screened for the presence of E1C Abs. Screening was performed using a standard mAb screening ELISA established in the monoclonal antibody core facility with E1C sp as antigen. Six clones were tested positive for the presence of E1C Abs, which were then confirmed by GST-E1C-Tag ELISA and WB with GST-E1C and GFP-E1C fusion proteins.Both GST-E1C and GFP-E1C fusion proteins were visible in the blots stained with cell culture supernatant of the six clones (Figure 4.5. (B)). After two sub-cloning rounds of these 6 clones and testing 277 supernatants one sub-clone (61/1/2) which showed highest reactivity in GST-E1C-Tag ELISA and reacted in WB (Figure 4.5. (C)) and IF (Figure 4.13. (B)) was chosen for the production of E1C mAb. In GST-E1C-Tag ELISA, the mean absorbance measured for GST-E1C antigen stained with E1C mAb was 0.85 and GST-Tag antigen (positive control) stained with α-Tag mAb was 1.03. Thus, E1C mAb showed more than half of the reactivity when compared to the positive control.

After performing an anti-IgG ELISA, E1C mAb was found to be of IgG1 subtype (data not shown). Optimal working dilutions of the purified E1C mAb were determined in dilution series and found to be 1:5 for ELISA (lowest concentration giving at least half of the maximal signal intensity as that of positive control) and WB and 1:3. Thus, a monoclonal antibody suited for the specific detection of HPV16 E1C protein was successfully produced.

4.2.2. Detection of overexpressed E1C protein

The newly generated E1C mAb was utilized for the detection of the small putative 9 kDa E1C protein overexpressed without additional protein fusions in HEK-293T and W12 cells.

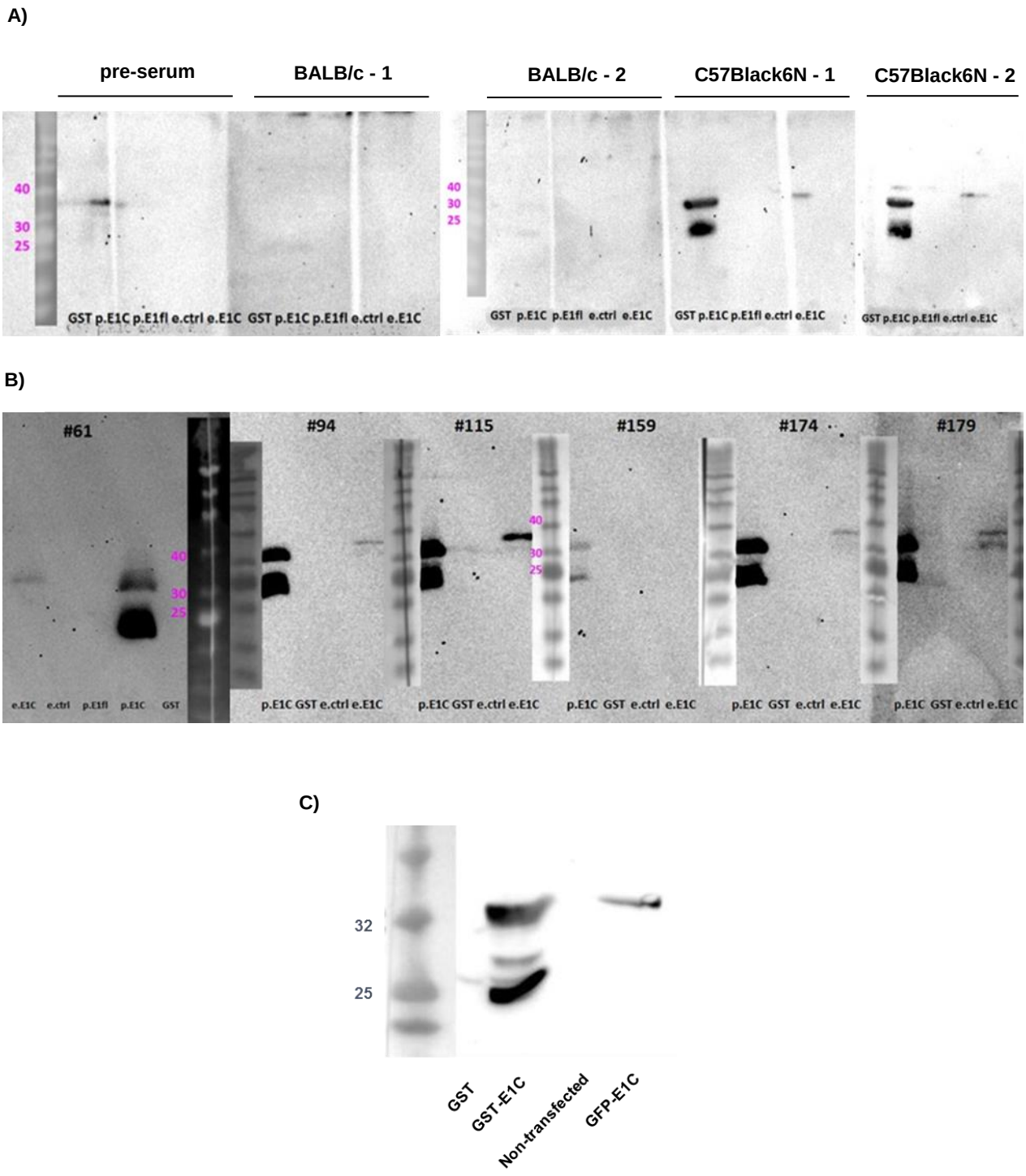


Figure 4.5. **E1C antibody reactivity in western blot.** Western blot lanes contain lysates from *E. coli* overexpressing GST-Tag (GST), GST-E1C (p.E1C), or GST-E1fl (p.E1fl) fusion protein and lysates from HeLa cells without (e.ctrl) or with (e.E1C) GFP-E1C fusion protein overexpression. Prestained Protein Ladder, Broad Range (10-230 kDa) was used

as size marker in all blots. A) Reactivity of mouse pre-serum and immune sera of E1C sp-immunized BALB/c and C57Black6N mice (2 each). B) Reactivity of supernatants of 6 hybridoma clones (#61, #94, #115, #159, #174 and #179) from E1C sp-immunized mice. C) Reactivity of the finally selected E1C hybridoma sub-clone 61/1/2. GST and GST-E1C denote the *E. coli* lysates overexpressing GST-Tag and GST-E1C fusion proteins. Non-transfected denotes HeLa cells without any protein overexpression and GFP-E1C denotes HeLa cells overexpressing GFP-E1C fusion protein.

When the western blots were performed under standard conditions, E1C protein was not detectable in p2-E1C-transfected HEK-293T cellular lysates while the 36 kDa GFP-E1C protein was visible in p1-His-GFP-E1C-transfected HEK-293T cellular lysates (data not shown). After literature survey and several personal communications with expert colleagues at DKFZ, the WB protocol was modified to enhance the detection of small proteins and the 10 kDa ABCB5 protein carrying a 6x-His tag was used as positive control. The polyacrylamide gels were enlarged to around 30 cm x 30 cm to allow good separation of small proteins, the loading well capacity was increased to around 50 µL to increase the amount of loaded protein. Protein transfer was performed using a transfer buffer without SDS and at lower current for shorter time (250 mA for 30 min) to reduce the loss of E1C protein through the nitrocellulose membrane.

4.2.2.1. Production and purification of ABCB5

Bacterial lysate overexpressing truncated ABCB5 protein was purified by affinity chromatography. The different eluates obtained during the purification process are Eluate 1 to Eluate 7. The concentrations of the eluates are measured using a NanoDrop™ spectrophotometer at 280 nm and ranged between 0.85 – 0.11 mg/mL (Table 4.9.) and were tested for the presence of ABCB5 by WB (Figure 4.6.). All the eluates contained ABCB5 protein, where Eluates 1 – 4 contained more protein compared to Eluates 5 – 7 (Figure 4.6.). Thus, a positive control protein for the detection of 9 kDa E1C protein was successfully established.

Table 4.9. **Protein concentrations of the eluates from ABCB5 purification.**

Eluate	Concentration of eluate (mg/mL)
Eluate 1	0.85
Eluate 2	0.33
Eluate 3	0.15
Eluate 4	0.84
Eluate 5	0.13
Eluate 6	0.12
Eluate 7	0.11

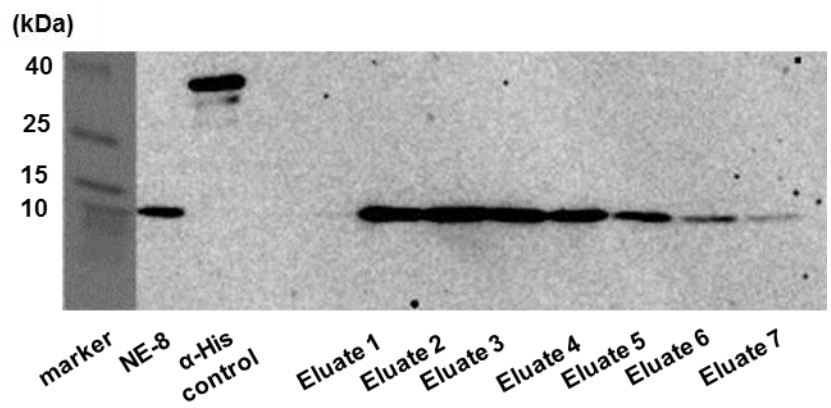


Figure 4.6. **Testing of all the eluates from ABCB5 purification in WB.** 50 ng of Eluate 1 to eluate 7 were applied to WB from left to right. The blot was stained with α-His Ab. NE-8 is a purified 6x-His tagged protein of 8 kDa which was loaded as the positive control and 50 ng of NE-8 was loaded. α-His control was a HeLa cellular lysate containing a His-tagged protein, which served as the positive control for α-His Ab, and 20 µg was applied. Spectra™ Multicolor Low Range Protein Ladder was used as the marker.

4.2.2.2. Testing the sensitivity of the modified WB protocol for the detection of small proteins

The sensitivity of WB protocol for small proteins was tested to analyze the least amount of purified protein detected with the protocol.

The sensitivity test was performed using ABCB5 and GFP-E1C fusion protein. Eluate 1 of ABCB5 (Table 4.9.) was chosen for the sensitivity test. For the preparation of GFP-

E1C fusion protein, HeLa cells were transfected with p1-His-GFP-E1C plasmid by PEI transfection. The samples were prepared for WB as described in Methods section 3.9.4.. A 1:10 dilution series ranging from 330 – 0.0033 ng were used to prepare ABCB5 samples whereas a dilution ranging from 10 – 0.1 µg was used for GFP-E1C sample. ABCB5 blot was stained with α-His Ab and GFP-E1C blot with α-E1C Ab.

ABCB5 band was observed at 10 kDa in 330 ng and 33 ng lanes and not in the rest of the lanes (Figure 4.7. (A)). For GFP-E1C, 36 kDa bands were observed in 10 µg and 5 µg lanes and not in the rest of the lanes (Figure 4.7. (B)). Therefore, the sensitivity or the detection limits of the modified WB protocol were 33 ng of purified ABCB5 protein. GFP-E1C protein from p1-His-GFP-E1C-transfected HeLa cellular lysate could be detected in a minimum of 5 µg total lysate protein. Hence the WB protocol was capable of detecting 33 ng of small proteins of size 10 kDa and E1C in 5 µg of cellular lysate protein.

4.2.2.3. Detection of overexpressed E1C protein in HEK-293T cells

For the detection of overexpressed E1C protein in HEK-293T cells, E1C and GFP-E1C were overexpressed using p2-E1C and p1-His-GFP-E1C plasmids in HEK-293T cells by PEI transfection. The negative control samples were HEK-293T cells transfected with the corresponding empty vector backbones, p2 and p1-His-GFP.

4.2.2.3.1. E1C mRNA quantification in p2-E1C-transfected HEK-293T cells

RNA was extracted from p2- and p2-E1C-transfected HEK-293T cells according to the protocol described in Methods 3.8.1.. The overexpressed E1C mRNA was quantified by singleplex RT-qPCR as explained in Methods 3.8.2.. 100 ng of RNA was applied to the PCR. The absolute E1C mRNA copies were analyzed using a 10^6 – 10^0 copies/µL dilution series of p2-E1C plasmid (Figure 4.8.).

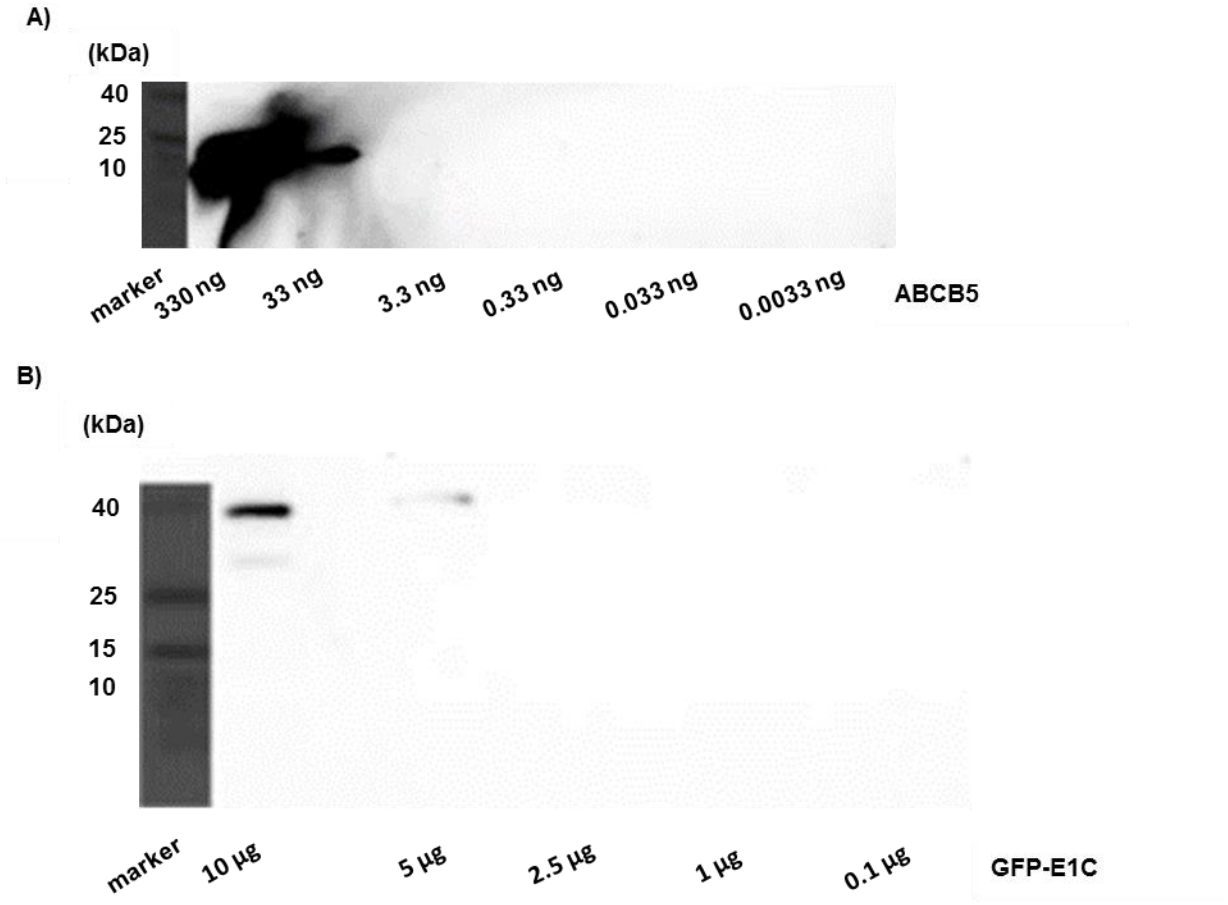


Figure 4.7. **Sensitivity of protein detection in the modified WB protocol**. A) Different amounts of ABCB5 protein (Eluate 1) were applied to WB and detected by 6x-His immunostaining. B) Different amounts of total lysate protein from HeLa cells overexpressing GFP-E1C fusion protein were applied to WB and detected by E1C immunostaining. Spectra™ Multicolor Low Range Protein Ladder was used as the marker.

The sensitivity of the PCR was 10^3 copies/µL and the specificity of the PCR was tested by applying 10^6 copies/µL of p2 plasmid. No Cp values were detected for p2 plasmid. The efficiency of the PCR was 2.54. Both the standards and samples were applied as duplicates.

The total no: of E1C mRNA copies per PCR was calculated from the following formula:

Total no: of E1C mRNA copies per PCR =2.71828182845904^((a-b)/c)

Where, a = average Cp value of the sample duplicates, b = constant from the standard curve, c = natural log of slope

The total no: of cells applied to the PCR was 5000 and was calculated under the assumption that one mammalian cell consists of 20 pg RNA. $9.24x10^7$was the total number of E1C mRNA copies detected per PCR in p2-E1C-transfected HEK-293T RNA sample and no E1C mRNA copies were detected in p2-transfected HEK-293T RNA sample. Therefore, the absolute no: of E1C mRNA copies per cell for p2-E1C-transfected HEK-293T cells was calculated as $1.85x10^4$.

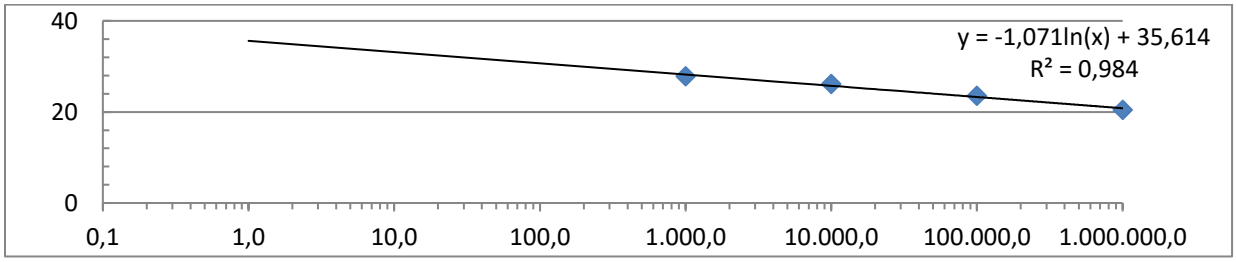


Figure 4.8. **Standard curve for q-PCR amplification of p2-E1C plasmid DNA.** X-axis was plotted with copies/µL and y-axis with Cp values. The slope is calculated from the curveas natural log and is given on the top right hand side. Efficiency of the PCR is calculated from base-10 log of slope, Efficiency = $10^{(-1/\text{slope lg }(x))}$. R^2 is the correlation coefficient.

4.2.2.3.2. E1C protein detection in p2-E1C-transfected HEK-293T cells

For the detection of overexpressed E1C protein in HEK-293T cells, WB samples were prepared from cellular lysates of p2-, p2-E1C-, p1-His-GFP-, and p1-His-GFP-E1C-transfected HEK-293T cells. GFP-E1C protein from p1-His-GFP-E1C-transfected HEK-293T cells acted as the positive control for E1C mAb staining and 50 ng of ABCB5 (in combination with α-His staining) for 10 kDa protein detection (Figure 4.9.).

A 9 kDa band corresponding to E1C protein was not detectable in the p2-E1C-transfected HEK-293T cellular lysate, whereas a 36 kDa GFP-E1C protein band was visible in the p1-His-GFP-E1C-transfected HEK-293T cellular lysate (Figure 4.9.).

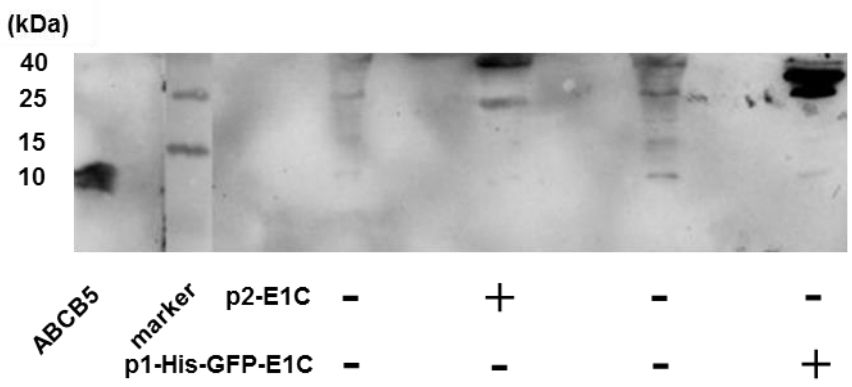


Figure 4.9. **Detection of overexpressed E1C protein in HEK-293T cells by WB.** Lysates of HEK-293T cells (100 µg total lysate protein) transfected with p2 (denoted as – in the p2-E1C row), p2-E1C (denoted as + in the p2-E1C row), p1-His-GFP (denoted as – in the p1-His-GFP-E1C row) and p1-His-GFP-E1C (denoted as + in the p1-His-GFP-E1C row) were applied to the western blot together with 50 ng ABCB5, loaded in the lane left to the Spectra™ Multicolor Low Range Protein Ladder.

4.2.2.4. Detection of overexpressed E1C protein in W12-epi and W12-int cells

For the detection of overexpressed E1C protein in W12-epi and W12-int cells, E1C was overexpressed using LV-E1C plasmid by lentiviral transduction. The negative control samples were cells transfected with the corresponding empty vector LV.

4.2.2.4.1. E1C mRNA quantification in LV-E1C-transduced W12-epi and W12-int cells

RNA was extracted from LV- and LV-E1C-transduced W12-epi and W12-int cells. The overexpressed E1C mRNA was quantified by singleplex E1C RT-qPCR. 100 ng of RNA was applied to the PCR. The absolute E1C mRNA copies were analyzed using a 10^6–10^0 copies/µL dilution series of LV-E1C plasmid (Figure 4.10.). The sensitivity of the PCR was 10copies/µL and the specificity of the PCR was tested by applying 10^6copies/µL of LV plasmid. No Cp values were detected for LV plasmid. The efficiency of the PCR was 2.11. Both the standards and samples were applied as duplicates.

The total no: of cells applied to the PCR was 5000. $7.7x10^4$was the total number of E1C mRNA copies detected per PCR in LV-E1C-transduced W12-epi RNA sample and $4.3x10^4$copies per PCR in LV-E1C-transduced W12-int RNA sample. No E1C mRNA copies were detected in LV-transduced W12-epi and W12-int RNA samples. Therefore, the absolute no: of E1C mRNA copies per cell was calculated as 16 for LV-E1C-transduced W12-epi cells and 9 for LV-E1C-transduced W12-int cells.

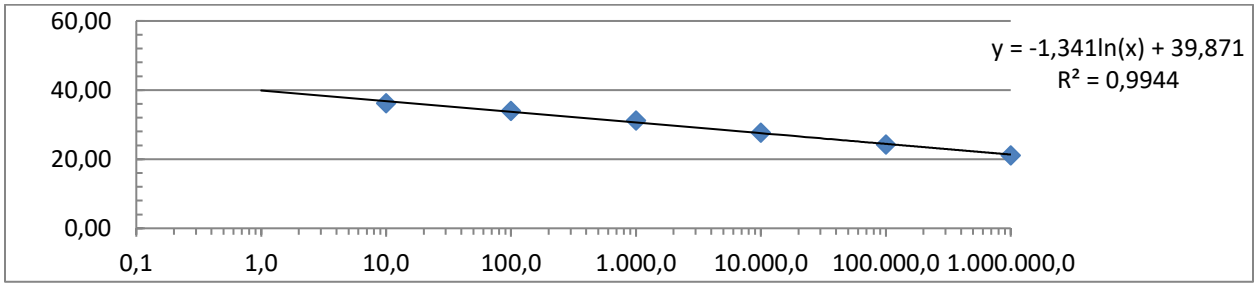


Figure 4.10. **Standard curve for q-PCR amplification of LV-E1C plasmid DNA.** X-axis was plotted with copies/µL and y-axis with Cp values. The slope is calculated from the curve as natural log and is given on the top right hand side. Efficiency of the PCR is calculated from base-10 log of slope, Efficiency = $10^{(-1/\text{slope lg }(x))}$. R^2 is the correlation coefficient.

4.2.2.4.2. E1C protein detection in LV-E1C-transduced W12-epi and W12-int cells

For the detection of overexpressed E1C protein in W12-epi and W12-int cells, WB samples were prepared from cellular lysates of LV- and LV-E1C-transduced W12-epi and W12-int cells. GFP-E1C protein from p1-His-GFP-E1C-transfected HEK-293T cells acted as the positive control for E1C mAb staining and ABCB5 for 10 kDa protein detection (Figure 4.11.).

A 9 kDa band corresponding to E1C protein was not detectable in the LV-E1C-transduced W12-epi and W12-int cellular lysates, whereas a 36 kDa GFP-E1C protein band was visible in the p1-His-GFP-E1C-transfected HEK-293T cellular lysate (Figure 4.11.).

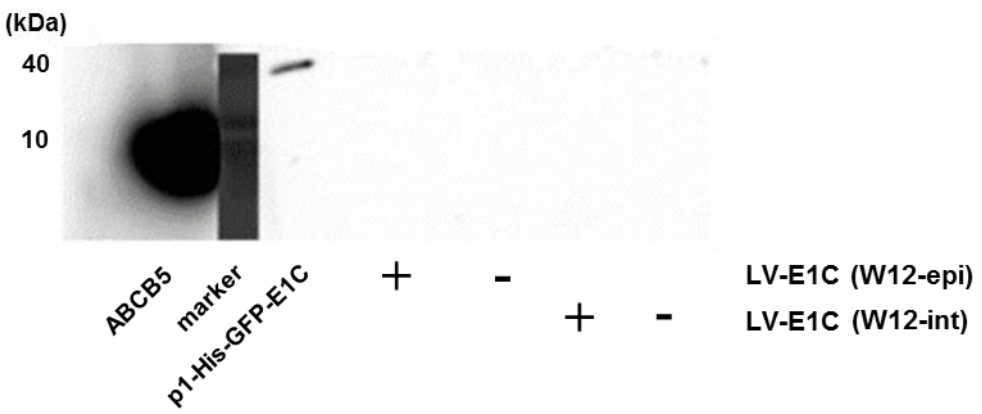


Figure 4.11. **Detection of overexpressed E1C protein in W12 cells by WB.** Lysates of HEK-293T cells (30 µg total lysate protein) transfected with p2 (denoted as – in the p2-E1C row), p2-E1C (denoted as + in the p2-E1C row), p1-His-GFP (denoted as – in the p1-His-GFP-E1C row) and p1-His-GFP-E1C (denoted as + in the p1-His-GFP-E1C row) were applied to the western blot together with ABCB5, loaded in the lane left to the Spectra™ Multicolor Low Range Protein Ladder.

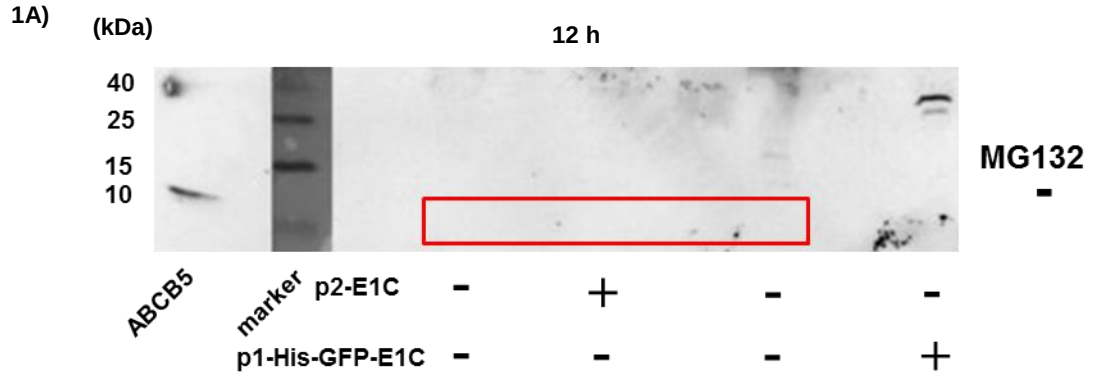
4.2.2.5. Detection of overexpressed E1C protein after proteasome inhibition

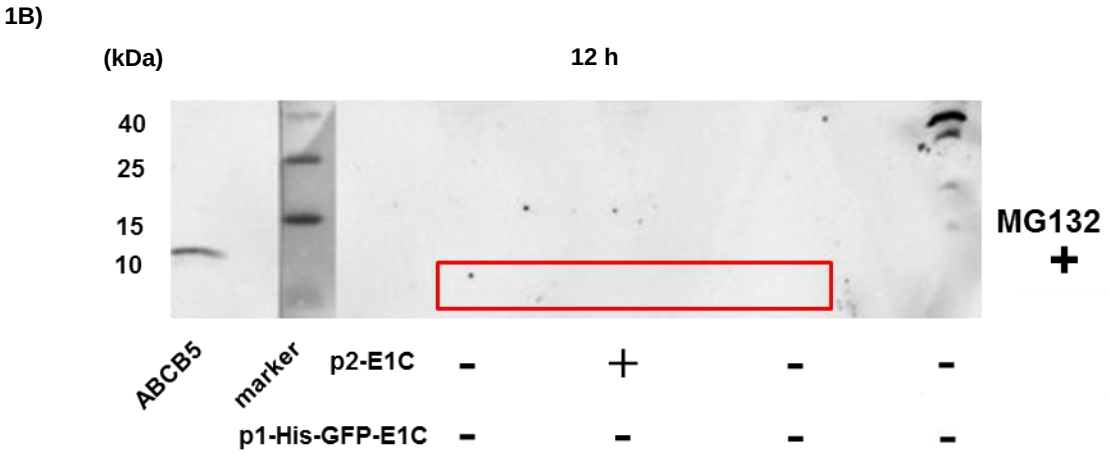
4.2.2.5.1. Detection of overexpressed E1C protein after proteasome inhibition: Western blot

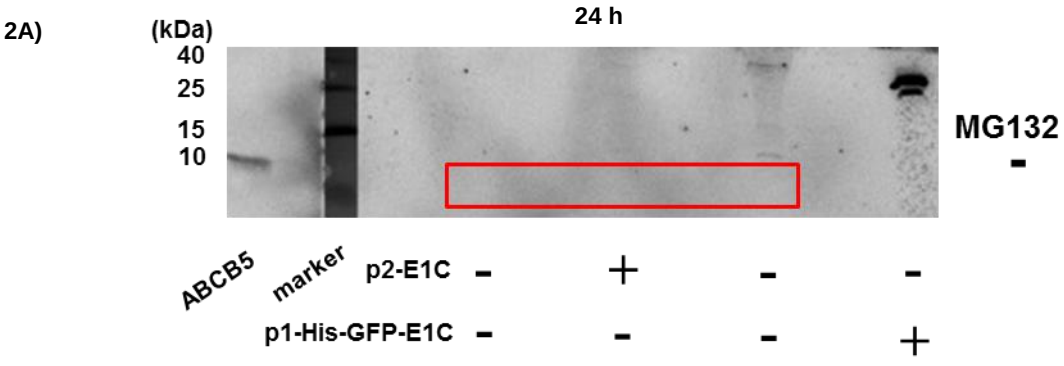
In order to detect the overexpressed E1C protein after proteasome inhibition, HEK-293T cells were treated with and without the proteasome inhibitor MG132 for 6 h prior to their harvest at 12, 24 and 48 h post transfection with p2 and p2-E1C. Afterwards, cellular lysates were prepared and applied to WB.

A 9 kDa band corresponding to E1C protein was detectable 48 h post transfection when treated with MG132 in the p2-E1C-transfected HEK-293T cells (Figure 4.12. (3B) and (3D)), and not in the cases of cells harvested post 12 and 24 h transfection (Figure 4.12. (1B) and (2B) respectively). No 9 kDa band was visible in the p2-E1C-transfected HEK-293T cells which were not treated with MG132 and were harvested at 12, 24 and 48 h

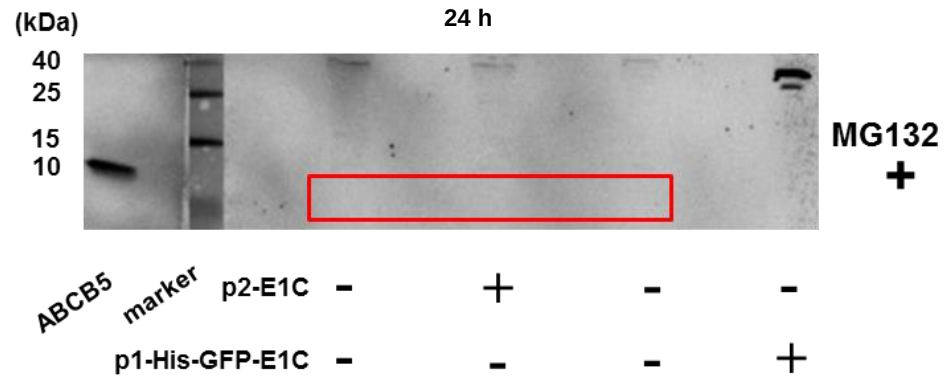
post transfection (Figure 4.12. (1A), (2A), (3A) and (3C). 36 kDa GFP-E1C protein band was visible in all the samples.



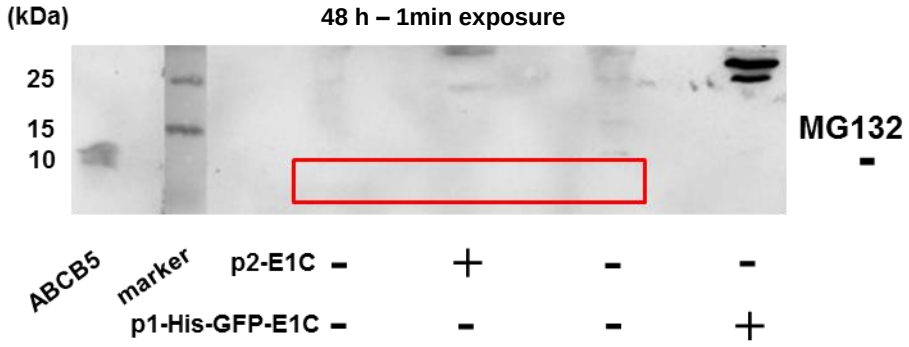




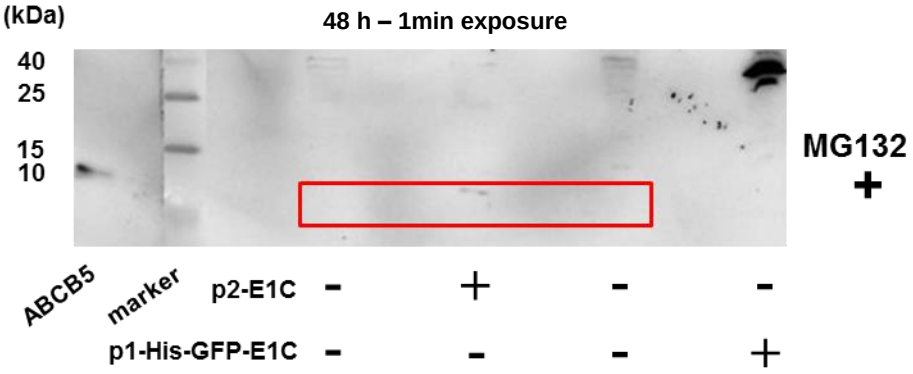
2B)


(kDa)
24 h
40
25
15
10
MG132
+
ABCB5
marker
p2-E1C - + - -
p1-His-GFP-E1C - - - +

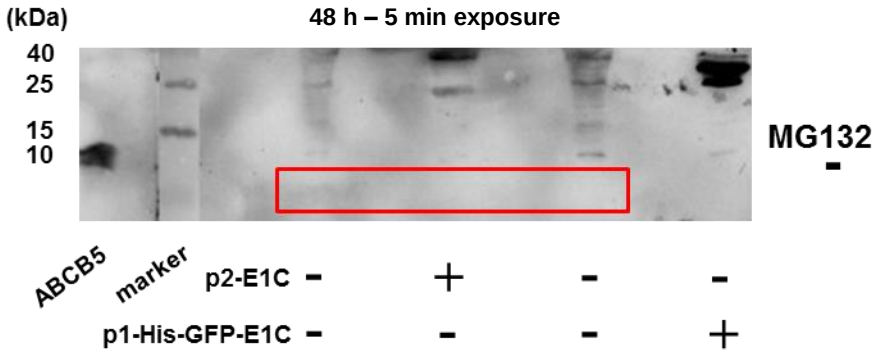
3A)


(kDa)
48 h – 1min exposure
25
15
10
MG132
-
ABCB5
marker
p2-E1C - + - -
p1-His-GFP-E1C - - - +

3B)


(kDa)
48 h – 1min exposure
40
25
15
10
MG132
+
ABCB5
marker
p2-E1C - + - -
p1-His-GFP-E1C - - - +

3C)



3D)

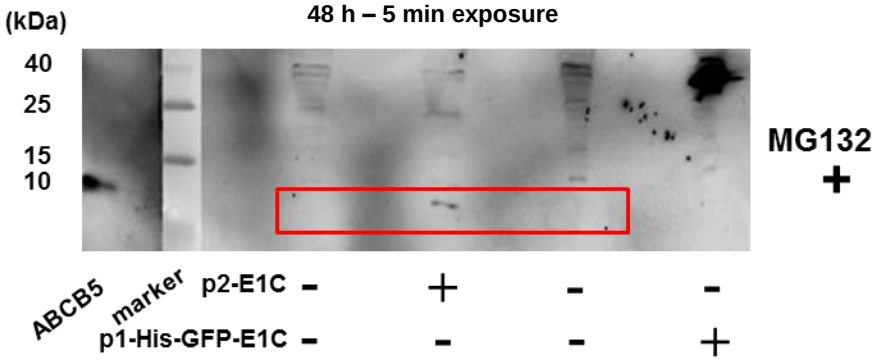


Figure 4.12. **WB detection of overexpressed E1C protein after proteasome inhibition.** HEK-293T cell lysates (100 µg total lysate protein) transfected with p2 (denoted as – in the p2-E1C row), p2-E1C (denoted as + in the p2-E1C row), p1-His-GFP (denoted as – in the p1-His-GFP-E1C row) and p1-His-GFP-E1C (denoted as + in the p1-His-GFP-E1C row) and treated with (1B, 2B, 3B, 3D) and without MG132 (1A, 2A, 3A,3C) have been applied to the western blot together with 50 ng ABCB5, which is loaded in the left lane to the Spectra™ Multicolor Low Range Protein Ladder. 1A and 1B are blots of lysates prepared post 12h transfection, 2A and 2B post 24h transfection and 3A, 3B, 3C and 3D post 48h transfection. 3A and 3B blots were exposed for 1 min during development and 3C and 3D for 5 min. 1A – 2B blots were exposed for 5 min. The red frame on the blots marks the region where the endogenous 9 kDa E1C protein would migrate to. E1C band was observed in 3B and 3D blots where p2-E1C transfected HEK-293T cellular lysates were loaded. 3D blot had a more intense E1C band compared to 3B blot. Also, GFP-E1C bands were intense in p1-His-GFP-E1C-transfected HEK-293T cellular lysates with MG132 treatment compared to the ones without treatment. All blots of lysates without MG132 treatment showed no 9 kDa protein.

E1C RNA was also quantified in p2 and p2-E1C-transfected and MG132-treated HEK-293T cells by singleplex E1C RT-qPCR. 100 ng of total RNA were applied to the PCR, the equivalent of 5000 cells. All cells overexpressed E1C and the abundance of E1C RNA increased with the increase in harvest time point from 12 to 48 h post transfection (Table 4.10). MG132-treated cells showed half the number of E1C RNA copies/cell as compared to DMSO-treated cells at 12 h and 24 h time points.

Table 4.10. **E1C RNA copies/cell in p2-E1C-transfected and MG132 treated HEK-293T cells.**

	E1C RNA copies/cell		
HEK-293T cells	**p2-E1C (12h)**	**p2-E1C (24h)**	**p2-E1C (48h)**
DMSO	$6x10^4$	$2x10^5$	$4x10^5$
MG132	$6x10^4$	$1x10^5$	$2x10^5$

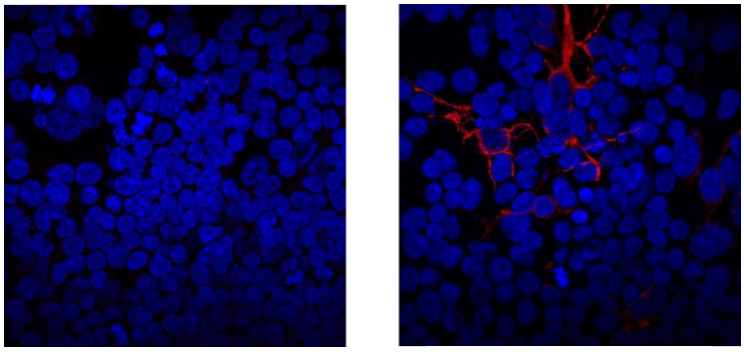
4.2.2.5.2. Detection of overexpressed E1C protein after proteasome inhibition: Immunofluorescence

In order to detect overexpressed E1C protein in p2-E1C-transfected and MG132-treated HEK-293T cells by IF, HEK-293T cells were treated with DMSO or MG132 for 6 h prior to their harvest at 48 h post transfection with p2 or p2-E1C. Afterwards, cells were fixed with 3.7% PFA and prepared for IF as described in Methods 3.9.5..

Before proceeding to E1C protein detection, the IF protocol was tested for the detection of small proteins of about 9 kDa size. For this purpose, HEK-293T cells transfected with pcDNA3.1(+) and pcDNA3.1(+)-CASIMO were analysed by IF. pcDNA3.1(+)-CASIMO encodes CASIMO protein tagged at the C-terminus with a Flag epitope and is of 9 kDa size[192]. CASIMO protein was successfully detected in the cytoplasm using α-Flag Ab (Figure 4.13. (A)). Thus, the IF protocol was suitable for the detection of E1C protein. As a positive control, HEK-293T cells transfected with p1-His-GFP and p1-His-GFP-E1C were also analyzed by IF. The cells were stained with α-E1C mAb and GFP-E1C protein

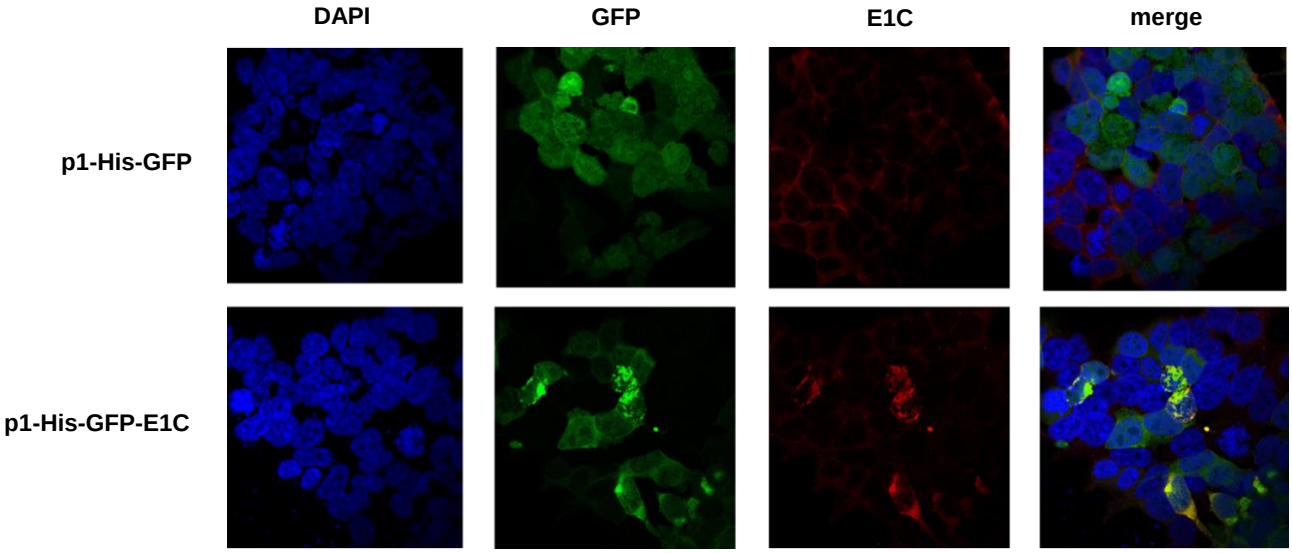
was found to be localized in the cytoplasm (Figure 4.13. (B)). When p2- and p2-E1C-transfected HEK-293T cells were stained with α-E1C mAb, E1C protein was found to be localized in the cytoplasm in the case of MG132-treated p2-E1C transfected HEK-293T cells and not in the p2-transfected HEK-293T cells (Figure 4.13. (C)). However, not all GFP-positive cells were stained positive for E1C. When ten images of MG132 treated p2-E1C transfected HEK-293T cells were analyzed from one microscopic cover slip, only 31 cells were found to be E1C-positive out of 175 GFP-positive cells. i.e. only 20% of the p2-E1C transfected cells were stained positive for E1C in IF.

A)



pcDNA3.1(+) **pcDNA3.1(+)-CASIMO**

B)



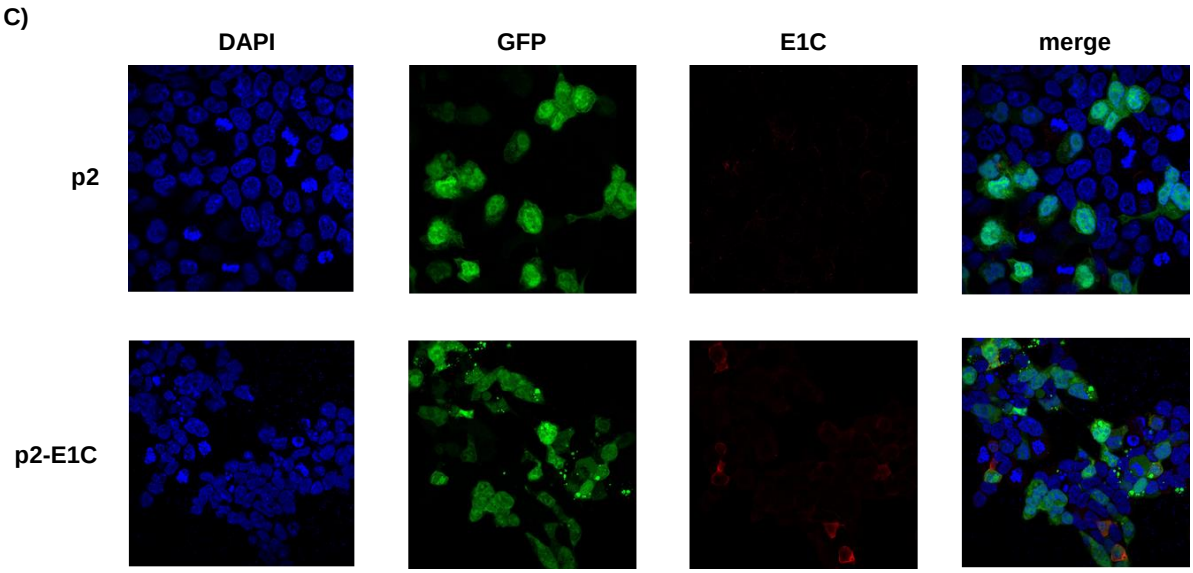


Figure 4.13. **IF detection of overexpressed E1C protein after proteasome inhibition.** A) HEK-293T cells transfected with pcDNA3.1(+) (left image) and pcDNA3.1(+) –CASIMO (right image) and stained with DAPI and α-Flag Ab are shown. Nuclei were stained blue and CASIMO protein was stained red (right image). B) HEK-293T cells transfected with p1-His-GFP (top panel) and p1-His-GFP-E1C (bottom panel) stained with DAPI and α-E1C Ab are shown. Nuclei were stained blue and GFP-E1C fusion protein visualized through green and red channels can be seen under 'GFP' and 'E1C' in p1-His-GFP-E1C panel. C) HEK-293T cells transfected with p2 (top panel) and p2-E1C (bottom panel) stained with DAPI and α-E1C Ab are shown. Nuclei were stained blue and E1C protein visualized through red channel is visible under 'E1C' in p2-E1C panel. As p2 and p2-E1C also encodes for GFP protein, GFP-positive cells are shown under 'GFP'. A merge of all the channels can be seen under 'merge' in both B) and C).

4.2.2.5.3. Detection of overexpressed E1C protein after proteasome inhibition: Mass Spectrometry

In order to detect overexpressed E1C protein in cells by MS, cellular lysates were prepared from p2-E1C-transfected HEK-293T cells with and without MG132 treatment prior to their harvest 12, 24 and 48 h post transfection. Lysate from p1-His-GFP-E1C transfected HEK-293T cells harvested 48 h post transfection without MG132 pre-treatment was used as positive control. Afterwards, cells were subjected to MS analysis as described in Methods 3.9.6..

A total of 69,796 peptides and 4574 proteins were identified by MS. Out of which four peptides of varying length and specific for HPV16 E1 or E1C protein sequence were identified. LFQ intensity which is a proxy for protein abundance was calculated taking all

the four peptides together in all the cellular lysates as described in Methods 3.9.6.. A E1fl- or E1C–specific peptide was detected, in p2-E1C-transfected and MG132-treated HEK-293T cells harvested 12, 24 and 48 h post transfection (Figure 4.14.) with intensities at least 50 times lower than in p1-His-GFP-E1C-transfected cells. No E1fl- or E1C-specific peptide was detected in p2-E1C transfected HEK-293T cells without MG132 and harvested 12 and 48 h post transfection while in cells harvested 24 h post transfection, a signal close to the sensitivity threshold was detectable (Figure 4.14.).

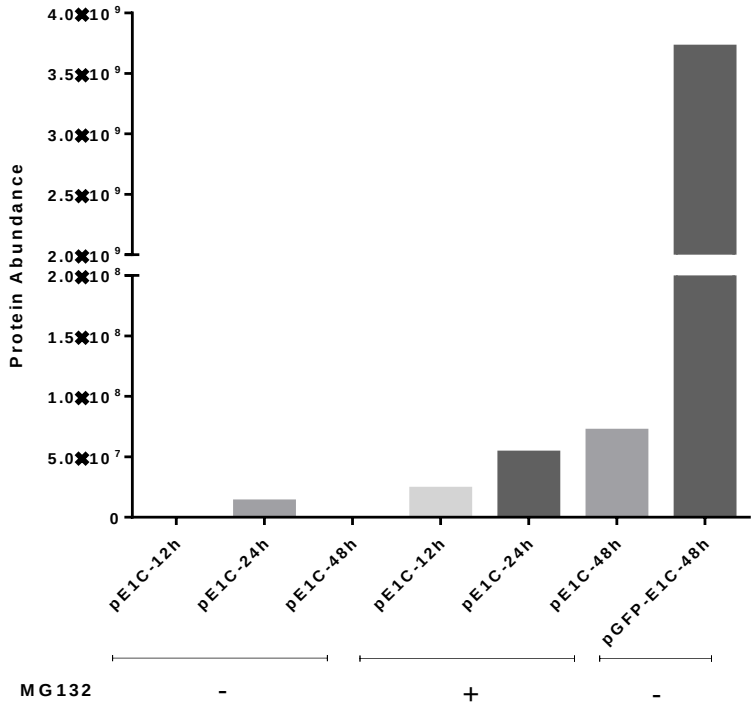


Figure 4.14. **MS detection of overexpressed E1C protein after proteasome inhibition.** Trypsin-digested lysates of HEK-293T cells transfected with p2-E1C or p1-His-GFP-E1C, harvested 12, 24 and 48 h post transfection with or without MG132 pretreatment were applied to MS analysis. Y axis shows the combined abundance (LFQ intensity) of four HPV16 E1- and E1C-specific peptides.

4.2.3. Detection of endogenous E1C protein in HPV16-positive cell lines

4.2.3.1. Endogenous E1C RNA in HPV16-positive cell lines

The absolute endogenous mRNA copies per cell were quantified for E1C by singleplex E1C RT-qPCR and for E6*I by the triplex RT-qPCR in HPV16 positive cell lines CaSki, SiHa, MRI-H186, MRI-H196, HPK-1A, W12-epi and W12-int (Table 4.11). RNA of the HPV16 negative cervical carcinoma cell line C33A served as negative control and as excepted no E6*I and E1C transcripts were identified. In the HPV16 positive cells E1C RNA was found in very low levels ranging from 0.0004 to 0.014 copies/cell, whereas E6*I RNA was found highly abundant ranging from 19 to 849 copies/cell.

Table 4.11. **E1C and E6*I transcript copies/cell in HPV16-positive cell lines.**

	Copies/cell							
Transcript	**CaSki**	**SiHa**	**MRI-H186**	**MRI-H196**	**HPK-1A**	**W12-epi**	**W12-int**	**C33A**
E1C	0.014	0.001	0.001	0.003	0.0004	0.0005	0.0004	-
E6*I	849	122	294	661	121	19	197	-

4.2.3.2. Endogenous E1C protein in HPV16-positive cell lines

In order to detect the endogenous E1C protein, cell lysates of cultured (without MG132 treatment) HPV16 positive cell lines CaSki, MRI-H186, MRI-H196, HPK-1A were analysed in E1C western blot together with lysates from p2-E1C- (MG132-treated) and p1-His-GFP-E1C-transfected HEK-293T cells harvested post 48 hours transfection as positive controls (Figure 4.15.). Cellular lysate from C33A, which is a HPV16 negative cervical cancer cell line, was used as negative control. No endogenous HPV16 E1C protein was detected in the MG132 non-treated HPV16 positive cell lines (Figure 4.15.).

To detect the endogenous E1C protein by IF, CaSkiand SiHacells were cultured for three days and treated with DMSO and MG132 for 6 hours prior to fixation by 3.7% PFA on the fourth day. The fixed CaSki and SiHa cells were prepared for IF analysis afterwards. Endogenous E1C protein was not detectable in both DMSO and MG132 treated CaSki and SiHa cells (Figure 4.16.).

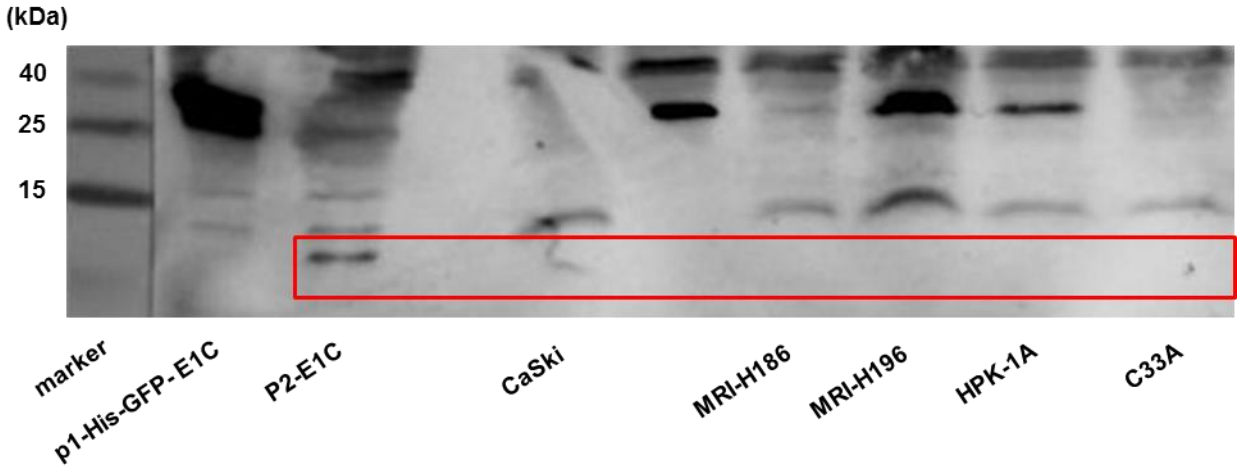


Figure 4.15. **WB detection of endogenous E1C protein in HPV16 positive cell lines**. Cell lysates from CaSki, MRI-H186, MRI-H196, HPK-1A and C33A were applied to the E1C western blot. Lysates of HEK-293T cell transfected with p2-E1C (with MG132 treatment) and p1-His-GFP-E1C and harvested 48 h post transfection served as positive controls. The red frame marks the region on the blot where the endogenous 9 kDa E1C protein would migrate.

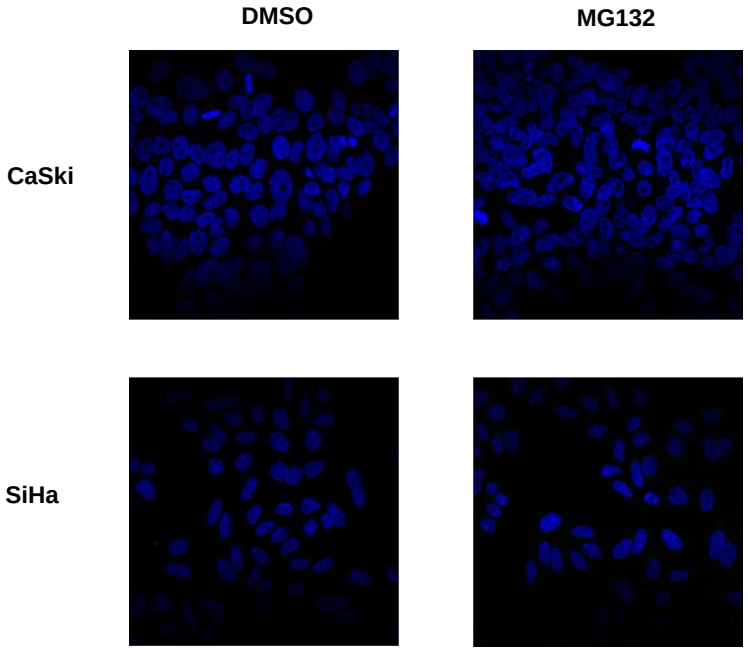


Figure 4.16. **IF detection of endogenous E1C protein in CaSki and SiHa cells.** DMSO and MG132 treated CaSki (top panel) and SiHa (bottom panel) cells stained with α-E1C Ab. Nuclei were stained blue and E1C protein was not visible when visualized through red channel both in DMSO and MG132 treated cells. Pictures are merged images of DAPI and E1C immunostaining.

4.3. Effect of E1C on HPV16 URR** activity

Potential effects of E1C RNA and/or protein on the HPV16 upstream regulatory region (URR) in the absence or presence of HPV16 E2 (E2) were analysed by a dual luciferase reporter assay (DLRA) in HEK-293T cells co-transfected with p-HPV16-URR-FLuc reporter plasmid, p-TATAbox-RLuc control plasmid and different experimental plasmids such as p-EV, p-HPV16-E2, p-HPV16-E1C, p-HPV16-E1C-RNA mut and p-HPV16-E1C-Pro mut. In the DLRA, the Renilla luciferase activity of co-transfected p-TATAbox-RLuc control plasmid is thought to provide a measure for transfection efficiency and the measured raw Firefly luciferase activity of the p-HPV16-URR-FLuc reporter should be normalized by the renilla activity. In the course of the DLRA experiments, it became obvious that presence of the p-TATAbox-RLuc control plasmid affected the p-HPV16-URR-FLuc reporter activity, so a second set of experiments was performed without the control plasmid.

** - HPV16 URR and HPV16 E2 are designated as URR and E2 hereafter in the 'Results' section.

4.3.1. Selection of HEK-293T cells for DLRA

To identify a cell line which supports basal transcription from p-HPV16-URR-FLuc reporter plasmid DLRA was initially performed in C33A, NIH-3T3 and HEK-293T cells. As C33A and NIH-3T3 cells are hard-to-transfect cells, they were transfected using calcium phosphate while PEI was used for HEK-293T.

C33A cells in numbers ranging from 75,000 – 200,000 were transfected p-HPV16-URR-FLuc and firefly luciferase activity expressed as relative luminescence units (RLU) was measured 48 h post transfection . Mean RLU signals ranged from 671 for 75,000 cells to 352 for 200,000 cells with a maximum of 777 for 125,000 cells (Figure 4.17. (A)) while non-transfected C33A cells showed 0 RLU for 75,000 and 9 for 200,000 cells.

For NIH-3T3 and HEK-293T cells, a seeding cell density for 70% confluency on the day of transfection of 40,000 and 75,000 cells, respectively, had been determined previously. The Mean RLU signals were 1,150 for NIH-3T3 and 11,736 for HEK-293T cells (Figure 4.17. (B)) while non-transfected cells showed a mean RLU of 16 and 5, respectively.

Since HEK-293T cells showed the highest basal FLuc activity they were selected for further DLRAs.

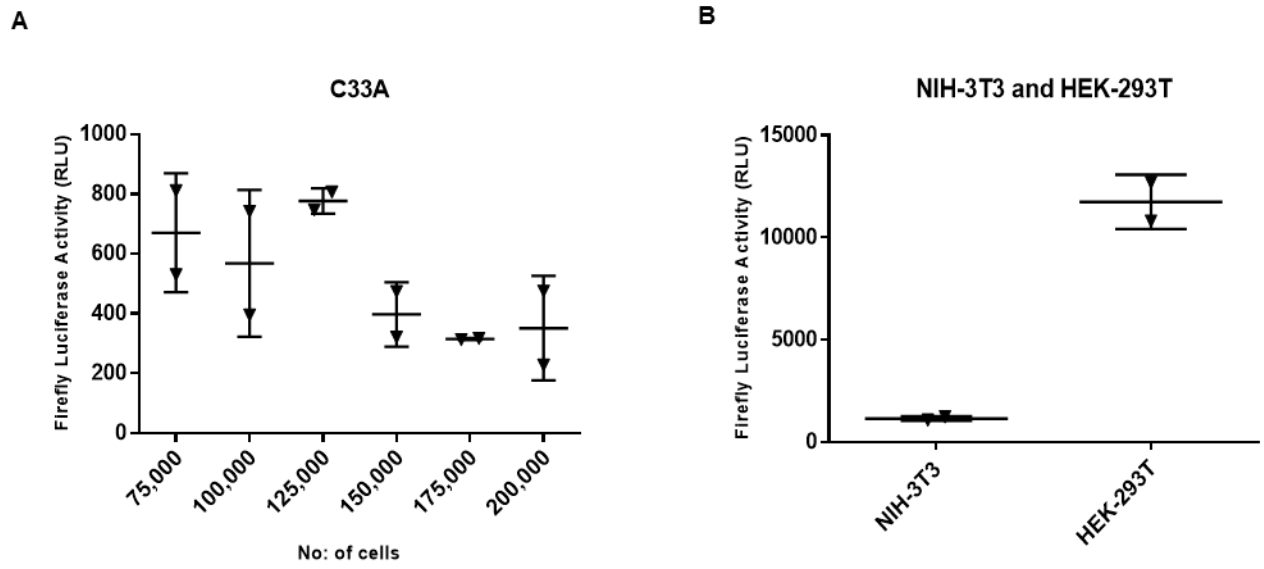


Figure 4.17. **Firefly luciferase activities measured for C33A, NIH-3T3 and HEK-293T cells by DLRA.** A) No: of C33A cells seeded is plotted on the x-axis and the firefly luciferase activity measured is plotted on the y-axis. B) Firefly luciferase activities measured for NIH-3T3 and HEK-293T cells are plotted. Mean together with standard deviation is shown for all the data points.

4.3.2. Time point determination for the measurement of firefly luciferase activity

To further optimize DLRA, the best time point for the measurement of FLuc activity after transfection of 639 ng p-HPV16-URR-FLuc in HEK-293T cells was determined. FLuc activities were measured 12, 24 and 48 h post transfection.

The mean RLU signals increased from 362 at 12 h, and 2302 at 24 h to 11,736 at 48 h (Figure 4.18.). Non-transfected cells at all 3 time points showed 23 or less RLU. Since, the 48 h time point showed the maximum FLuc activity it was chosen for further experiments.

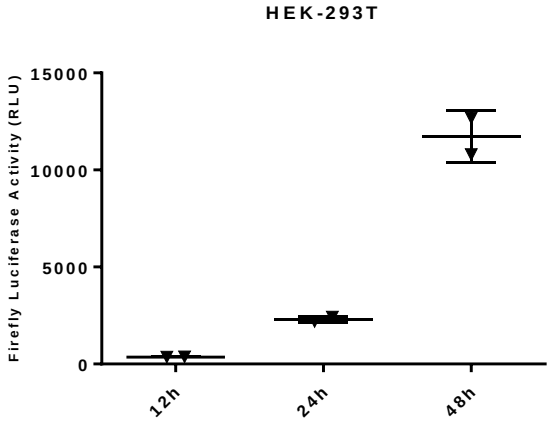


Figure 4.18. **Time point determination for firefly luciferase activity measurement in HEK-293T cells by DLRA.**12 h, 24 h and 48 h time points are plotted on the X-axis and the firefly luciferase activity measured is plotted on the Y-axis. Mean together with standard deviation is shown for all the data points.

4.3.3. Determination of a control plasmid for DLRA

In order to determine a proper control plasmid, both p-SV40-RLuc and p-TATAbox-RLuc plasmids encoding renilla luciferase (RLuc) were tested in DLRA. Transfection of a control plasmid together with the reporter and other experimental plasmids is essential for the normalization of reporter activity, which could vary depending on transfection efficiency. The aim was to find out the least amount of control plasmid and thereby the best feasible reporter-control plasmid amount combination for co-transfection which avoids the occurrence of *trans* effects between promoter elements.

At first, 2.5, 5 and 10 ng of p-SV40-RLuc or p-TATAbox-RLuc plasmids were transfected in HEK-293T cells and RLuc activities were measured. In general, RLuc activities from p-SV40-RLuc were higher than from p-TATAbox-RLuc. Also, mean RLU signals increased with increase in amount of plasmid 12,450 with 2.5 ng, 45,969 with 5 ng and 46,829 with 10 ng p-SV40-RLuc (Figure 4.19. (A)) and from 212 with 2.5 ng, 385 with 5 ng and 430 with 10 ng p-TATAbox-RLuc (Figure 4.19. (B)). Since 2.5 ng of both p-SV40-RLuc and p-TATAbox-RLuc plasmids showed the lowest, but robustly measurable RLuc signals, this amount was chosen to test the combination of p-HPV16-URR-FLuc reporter and p-SV40-RLuc or p-TATAbox-RLuc control plasmids for DLRA.

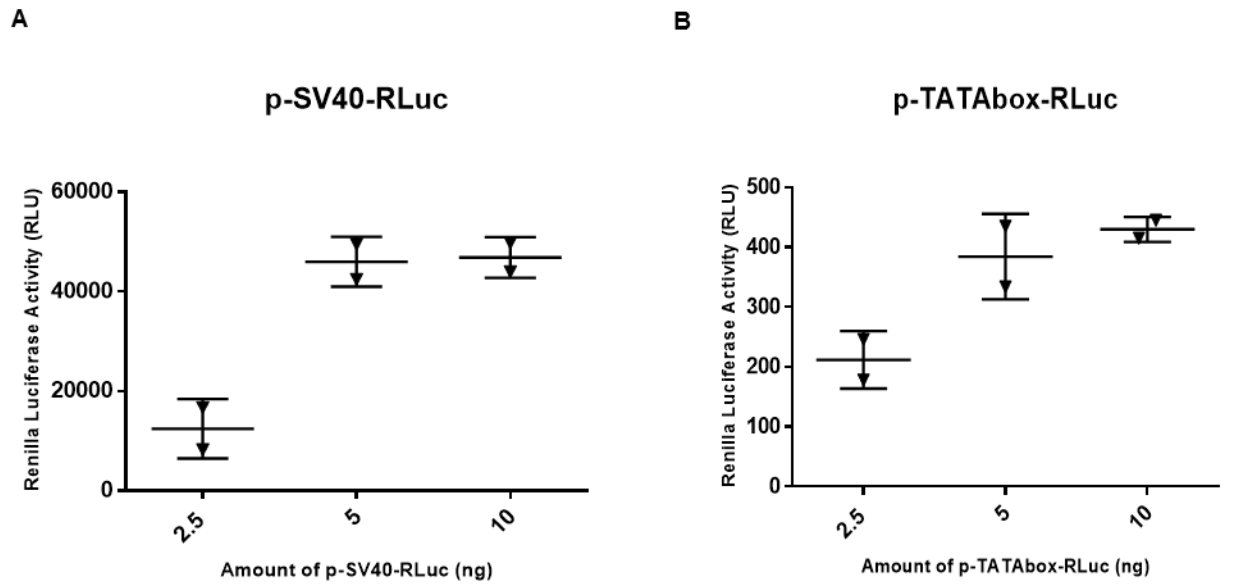


Figure 4.19. **Determination of a control plasmid for DLRA in HEK-293T cells.** Amounts of p-SV40-RLuc (A) and p-TATAbox-RLuc (B) plasmids in ng are plotted on the X-axis and the renilla luciferase activity measured is plotted on the Y-axis. Mean together with standard deviation is shown for all the data points.

For testing the combination of the reporter and control plasmids, 75,000 HEK-293T cells were co-transfected with 500 ng of p-HPV16-URR-FLuc and 2.5 ng of p-SV40-RLuc or p-TATAbox-RLuc, Each plasmid was also transfected separately for measuring the respective luciferase signals without any promoter *trans* effects.

Co-transfection of p-HPV16-URR-FLuc with p-SV40-RLuc (Figure 4.20. (A)) lowered the mean FLuc RLU signal to 875 relative to 1742 for p-HPV16-URR-FLuc alone. The mean RLuc RLU signal in co-transfection increased to 83,383 relative to 16,717 for p-SV40-RLuc alone. Co-transfection with p-TATAbox-RLuc (Figure 4.20. (B)) reduced the mean FLuc RLU signal to 1024. The mean RLuc RLU signal for p-TATAbox-RLuc alone was 23 and in co-transfection, increased to 308.

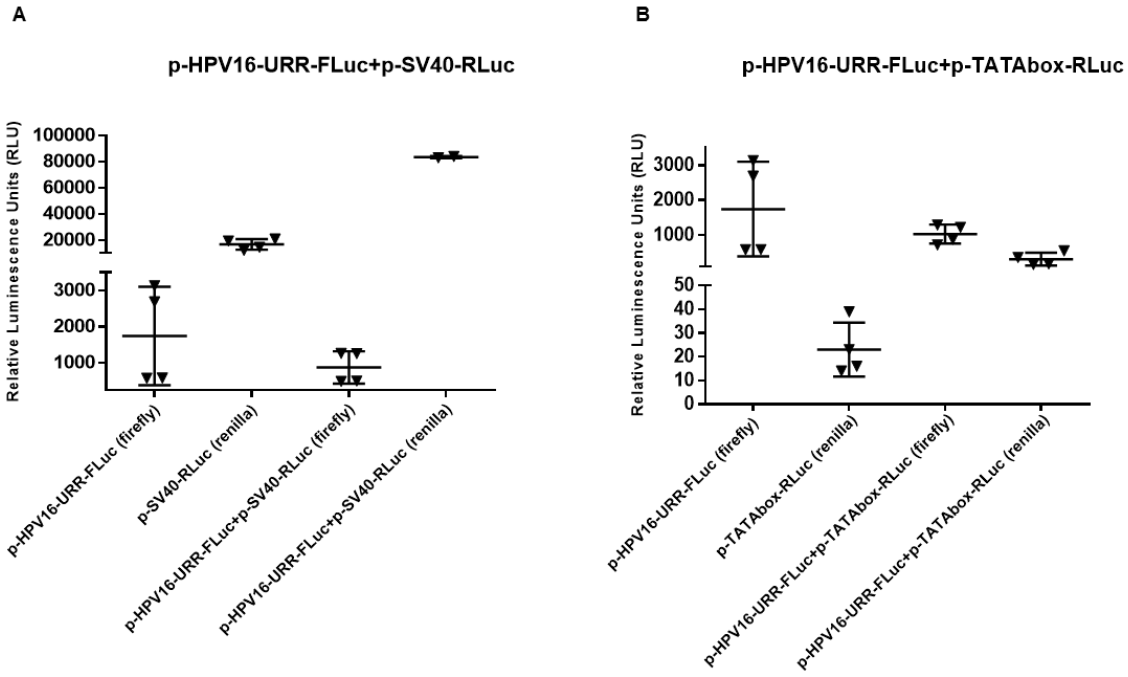


Figure 4.20. **Testing the combination of reporter and control plasmids for DLRA in HEK-293T cells.** Combinations of p-HPV16-URR-FLuc and p-SV40-RLuc (A) and p-HPV16-URR-FLuc and p-TATAbox-RLuc (B) plasmids were tested in DLRA. X-axis indicates the transfected plasmid(s) and in brackets, the luciferase analysed. RLU values are plotted on the Y-axis. Bars indicate arithmetic mean and standard deviation. In A) the mean renilla RLU for co-transfection is an underestimation because it was calculated from two data points only, since the luminometer showed 'OVRFLW' during measurement of two other samples (which means the renilla signals were beyond the measurement threshold of the luminometer).

Thus, FLuc activity was reduced 50% by co-transfection with p-SV40-RLuc and 41% with p-TATAbox-RLuc. Co-transfection with p-HPV16-URR-FLuc increased RLuc activity of p-SV40-RLuc by 398% and that of p-TATAbox-RLuc by 1239%. Since the promoter *trans*

effect on FLuc was less in co-transfections with p-TATAbox-RLuc, this control plasmid was used in subsequent experiments.

4.3.4. Effect of E2 on URR activity

Alloul and Sherman[63] for HPV16 have shown with a chloramphenicol acetyltransferase reporter assay that an expression construct for E2 represses and one for E1C activates URR activity. They have also shown that co-transfection of E1C reduces the repressive effect of E2, owing to the simultaneous activation by E1C. Before proceeding to determine in a more sensitive assay like DLRA whether E1C protein or E1C RNA activates URR, I decided to test first the repressive effect of E2 on URR.

HEK-293T cells were transfected with p-HPV16-URR-FLuc, p-TATAbox-RLuc and p-HPV16-E2 or empty p-EV plasmids. In the total of ten experiments the mean RLU signal of both FLuc and RLuc activity for non-transfected cells was 15 or less. In p-EV co-transfected cells the mean FLuc RLU signal ranged from 2423 to 6238 and with co-transfected p-HPV16-E2 the signal was always reduced, ranging from 1389 to 5179 (Figure 4.21. (A)). In summary, in these 10 experiments E2 reduced FLuc activity significantly ($p = 0.0001$) by about 35% (Figure 4.21. (B)), indicating that E2 in concordance with the data obtained by Alloul and Sherman[63] also in this experimental design, represses HPV16 URR.

However, in all experiments with p-HPV16-E2 transfected HEK-293T cells, RLuc activities of the control plasmid were also repressed compared to p-EV (Figure 4.22.). Mean RLuc RLU signals ranged between 160 and 2167 for the p-EV co-transfected cells and 55 - 907 for the p-HPV16-E2 co-transfected cells (Figure 4.22.). Therefore, only raw FLuc data without normalization by RLuc data were considered in the analysis of effects of E2 on URR activity. As a further consequence, in further DLRA esperiments, co-transfections were performed with and without p-TATAbox-RLuc and firefly luciferase activities in both cases were analyzed separately. p-HPV16-URR-FLuc and p-TATAbox-RLuc were also transfected separately and in combination to analyze the background signals and

corresponding data of all the DLRA performed is given in the supplementary information 9.6..

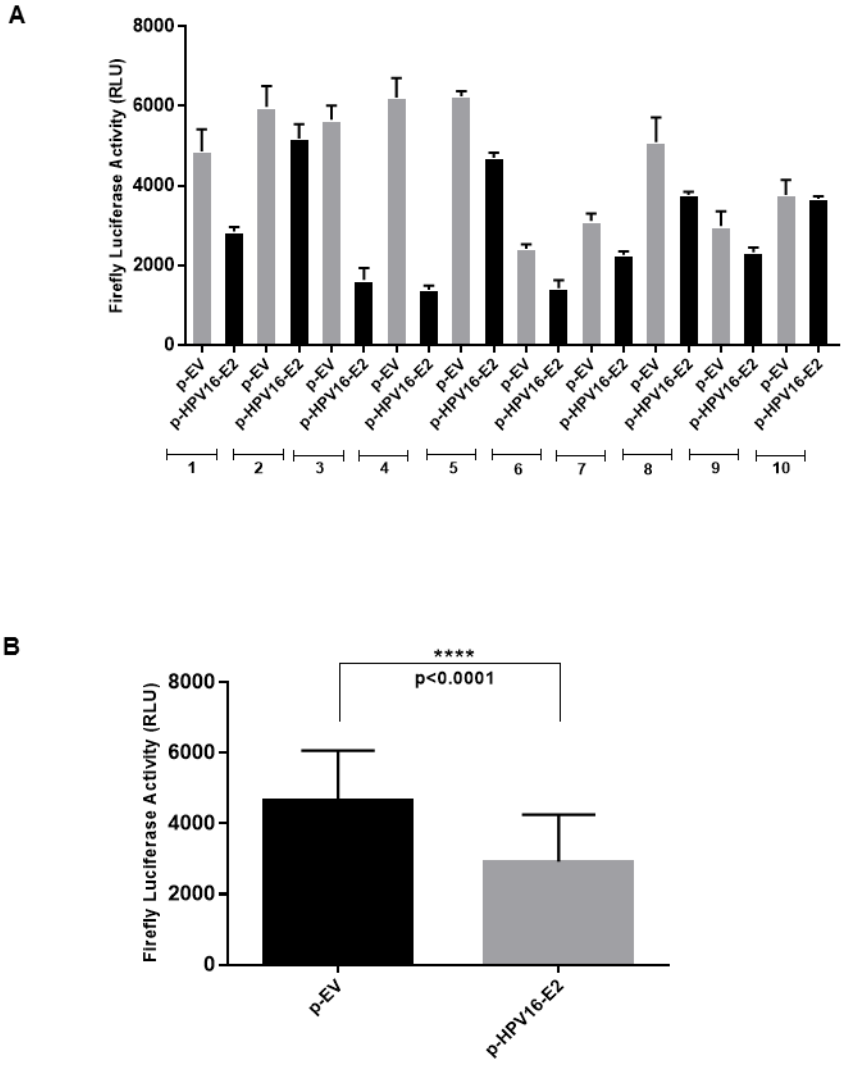


Figure 4.21. **Measurement of firefly luciferase activity in p-HPV16-E2 co-transfected HEK-293T cells.** A) HEK-293T cells were co-transfected with p-HPV16-URR-FLuc and p-TATAbox-RLuc and p-EV or p-HPV16-E2. Mean and SD of raw FLuc RLU signals of 10 independent experiments are shown. B) Mean and SD of the FLuc RLU values of p-EV and p-HPV16-E2, respectively, from the 10 individual experiments shown in (A). The two-tailed p value of 0.0001 was calculated by paired t test

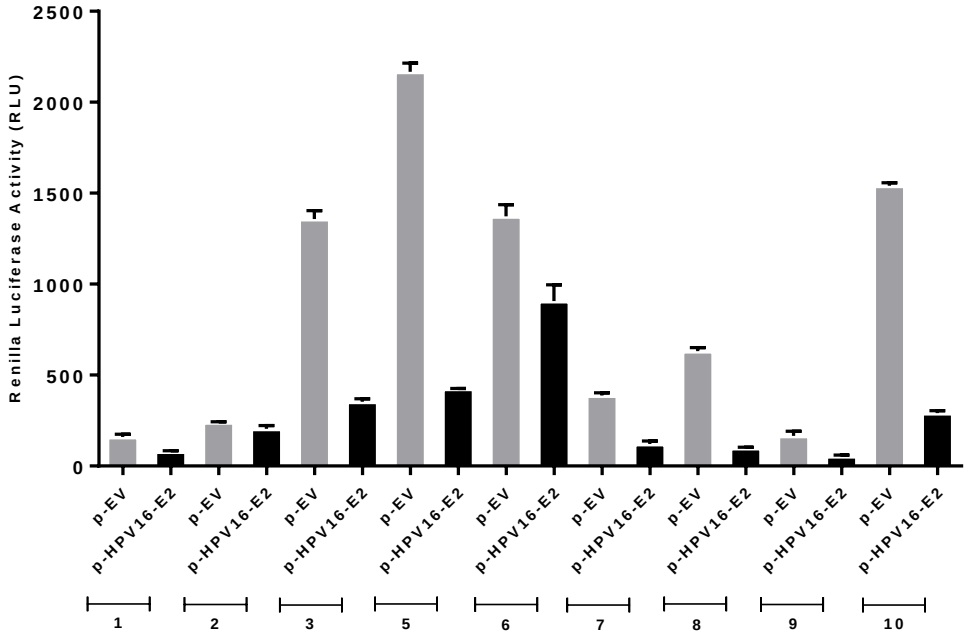


Figure 4.22. **Measurement of renilla luciferase activity in p-HPV16-E2 co-transfected HEK-293T cells.** HEK-293T cells were co-transfected with p-HPV16-URR-FLuc and p-TATAbox-RLuc, and p-EV or p-HPV16-E2. Mean and SD of raw RLuc RLU signals of 9 independent experiments are shown. Experiment 4 was excluded from the analysis because of outlier renilla signals obtained in the duplicates with both p-EV and p-HPV16-E2 (For data see supplementary information 9.6..)

4.3.5. Effect of E1C and E1C mutants on URR activity

In order to study by DLRA the effect of E1C on URR, HEK-293T cells were transfected in 4 independent experiments with p-HPV16 URR-FLuc, p-TATAbox-RLuc and p-HPV16-E1C or p-EV plasmids. Mean FLuc RLU signals in non-transfected cells were 1.5 - 15 and 0 – 8 for RLuc. With p-EV mean FLuc RLU signals were 2423 - 6238 and 4332 – 7444 with p-HPV16-E1C (Figure 4.23 (A)). Mean RLuc RLU signals were 160 – 2167 with p-EV and 150 - 1177 with p-HPV16-E1C (for RLuc data see supplementary information 9.6.).

The results demonstrated significant (p = 0.0005) activation of about 42% of the HPV16 URR by E1C (Figure 4.23. (B)) indicating that E1C, also in this experimental design, activates the HPV16 URR.

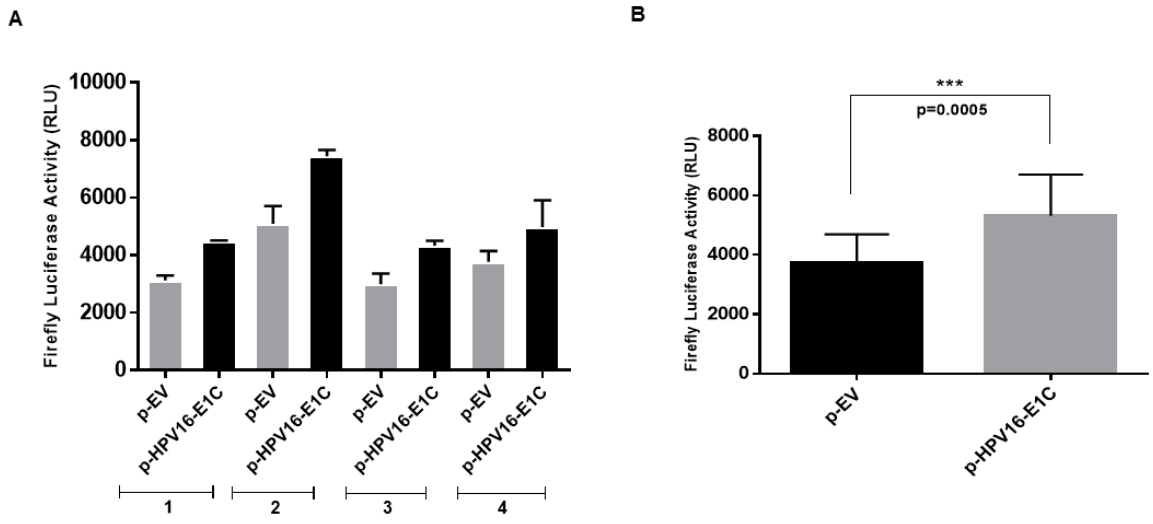


Figure 4.23. **Measurement of firefly luciferase activity in p-HPV16-E1C co-transfected HEK-293T cells.** A) HEK-293T cells were co-transfected with the reporter and control luciferase plasmids and with p-EV or p-HPV16-E1C. Mean and SD of raw FLuc RLU signals of 4 independent experiments are shown. B) Mean and SD of the raw FLuc RLU values of p-EV and p-HPV16-E1C, respectively, from the 4 individual experiments shown in (A). The two-tailed p value of 0.0001 was calculated by paired t test.

In parallel, additional four p-HPV16-E1C co-transfection experiments were carried out in duplicates without the addition of p-TATAbox-RLuc. Mean FLuc RLU signals were 3306 – 5084 with p-EV and 4667 – 6731 with p-HPV16-E1C (Figure 4.24 (A)). Here again E1C caused a significant (p = 0.0003) activation of about 29%.

FLuc activities measured in the presence of both p-EV and p-HPV16-E1C and with and without p-TATAbox-RLuc co-transfection are summarized in Table 4.12.. The ratio of FLuc activity without p-TATAbox-RLuc to the activity with p-TATAbox-RLuc calculated for each experiment varied only 0.97 and 1.37 for p-EV and 0.87 to 1.1 for p-HPV16-E1C so that the data were combined for statistical significance analysis. A two-tailed p value of <0.0001 was obtained by paired t test and a mean activation of 35% of FLuc activity was

observed from the total eight experiments (Figure 4.24. (C)).Taken together these results indicate that E1C activates the HPV16 URR and are in concordance with the data obtained by Alloul and Sherman[63].

Table 4.12. **Summary of firefly luciferase activity measured for p-HPV16-E1C co-transfected HEK-293T cells with and without p-TATAbox-RLuc co-transfection.** FLA – Firefly luciferase activity

Experiment No:	FLA with p-TATAbox-RLuc transfection (RLU)		FLA with p-TATAbox-RLuc transfection (RLU)		Ratio of FLA without to with p-TATAbox-RLuc transfection	
	p-EV	p-HPV16-E1C	p-EV	p-HPV16-E1C	p-EV	p-HPV16-E1C
1	3246	4505	4004	4971	1.23	1.1
	2996	4472	4116	4865	1.37	1.08
2	5531	7599	5610	6632	1.01	0.87
	4661	7288	4558	6829	0.97	0.93
3	2725	4455	3308	4787	1.21	1.07
	3251	4208	3304	4546	1.01	1.08
4	4037	4325	3948	4632	0.97	1.07
	3511	5640	4370	5445	1.24	0.96

In order to analyze whether it is the E1C RNA or E1C protein that activates URR, two mutant E1C expression plasmids were constructed. p-HPV16-E1C-RNA mut transcribes a random sequence with the same size and GC content as that of E1C RNA and in p-HPV16-E1C-Pro mut the first three E1C codons are replaced by TGA, TAG and TAA stop codons to abolish translation of E1C protein.

HEK-293T cells were co-transfected with p-HPV16 URR-FLuc, with or without p-TATAbox-RLuc and with p-HPV16-E1C or p-HPV16-E1C-RNA mut or p-HPV16-E1C-Pro mut or p-EV plasmids. Two experiments each were performed with and without p-TATAbox-RLuc. Since the FLuc activity variation with and without p-TATAbox-RLuc transfection was negligible, the data from all four experiments were combined and analyzed together.

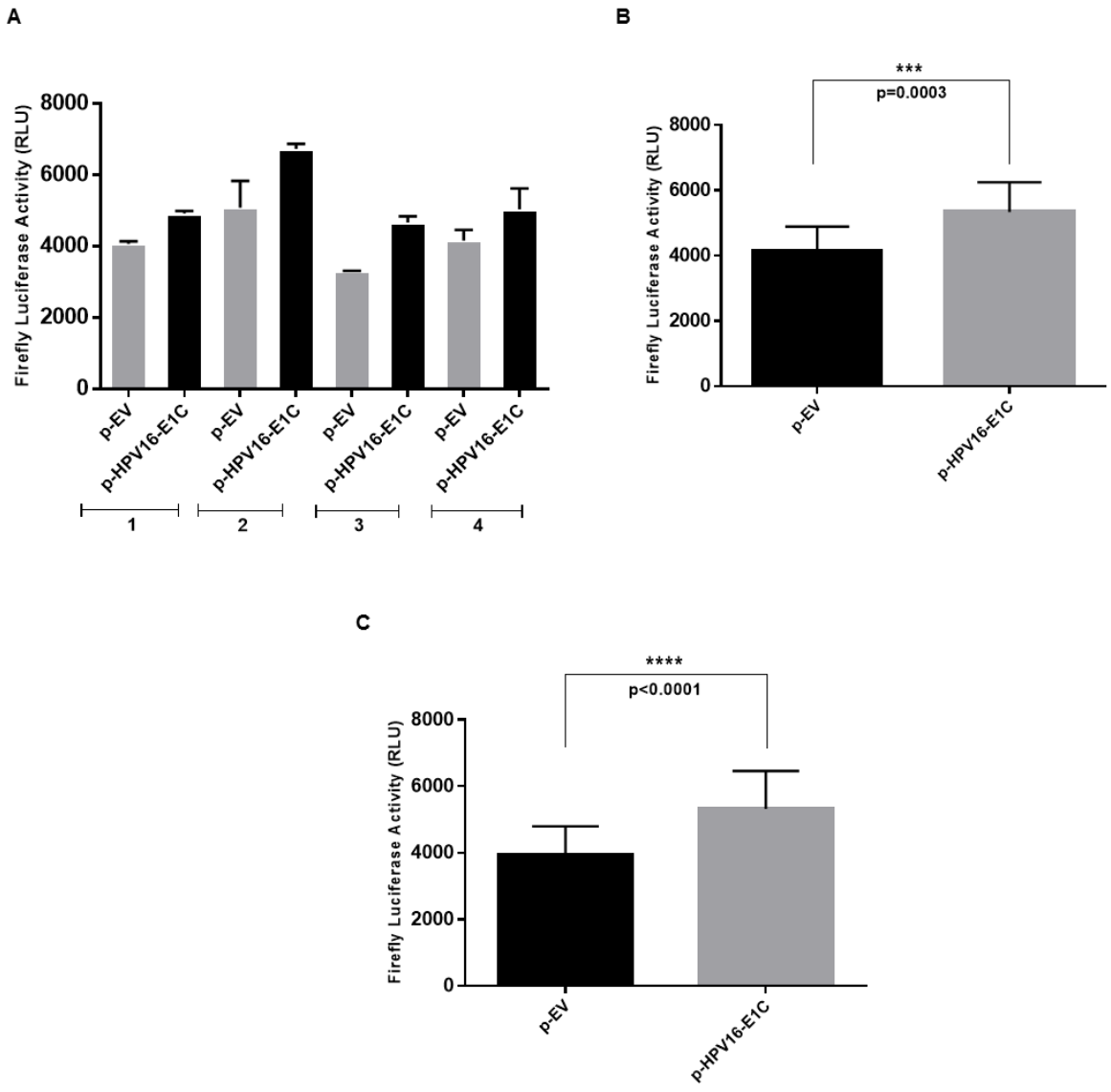


Figure 4.24. **Measurement of firefly luciferase activity in p-HPV16-E1C co-transfected HEK-293T cells without p-TATAbox-RLuc transfection.** A) HEK-293T cells were co-transfected with p-HPV16-URR- and p-EV or p-HPV16-E1C. Mean and SD of raw FLuc RLU signals of 4 independent experiments are shown. B) Mean and SD of the raw FLuc RLU values of p-EV and p-HPV16-E1C, respectively, from the 4 individual experiments shown in (A). The two-tailed p value of 0.0003 was calculated by paired t test. C) All FLuc values of p-EV and p-HPV16-E1C co-transfected cells from (A) and Figure 4.23. (A) are combined in a bar graph to evaluate the statistical significance. The two-tailed p value of 0.0001 was calculated by paired t test.

In non-transfected cells the mean raw FLuc RLU signal was 12.5, and 4.5 for RLuc. In p-EV-transfected cells mean FLuc RLU signals were 2988 - 4159, 4332 – 5039 with p-HPV16-E1C, 3917 - 4676 with p-HPV16-E1C-RNA mut and 4693 - 6207 with p-HPV16-E1C-Pro mut (Figure 4.25 (A)). (For mean RLuc RLU signals see supplementary information 9.6.)

With p-HPV16-E1C, a significant ($p = 0.0007$, two-tailed, paired t test) mean FLuc activation of 35% was observed, a significant ($p = 0.003$) mean activation of 19% with p-HPV16-E1C-RNA mut and a significant ($p = 0.0001$) mean activation of 33% with p-HPV16-E1C-Pro mut (Figure 4.25. (B)). Thus, E1C but also both p-HPV16-E1C-RNA mut and p-HPV16-E1C-Pro mut appeared to activate the HPV16 URR.

4.3.6. Effect of co-transfection of E1C and E2 on URR activity

In order to study the effect of E1C in combination with E2 on URR by DLRA, HEK-293T cells were co-transfected with p-HPV16 URR-FLuc, p-TATAbox-RLuc (in 4 of the 8 experiments), p-HPV16-E2, and p-EV or p-HPV16-E1C plasmids. The data from all eight experiments were combined and analyzed together.

In non-transfected cells, mean RLU signals for FLuc were 1.5 - 15 and 0 – 8 for RLuc. With p-EV alone, mean FLuc RLU signals were 2988 – 5096, 2265 – 3770 with p-HPV16-E2 alone, 4332 – 7444 with p-HPV16-E1C alone, 1077 – 3242 with p-HPV16-E2 and p-EV, and 1351 – 4082 with p-HPV16-E2 and p-HPV16-E1C (Figure 4.26. (A)). (For RLuc RLU signals see supplementary information 9.6.)

A significant ($p < 0.0001$, two-tailed, paired t test) mean reduction of FLuc activity of 24% was observed with p-HPV16-E2 alone, a significant ($p < 0.0001$) mean activation of 35% with p-HPV16-E1C alone, and a significant ($p = 0.0124$) mean activation of 19% with p-HPV16-E2 and p-HPV16-E1C combined (Figure 4.26. (B)). Thus, E1C co-transfected with E2 reduced the repressive effect of E2, which is in accordance with the data obtained by Alloul and Sherman[63].

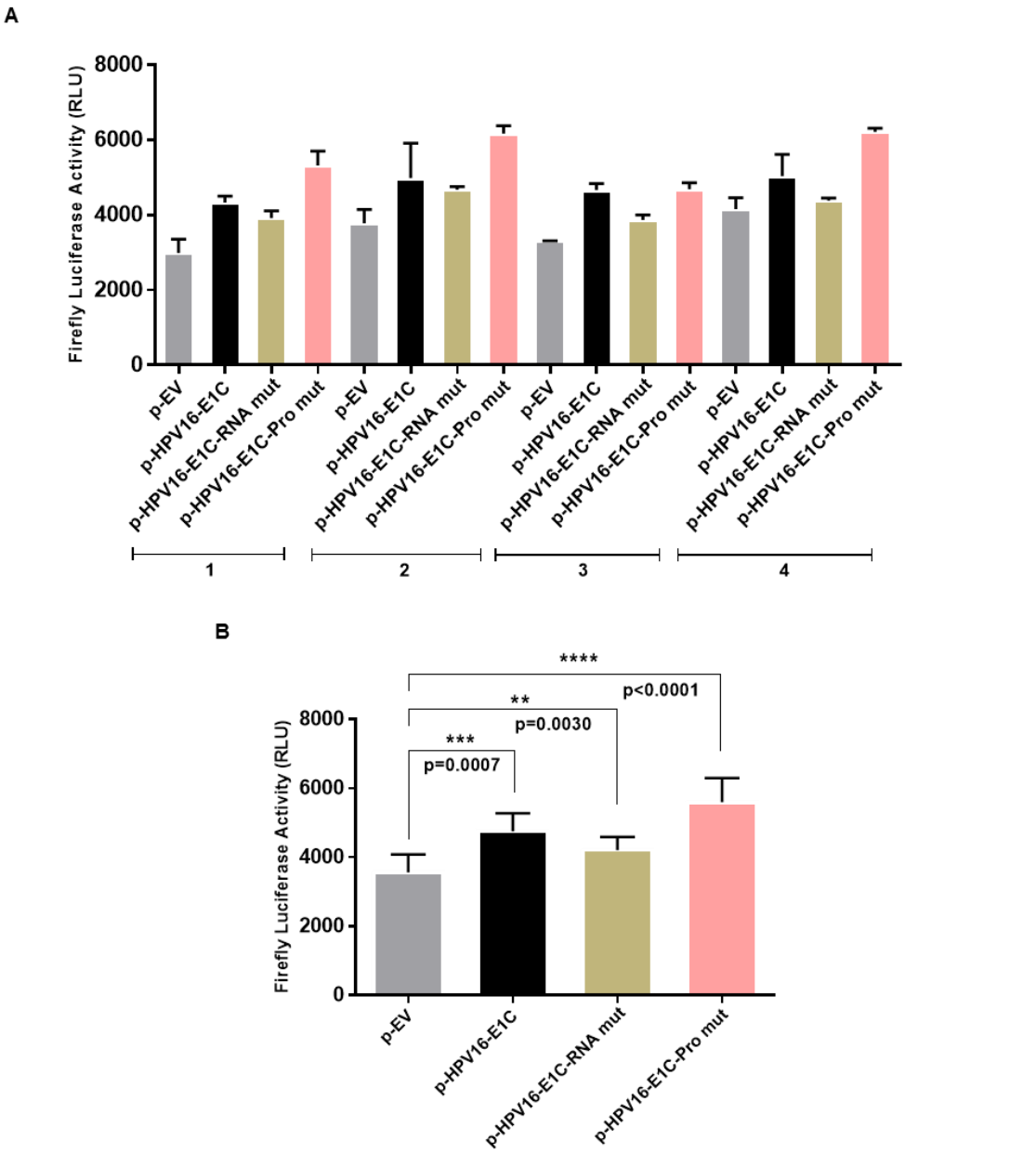


Figure 4.25. **Measurement of firefly luciferase activity in p-HPV16-E1C-RNA mut and p-HPV16-E1C-Pro mut co-transfected HEK-293T cells.** A) HEK-293T cells were co-transfected with one of the expression plasmid p-EV, p-HPV16-E1C, p-HPV16-E1C-RNA mut or p-HPV16-E1C-Pro mut and the reporter plasmid p-HPV16-URR-FLuc with or without the control plasmid p-TATAbox-RLuc. Mean and SD of raw FLuc RLU signals of 4 experiments are shown. B) Mean and SD of the raw FLuc RLU values obtained with p-EV, p-HPV16-E1C, p-HPV16-E1C-RNA mut or p-HPV16-E1C-Pro mut, respectively, from the 4 individual experiments shown in (A). The two-tailed p values were calculated by paired t test.

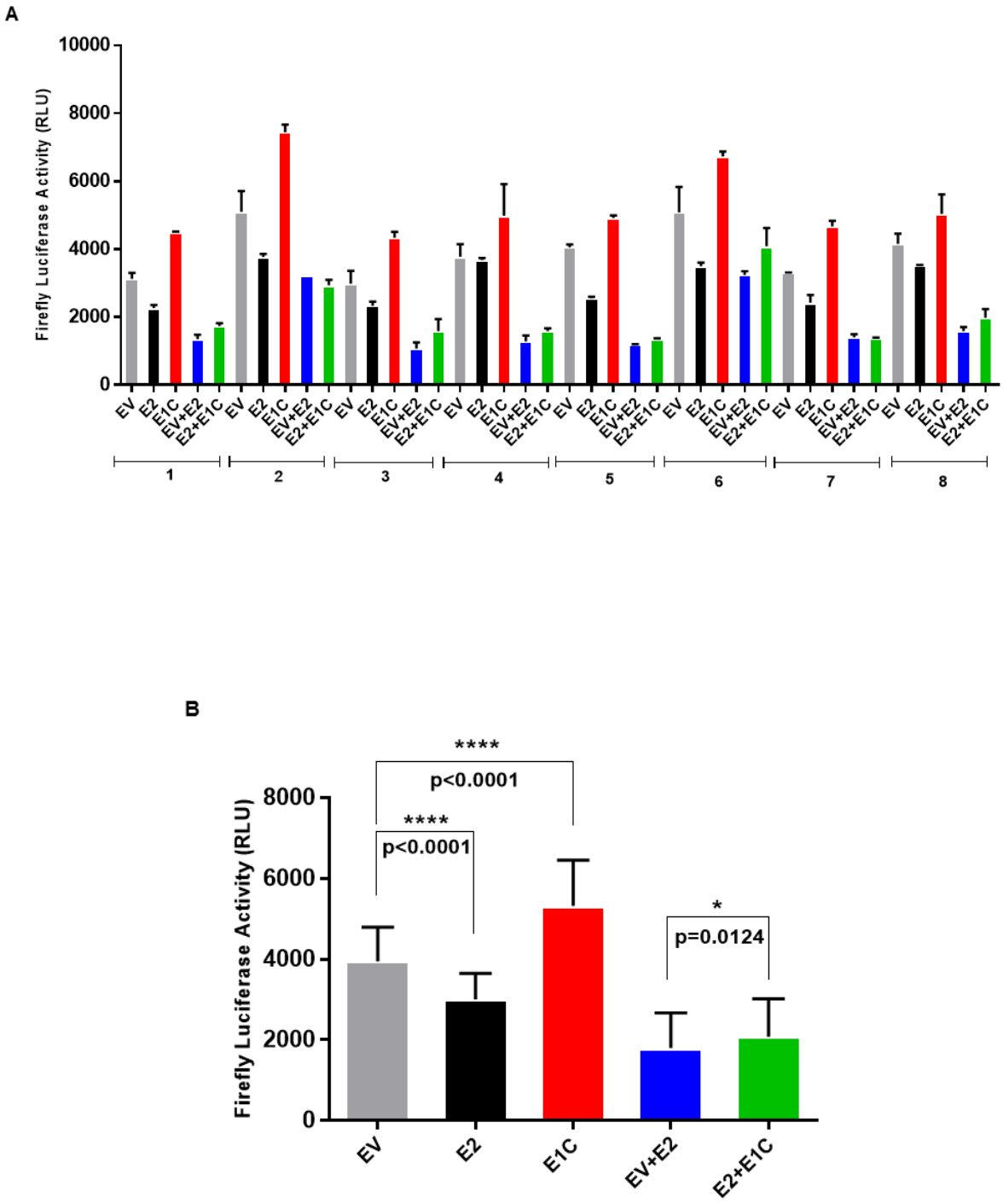


Figure 4.26. **Measurement of firefly luciferase activity in p-HPV16-E1C and p-HPV16-E2 co-transfected HEK-293T cells.** A) HEK-293T cells were co-transfected with reporter plasmid p-HPV16-URR-FLuc and with or without the control plasmid p-TATAbox-RLuc as well as with the expression plasmids p-EV (EV), p-HPV16-E2 (E2) and p-HPV16-E1C (E1C) alone or in combinations. Mean and SD of raw FLuc RLU signals of 8 experiments (4 with and 4 without p-

TATAbox-RLuc) are shown. (B) Mean and SD of the raw FLuc RLU values obtained with EV, E2, E1C, EV+E2 and E2+E1C, respectively, from the 4 individual experiments shown in (A). The two-tailed p values were calculated by paired t test.

In order to analyze whether it is the E1C RNA or E1C protein that is capable of reducing the repressive effect of E2 on URR, HEK-293T cells were transfected with p-HPV16 URR-FLuc, p-TATAbox-RLuc (in 2 of the 4 experiments), p-HPV16-E2, either p-HPV16-E1C, p-HPV16-E1C-RNA mut or p-HPV16-E1C-Pro mut plasmids. The data from all four experiments were combined for statistical analysis.

In non-transfected cells, mean RLU signal was 12.5 for FLuc and 4.5 for RLuc. With p-EV alone, the FLuc RLU signals were 2988 – 4159, 2333 – 3672 with p-HPV16-E2 alone, 4332 – 5039 with p-HPV16-E1C alone, 1077 – 1590 with p-HPV16-E2 and p-EV combined, 1390 – 1975 with p-HPV16-E2 and p-HPV16-E1C combined, 1526 – 2067 with p-HPV16-E2 and p-HPV16-E1C-RNA mut combined, and 1772 – 2499 with p-HPV16-E2 and p-HPV16-E1C-Pro mut combined (Figure 4.27. (A)). (For RLuc RLU signals see supplementary information 9.6.)

In comparison to p-EV, a significant ($p = 0.0042$, two-tailed, paired t test) mean FLuc reduction of 17% was observed with p-HPV16-E2 and a significant ($p = 0.0007$) activation of 35% with p-HPV16-E1C. In the presence of p-HPV16-E2 and compared to p-EV, p-HPV16-E1C significantly ($p = 0.0007$) activated FLuc by 24%, p-HPV16-E1C-RNA mut significantly ($p = 0.0052$) activated by 32% and p-HPV16-E1C-Pro mut significantly ($p = 0.0028$) activated by 69% (Figure 4.27. (B)).

Thus, not only p-HPV16-E1C but also both E1C mutants, p-HPV16-E1C-RNA mut and p-HPV16-E1C-Pro mut, reduced the repressive effect of co-transfected E2 on the HPV16-URR.

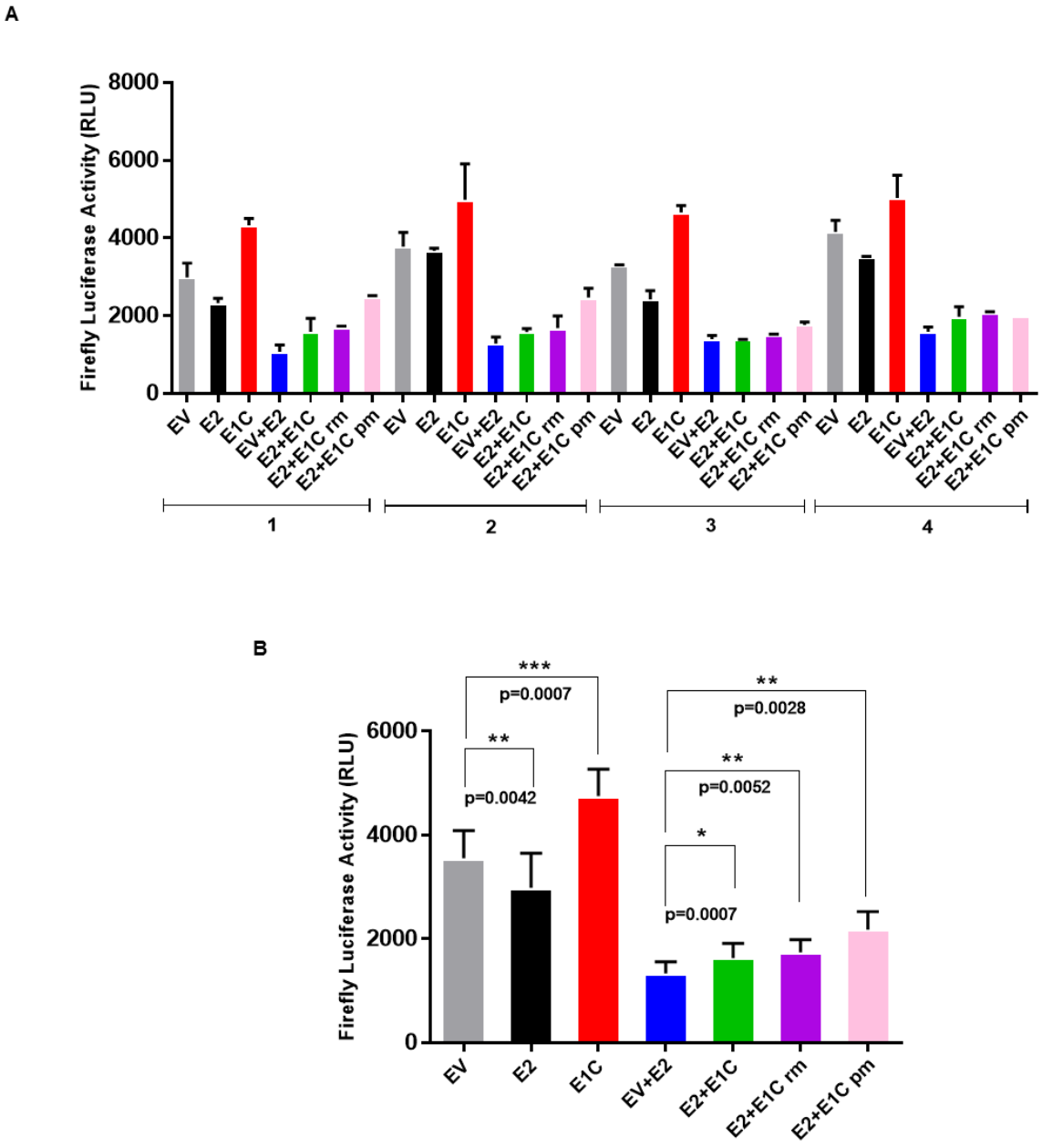


Figure 4.27. **Measurement of firefly luciferase activity in p-HPV16-E2, p-HPV16-E1C-RNA mut and p-HPV16-E1C-Pro mut co-transfected HEK-293T cells.** A) HEK-293T cells were co-transfected with the reporter plasmid p-HPV16-URR-FLuc and with or without the control plasmid p-TATAbox-RLuc (2 experiments each) in combination with either p-EV (EV), p-HPV16-E1C (E1C) or p-HPV16-E2 (E2) alone or p-HPV16-E2 together with either p-EV, p-HPV16-E1C, p-HPV16-E1C-RNA mut (E1C rm or p-HPV16-E1C-Pro mut ((E1C pm). Mean and SD of raw FLuc RLU signals of 4 experiments with duplicate samples are shown. B) Mean and SD of the raw FLuc RLU values obtained from the co-transfections indicated below the x-axis in the 4 individual experiments shown in (A). The two-tailed p values were calculated by paired t test.

5. Discussion

5.1. Rationale

The aim of this thesis was the characterization of the diagnostically relevant HPV16 E1C transcript (splice 880^2582). Schmitt and colleagues[152] had analyzed the HPV16 transcriptome including 10 spliced and 5 fl transcripts to identify novel RNA patterns for the grading of LSIL and HSIL, by NASBA-Luminex assay[152,153,158-160] (see Introduction section 1.4.1.1.). One of the important findings of this study was the identification of the frequent upregulation of E1C transcripts in HSIL and CxCa patients when compared with NIL/M. E1C transcript's prevalence gradually increased from NIL/M to CxCa. The study used the transcript ratio of E1C to L1 (splice 3632^5639) for the differentiation of <=LSIL and >= HSIL and found it to be diagnostically useful as the results were in concordance with the clinical grading of these lesions[153].

Another study which showed the diagnostic relevance of E1C transcript was from Höfler and colleagues[85]. They developed new HPV16 RNA patterns using RT-qPCR technology for the detection of the 3 spliced transcripts E6*I, E1^E4 and E1C for the discrimination of mild and severe lesions and thereby improving the positive predictive value of HPV DNA testing in screening settings. When the prevalence of E1C transcript was analyzed, it was found to be 3% or lower in NIL/M, CIN0 and CIN1, 17% in CIN2, 54% in CIN3 and 22% in CxCa. The study described E1C as a specific, but insensitive marker for the differentiation of ≤CIN1 and ≥CIN3[85].

Alloul and Sherman[63] studied a transcription regulating activity on the HPV16 LCR/URR of a'-type HPV16 mRNA (E1C) which contains the 880/2581 splice. They found that E1C cDNA can activate the LCR in a dose-dependent fashion and together with E2 cDNA can reduce the repressive effect of E2 on LCR activity. This is the only study so far that had shed some light to the functional side of E1C.

Based on these studies, my thesis aimed (i) to further understand the role of E1C in regulating transformation specific HPV16 transcripts E6*I, E1^E4 and E7, (ii) to analyse whether E1C effects are mediated by E1C RNA or E1C protein and (iii) to corroborate the findings of LCR activation by E1C cDNA from Alloul and Sherman using a more sensitive and better quantifiable assay like DLRA.

5.2. Effect of E1C overexpression on HPV16 transcripts

Experiments on silencing of E1C transcripts in HPV16 positive cell lines was purposefully avoided as the endogenous E1C mRNA levels present in the cell lines were found to be very low and were in the range of 0.0004 – 0.014 copies/cell. So the rationale was not to further knock down the already lowly abundant E1C transcript which might fail to detect the subsequent changes, e.g. in HPV16 transcript levels. Therefore, E1C overexpression experiments were performed by lentiviral transduction. Lentiviral transduction was chosen over transient transfection because there would not be an over-dosage of E1C transcripts in stably transduced cells; in line with this expectation the E1C transcript levels found in the lentivirally transduced cells was approximately 10 copies/cell. In contrast, in the case of transient transfection, say, by PEI in HEK-293T cells, the overexpressed E1C RNA level reached approximately 1.85×10^4 copies/cell which is roughly 2000 times more than lentiviral transduction. In addition, in clinical specimens, the number of E1C copies per HPV-infected cell is estimated in the range from 1 – 20[147]. Therefore, E1C overexpression by lentiviral transduction was the overexpression system closest to the physiological status.

W12-epi and W12-int cell lines were two out of the four cell lines chosen for E1C overexpression. The parental W12 cell line was established from a LSIL patient[162]. W12-epi and W12-int are two of the many clones generated from the parental W12 cell line in the Dr. Paul F. Lambert lab[163]. The HPV16 genome is integrated in almost 85% of HPV16 positive CxCas[161]. As the HPV16 genome is present as episomes in W12-epi, this cell line represents an early stage of transformation and is a good material to analyze the viral and cellular events that could trigger the transformation pathway further. Since the HPV16

genome is integrated in W12-int cells, this cell line is thought to be closer to the fully transformed CxCa stage. The other two cervical carcinoma-derived cell lines, CaSki and SiHa also have integrated HPV16 genomes. CaSki has around 400-600 complete and SiHa has one incomplete HPV16 genome(s) present per cell and thus are at the two extremes with respect to HPV16 genome copy numbers[164,165]. Therefore, the four cell lines chosen were assumed to be a good representation of HPV16 transformation stages and suited to test E1C overexpression effects on HPV16 transcripts.

The HPV16 transcripts E6*I, E1^E4 and E7 were analyzed upon E1C overexpression in W12-epi, W12-int, CaSki and SiHa because of their functional roles in cell cycle progression and transformation. The prevalence of HPV16 E6*I was found to be 8% in NIL/M and rising to more than 95% in CIN3 and CxCa[85]. Zhi-Ming Zheng's lab has studied HPV16 E6 splicing in depth. They have shown that E6*I transcripts which are spliced from 226 nt at the 5' to 409 nt at the 3' are abundant in CaSki and SiHa cells[166]. They have also shown that the majority of the E7 proteins are translated from E6*I mRNAs[166] in these cells. The E6*I splice is present only in high-risk HPVs and not in low-risk ones[167-169]. This could mean that E6*I might have a role in viral oncogenesis, another alternative suggestion being a role in viral life cycle completion[170]. Another group that has studied about both HPV16 E6*I transcript and protein is Penelope Duerksen-Hughes and colleagues. They have shown that HPV16 E6*I overexpression in SiHa cells increased the expression of p53 and sensitized the cells to apoptosis[175]. They have also observed a reduction in tumor formation in mice when HPV16 E6*I overexpressing SiHa cells were injected compared to the control mice. This is the first study which demonstrated a biological activity for E6*I[175].

The prevalence of HPV16 E1^E4 was found to be 6% in NIL/M, increasing to >90% in CIN1and CIN2 and slightly decreasing to 83% in CIN3 to a substantial decline of 33% in CxCa[85]. HPV16 E1^E4 protein was found to be abundantly present in the upper epithelium in productive infections[171]. It has also been found to associate with cytokeratin networks in keratinocytes leading to its collapse[172,173]. HPV16 E1^E4 can also bind to a DEAD box RNA helicase, also known as E4-DBP and change its ATPase activity. Davy and colleagues have shown that HPV16 E1^E4 can affect cell cycle progression by

causing the cells to arrest in the G_2 phase of the cell cycle[171]. HPV16 E7 protein is one of the two viral oncoproteins. It binds and degrades the cellular tumor suppressor protein pRb, induces genome instability and has immortalization and transformation functions.

E1C overexpression in stably lentivirally transduced W12-epi, W12-int, CaSki and SiHa cell lines has not changed E6*I and E7 expression. This means the transformation pathways that are affected by altered E6*I and E7 expression levels are not further affected by overexpressed E1C levels. Although E1C is described as a specific marker for ≥CIN3 cases[85], the occurence of the transcript in severe lesions seems to have no transformation promoting roles and works independent of E6*I and E7 transcript levels. Alternatively, as E6*I protein possesses an anti-oncogenic function, an increase in E6*I expression could inhibit cell proliferation and vice versa. Therefore, it is tempting to speculate that E1C expression maintains the cell cycle progression in CxCa cell lines by maintaining the homeostatic levels of E6*I and E7. However, this hypothesis demands further investigation.

The E1^E4 expression was reduced 20 - 30% in E1C overexpressing W12-epi, W12-int and CaSki cells. Dave and colleagues[171] have previously demonstrated the ability of HPV16 E1^E4 protein to inhibit cell proliferation. Also, it may be possible that this ability of E1^E4 is required to inhibit cell division during the virus life cycle. Therefore, a lower E1^E4 RNA expression induced by E1C could be to prevent the cells from G_2 arrest[171] and maintain cell cycle progression in these cells. However, it is not clear how E1^E4 transcription and splicing is targeted. It might be an HPV genome integration-independent mechanism as E1^E4 RNA expression was reduced in both W12-epi and W12-int cells upon E1C overexpression. It will be also interesting to look at the E1^E4 protein expression upon E1C overexpression. The C-terminal region of the E1C protein contains the E1 helicase domain which has ATPase activity. HPV16 E1^E4 protein is known to change the ATPase activity of a RNA helicase protein[171]. So far the data point towards an indirect role of E1C in transformation by not changing the RNA expression levels of the major transformation transcripts such as E6*I and E7 and by reducing E1^E4 levels, thereby aiding the cells to remain in proliferative stage. As next steps, it would be interesting to look at the entire HPV16 as well as cellular transcriptomes to analyze the

changes in the RNA expression profiles of both virus and host upon E1C overexpression. This would shed more light on the other key players that orchestrate either upstream or downstream of E1C to execute its functions.

When the E1C overexpressed W12-epi, W12-int, CaSki and SiHa cells were analysed to identify cellular proteome changes, no significant changes were observed except in the case of CaSki cells, where PDIA3 protein was found to be two fold upregulated in E1C overexpressed CaSki cells. Since it is just one protein out of the 4899 proteins analysed in MS, this result demands reproduction and further verification by WB. It will also be interesting to examine PDIA3 RNA levels in E1C overexpressed CaSki cells. PDIA3 belongs to protein disulfide isomerase (PDI) family of proteins. PDI proteins are important for the proteostasis in endoplasmic reticulum[229]. However, several studies have demonstrated the altered expression of PDI proteins in various cancers, for e.g., PDIA3 has been reported to be upregulated in primary ductal breast cancer[230]. Another PDI protein, AGR2 which is highly expressed in ovarian and prostate cancers, can stimulate the proliferation and metastasis of cancer cells[231]. Accordingly, PDI inhibitors have shown to enhance apoptosis in cancer cells[232]. Zhang and colleagues have shown that interrupting the interation between endoplasmic reticulum oxidoreductin-1α and PDI proteins can suppress cervical cancer progression[233]. Since PDIA3 was also observed to be upregulated in E1C overexpressed CaSki cells, it will be intriguing to look at the upstream and downstream components of PDIA3 pathway that are involved in tumorigenesis.

5.3. Detection of E1C protein

Usually viruses with small genomes as that of HPV (8 kb) use multiple strategies to increase the coding capacity of the genome to meet their various needs in the completion of life cycle. Alternative splicing of mRNA is one of the mechanisms used by HPV to synthesize multiple proteins. Some of the proteins from alternatively spliced HPV mRNAs that have been closely studied are E1^E4, E8^E2C and E6*I. As described in Chapter 4.2., E1^E4 protein plays a role in the productive life cycle of HPV. Also, HPV11 and

HPV16 E1^E4 can induce G_2 arrest[171]. The spliced E8^E2C fusion protein from HPV31 has a negative regulatory effect on viral transcription and replication[176,177]. As there is increasing and compelling evidence of biologically active proteins from alternatively spliced products of HPV, I decided to look for the translated product of the spliced E1C transcript, i.e. E1C protein. E1C protein is of 9 kDa size and consists of 82 aa, of which the first 5 aa are from the N-terminus and the remaining 77 aa are from the C-terminus of E1.

A mAb was generated by hybridoma technology for the detection of E1C protein in various immunological assays such as ELISA, WB and IF. The mAb generated is directed against the splice junction-derived peptide sequence of E1C to ensure the unique E1C specificity of the antibody. The newly generated E1C mAb could not detect overexpressed HPV16 GST-E1fl fusion protein in *E. coli*, but detected the overexpressed GST-E1C fusion protein in *E. coli* in ELISA and WB, and detected overexpressed GFP-E1C fusion protein in HEK-293T cells in WB and IF. Thus, a functional E1C-specific mAb was successfully generated.

However, for the detection of overexpressed untagged E1C protein in WB, the standard WB protocol had to be modified due to the very small size of E1C. The first WBs were run with mini SDS gels of 20 µL well volume capacity and loaded with a maximum of 20 µg lysates. The transfer of proteins from gels to nitrocellulose membrane was performed at a current of 200 mA for 1h using a SDS containing transfer buffer. When 20 µg of p2-E1C- and p1-His-GFP-E1C-transfected HEK-293T cellular lysates were run in a WB under these conditions, only a 36 kDa band corresponding to GFP-E1C fusion protein was visible and not the 9 kDa E1C protein. Therefore, the next step was to ensure the detection of a comparably small protein in WB. To this end, a truncated version of ABCB5, the ATP Binding Cassette Subfamily B Member 5 protein was used. The N-terminally His-tagged truncated ABCB5 is of 10 kDa size and thus suited as a size detection threshold for the WB protocol and as proxy for the detectability of the 9 kDa E1C protein. The truncated ABCB5 was overexpressed in *E. coli* and purified by affinity chromatography. In addition, the WB protocol was modified to use bigger gels with 50 µL well volume capacity and the protein transfer conditions were changed to 250 mA for 30 minutes using

a transfer buffer without SDS. The absence of SDS and the shorter transfer time ensured that small proteins do not migrate too fast and get lost through the membrane. The modified WB protocol conditions were also tested for the sensitivity of both purified ABCB5 protein and p1-His-GFP-E1C-transfected HEK-293T cellular lysate. The purpose of this sensitivity test was to analyze the detection limit of the modified WB protocol in terms of protein amounts. Results showed that 33 ng of purified ABCB5 and 5 µg of p1-His-GFP-E1C-transfected HEK-293T cellular lysate were detectable with the modified WB protocol. This means that it could be possible to detect E1C in lysate of p2-E1C-transfected HEK-293T cells with the newly modified WB protocol.

For the detection of overexpressed E1C protein in p2-E1C-transfected HEK-293T cells, the modified WB protocol was used together with 50 ng of ABCB5 and 100 µg of p1-His-GFP-E1C-transfected HEK-293T cellular lysate. 100 µg protein amount was also loaded for p2-E1C-transfected HEK-293T cellular lysate. However, only 10 kDa ABCB5 and 36 kDa GFP-E1C bands were visible and not the 9 kDa E1C band. There could be several reasons for not detecting E1C protein out of which I tested the following three: 1) insufficient E1C transcription from the E1C cDNA sequence present in p2-E1C plasmid 2) insufficient E1-Alpha promoter activity to transcribe the E1C cDNA sequence in p2-E1C plasmid 3) instability of E1C protein by treatment with MG132 proteasome inhibitor.

E1C transcription from E1C cDNA sequence in p2-E1C plasmid followed by transfection in HEK-293T cells was tested by E1C RT-qPCR. The E1C copies per cell were quantified to be 18,500. This means that the transfection was successful and there was no blockade present to transcribe the E1C cDNA. HEK-293T cells are easy-to-transfect cells as also demonstrated here by the high number of E1C RNA copies per cell measured. This also indicated that the overexpressed E1C mRNA was stable and abundantly present at least until 48 h post transfection.

Next, I tested for E1-Alpha promoter activity for its capability in transcribing E1C cDNA. The p2 plasmid has the same vector sequence as that of p2-E1C except that the E1C cDNA sequence is replaced with Gaussia luciferase cDNA sequence in p2. This means E1C cDNA sequence in p2-E1C plasmid and Gaussia luciferase cDNA sequence in p2 are driven by the E1-Alpha promoter. Gaussia luciferase activity was measured in p2-

transfected HEK-293T cells and was comparable to the positive control (data not shown). This means that E1-Alpha promoter is capable of driving the E1C RNA expression in p2-E1C-transfected HEK-293T cells.

The third parameter that has been tested as potential reason for not detecting E1C protein in p2-E1C-transfected HEK-293T cells was to check for instability of the putative E1C protein. This means that although E1C protein is translated from the highly overexpressed E1C RNA, the protein is not stable and gets degraded immediately upon translation. The involvement of the ubiquitin-proteasome pathway in the regulation of various cellular processes such as oncogenesis, cell cycle progression, apoptosis, protein quality control etc. has been well studied[178]. This pathway in mammalian cells is responsible for the degradation of both short- and long-lived proteins[179,180]. The main targets of this pathway are proteins that are incorrectly folded or damaged[181]. A protein is targeted for degradation by poly-ubiquitinylation which serves as a tracking signal for the proteasome. Hence, I decided to test whether in p2-E1C-transfected HEK-293T cells, that could translate E1C from the abundant E1C RNA, the E1C protein might be degraded by the proteasome pathway.

There are various proteasome inhibitors available which target multiple active sites of the proteasome like chymotrypsin-like, trypsin-like or caspase-like sites[178]. MG132 is a reversible inhibitor and belongs to the peptide aldehyde class. It can enter cells rapidly and the inhibitor effect can be rapidly reversed by removing MG132 from the cells[182]. In addition, MG132 is more potent and selective for proteasome rather than other cellular proteases compared to other peptide aldehyde inhibitors[183]. MG132 is also available at low cost and owing to all these characteristics, it is the favorite candidate to study the involvement of proteasome in a biological process in cell culture systems [178]. Hence, I decided to treat p2-E1C-transfected HEK-293T cells with MG132 proteasome inhibitor to analyze whether the E1C protein translated is being degraded via ubiquitin-proteasome pathway. 12 h, 24 h and 48 h post transfection time points were chosen for harvesting MG132-treated cells to ensure the capture of the E1C protein soon after translation. A 9 kDa E1C protein was detected in MG132-treated p2-E1C-transfected HEK-293T cells which were harvested 48 h post transfection but not 12 h and 24 h post transfection in

WB. However, with an apparently more sensitive method like MS, E1C protein was detectable in all these three cases with increasing intensity from 12 h to 48 h, which indicates E1C protein accumulation over time. This means E1C RNA translation started less than 12 h post transfection and the protein accumulated over the next 36 h. This means that at 48 h post transfection, there were more cells present due to cell division, therefore the p2- E1C plasmids were replicated accordingly and thereby more E1C RNA transcribed and more E1C protein translated which was prevented from degradation by MG132. E1C protein was found to be localized in cytoplasm after MG132 treatment in p2-E1C-transfected HEK-293T cells in IF. However, only 20% of the p2-E1C-transfected cells were stained positive for E1C. This could be due to the heterogeneity in proteasome inhibition in different cells, which means E1C protein was visible in the cells where there was higher proteasome inhibition and the IF protocol might not be sensitive enough to detect low E1C protein amounts. However, it would be interesting to quantify the E1C positive staining signal intensity to understand the IF detection thresholds.

In summary, the proteasome inhibitor experiments show that E1C protein is unstable which could impair detectability of endogenous E1C protein in CxCa cell lines or ≥CIN3 clinical cases. This could also mean that E1C executes its functions rather via RNA and not via protein. Accordingly, no endogenous E1C protein was detected in HPV16 positive cell lines CaSki, MRI-H186, MRI-H196 and HPK-1A. However, in these analyses they were not treated with MG132. But, the endogenous E1C RNA present in these cell lines are in the range of 0.0004 – 0.014, which makes the detection of E1C protein after MG132 treatment less likely in light with the results obtained so far.

There are several factors that determine the stability of a protein like pH, temperature, salt concentration, protein-protein interactions etc. A folded protein has higher stability than the one in its native state. Folding of a protein after its synthesis is necessary to acquire its functions[195], for e.g. post-translational modifications on the aa sequence of the primary structure of a protein can alter the structure and function and also stability of the protein. Also, secondary structures of the protein can be unstable if there is a proline aa on the α-helix, because proline misses an amide H atom and therefore cannot form stable hydrogen bonds[205]. Some of the factors which affect the stability of tertiary structure of a

protein which means the fold of the protein are pH, temperature, van der Waals forces, hydrogen bonding, primary and secondary structure etc.[194]. With respect to quaternary structure, which is a complex of individual folded protein sub-units, the stability is determined by the conformational changes. It is known that the different stability states of a protein is associated with its function. The best example studied is hemoglobin, switching between tense and relaxed states during the oxygenation process. Deoxy-hemoglobin is in relaxed state and oxyhemoglobin is in tense state[206-208].

Degradation of proteins is part of cellular metabolism. It has been observed that proteases have an increased preference for denatured proteins *in vitro*[210-213]. Accordingly, Parsell and Sauer[209] have shown that thermal stability and sequence composition are key determinants to understand the susceptibility of a protein to proteolysis. Post-translational modifications, cellular environment and protein-protein interactions are some of the factors that also affect the stability of HPV proteins. Thomas et al. have shown that cellular E6AP is destabilised in the presence of HPV-10 and HPV-11 E6 proteins[214]. Chang et al. have shown that when HPV16 E2 is phosphorylated at serine 243, the half-life is reduced compared to wild type E2. However, this modification of E2 is necessary for its binding to cellular Brd4 protein, which facilitates the tethering of viral genome to mitotic chromosomes[215]. Selvey and colleagues have demonstrated that HPV18 E7 is a short-lived protein with a half-life of 13.5 min in HeLa cells[216]. The phosphorylated form (at serine 33) of E7 showed longer half-life than de-phosphorylated form, when overexpressed in Sf21 cells[216]. They have also shown that the steady state levels of E7 increased in the presence of hydrocortisone and not with progesterone, oestrogen or testosterone[216], which means cellular hormonal environment also plays a role in the steady state levels of proteins. Another HPV protein which has a short half-life is HPV18 E6*I. When untagged E6*I proteins were expressed in human 293 cells and treated with cycloheximide 24 h later to prevent *de novo*protein synthesis, the half-life of E6*I was determined to be 6 h[217].

5.4. Effect of E1C transcript and E1C protein on URR activity

Since the HPV URR contains the regulatory elements such as promoters and enhancers which are required for the viral gene regulation, extensive studies have been undertaken to understand the effect of various viral genes on the activity of URR. Accordingly, Alloul et al. studied the effects on HPV16 URR of various E2 transcripts from HPV16, which contain E2 ORFs connected to variable upstream 5' ORFs by alternative spliceing[63]. a'-type mRNA which has the coding potential of both E2 and E1C, had a repressive effect on URR activity. But a mutant of a'-type mRNA truncated in the E2 ORF activated the URR, which they suggested might have been driven by the putative E1C protein. Accordingly, they showed that E1C cDNA activated the URR and suppressed the repression of URR by E2[63]. I decided to confirm the URR activation by E1C together with E1C mutants to corroborate the findings from Alloul et al.. They had performed the experiments in a chloramphenicol acetyltransferase (CAT) assay, a classic method used in the past in many promoter studies. However, although CAT assay has a good signal-to-background ratio, but it is a rather insensitive and time-consuming procedure requiring radioactivity[184,185]. Therefore, I decided to use DLRA, a simple 2 h procedure which is widely used in the last years to study the promoter activity and DNA replication of HPVs and has an exceptional sensitivity and wide dynamic range.

Of the three cell lines initially tested for DLRA, HEK-293T cells, an epithelial cell line derived from human embryonic kidney by adenovirus immortalization and additional expressing SV40 T antigen, upon transfection of an HPV16-URR luciferase construct showed the maximum HPV16 URR activity. URR activity was lowest in human cervical carcinoma-derived C33A cells and only 1.5 times higher in mouse NIH-3T3 fibroblasts. Alloul et al. had described that the basal transcription of HPV16 URR-CAT reporter plasmid in NIH-3T3 and canine Cf2Th fibroblasts is higher than in C33A cells. I chose only HEK-293T cells for DLRA since it showed the highest URR activity and therefore will be possible to detect a wide range of repressed as well as activated URR signals[63].

In DLRAs, a control plasmid is always co-transfected with the reporter plasmid to allow normalization for transfection efficiency. Alloul et al. co-transfected a CMV-ß-galactosidase plasmid as control plasmid for normalization. Of two control plasmids

tested in DLRA, p-TATAbox-RLuc was chosen over p-SV40-RLuc because of its lower promoter *trans* effect. However, when the effect of E2 on URR was analyzed, RLuc activities were found to be consistently repressed in ten consecutive transfections and therefore normalization of FLuc activites by RLuc was omitted. Only raw FLuc activities were used for DLRA data analysis. However, when FLuc activities were analyzed for the DLRAs with and without p-TATAbox-RLuc co-transfections, the variations observed were negligible. This indicates the consistency of the transfections and the entire DLRA protocol performed and therefore, analyzing URR activity without the normalization to a control plasmid was considered acceptable.

E2 repressed URR activity in p-HPV16-E2 co-transfected HEK-293T cells which was in accordance with Alloul et al.[63]. In DLRA, 100 ng of p-HPV16-E2 produced a mean reduction of 35% in URR activity, whereas in CAT assay, Alloul et al. observed 69% reduction with 2000 ng of E2 plasmid co-transfected with LCR-CAT plasmid in Cf2Th cells, which is roughly ten times less reduction when compared to DLRA and assuming a linear dose-response relationship. This could be due to the different cell lines and different transfection efficiencies and/or higher sensitivity of DLRA. For e.g., Schenker and colleagues observed that the extent of repression of different HPV URRs by corresponding E2 proteins is cell-type dependent[186].

E1C activated URR activity in p-HPV16-E1C co-transfected HEK-293T cells which was in accordance with Alloul et al.[63]. In DLRA, 100 ng of p-HPV16-E1C increased URR activity by a mean of 42%, whereas in CAT assay, Alloul et al. observed 48% activation with 2000 ng of E1C plasmid, which is roughly 18 times less activation when compared to DLRA and assuming a linear dose-response relationship. This could be due to the same reasons as discussed for E2 above.

Since E1C protein was found to be unstable, it was interesting to test whether URR activation by E1C is mediated through transcript and/or protein. p-HPV16-E1C-Pro mut co-transfected HEK-293T cells showed a mean URR activation of 33%. URR activation by p-HPV16-E1C-Pro mut indicates that E1C RNA is capable of activating URR because p-HPV16-E1C-Pro mut plasmid cannot encode an E1C protein. However, the mechanism by which E1C RNA activates URR for e.g., by URR binding needs further investigation.

There are studies about different cellular proteins that can bind to HPV mRNAs that control HPV gene expression[187], for e.g., HuR binding to the late 3'-UTR of HPV16[191]. Also, it has been suggested that HPV E4 might have a role in HPV mRNA splicing regulation[188]. HPV E2 is a DNA binding protein which binds to URR to control transcription. There is also increasing evidence which supports the idea that E2 might also control HPV- and cellular-RNA splicing[189] and HPV early polyadenylation[190], thus acting as a RNA processing factor. Although there is a wealth of information about cellular proteins binding to HPV mRNAs or HPV proteins binding to HPV genomes, there is no knowledge about a HPV RNA binding to URR, and therefore this is the first study that hints to such a possibility.

p-HPV16-E1C-RNA mut co-transfected HEK-293T cells unexpectedly also activated the URR by a mean of 19%. E1C-RNA-mut transcribes a random RNA sequence with length and GC content identical to E1C RNA. This could mean that URR activation by E1C RNA is not really E1C specific and URR could merely be activated by a RNA which possess the same GC content as that of E1C RNA. However, this should be tested with more random sequence RNAs with the same length GC content as E1C RNA.

For E2 and E1C co-transfections, p-HPV16-E2 alone repressed URR activity by a mean of 17%. However, p-EV and p-HPV16-E2 together repressed by a mean of 62% compared to p-EV alone. This means that the addition of p-EV introduced some promoter *trans* effect causing additional URR repression apart from repression by E2. This also indicates susceptibility of the complex DLRA co-transfection system with subtle changes causing high oscillations and the need for multiple repetitions of the experiment to generate robust and consistent results.

p-HPV16-E1C and p-HPV16-E2 co-transfection activated the URR by a mean of 24% when compared to p-EV and p-HPV16-E2 co-transfection which is the specific negative control. However, p-EV and p-HPV16-E2 co-transfection repressed by a mean of 54% when compared to p-HPV16-E2 alone. And, p-HPV16-E1C and p-HPV16-E2 together repressed the URR by a mean of 44% when compared to p-HPV16-E2 alone and 66% when compared to p-HPV16-E1C alone. Together these data show that E1C reduces the repressive effect of E2 on URR in E2 and E1C co-transfections and is in accordance with

data shown by Alloul et al.. They have observed a reduction in the E2-mediated transrepression of URR in E2 and E1C co-transfection experiments in Cf2Th cells in a dose-dependent fashion[63]. When E2 was co-transfected with p-HPV16-E1C-RNA mut plasmid, there was a mean of 40% repression of URR and 24% repression with p-HPV16-E1C-Pro mut plasmid which are 14% and 30% lower repressions compared to E2 transfection alone. These data again indicate that it is the E1C RNA which is activating URR and not E1C protein.

6. Outlook

In this thesis, HPV16 E1C, which had been identified as a useful diagnostic marker transcript for the discrimination of low grade and high grade cervical lesions[85,152], has been characterized. E1C was overexpressed by lentiviral transduction in four HPV16 positive cell lines and E1^E4 transcript was found to be 20 – 30 % significantly downregulated. As E1^E4 protein is shown to inhibit cell proliferation[171] it would be interesting to check for proliferation in E1C overexpressed cells by quantifying DNA synthesis or cellular metabolic changes or proliferation proteins such as PCNA, Ki67 or MCM2. The tumorigenic potential of E1C overexpressed cells could also be assessed by checking for clonogenicity in a colony forming assay or anchorage-independent growth potential in a soft agar assay[193]. In addition, analyzing changes in global cellular transcriptome under overexpressed E1C condition by microarray or RNA-seq technologies is also intriguing.

Although E1C RNA can be stably overexpressed, E1C protein is found to be unstable. As there are several factors, which can be attributed to the instability of a protein, quinary interactions, cellular crowding environment or primary – secondary – tertiary – quaternary structure of E1C protein could be studied to understand its instability[194,195]. Also looking for post-translational modifications such as phosphorylation, acetylation or methylation which can regulate protein stability is also compelling[196].

E1C RNA was found to activate HPV16 URR in a dual reporter luciferase assay. Different cellular RNA binding proteins have been identified that could bind to HPV16 mRNAs to control HPV16 gene regulation[197]. In this context, in order to unravel the mechanism with which the URR activation is mediated by E1C RNA, it would be interesting to look for E1C RNA interactions with cellular proteins. It is also intriguing to check whether E1C RNA interacts with other HPV16 proteins or is able to bind to the URR. Results from these experiments should provide evidence whether the URR activation by E1C RNA is a direct or indirect mechanism.

Furthermore, the C-terminal module of E1fl protein from papillomaviruses was shown to play a role in the stabilization of E1 hexamer during viral DNA replication[198-200]. Therefore,

E1C protein overexpressed under proteasome inhibition conditions could also be validated for its capability to support viral DNA replication.

7. Abbreviations

Bidest.	Bidestillatus
BSA	Bovine Serum Albumin
bp	base pair
BSA	Bovine Serum Albumin
ºC	degree Celsius
CIN	Cervical Intraepithelial Neoplasia
CMV	Cytomegalovirus
Cp	Crossing point
CV	Coefficient of Variation
CxCa	Cervical Cancer
DKFZ	Deutsches Krebsforschungszentrum
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
e.g.	exempli gratia
EtOH	Ethanol
FBS	Fetal Bovine Serum
fl	full length
h	hour(s)
HPV	Human Papillomavirus
hr	high-risk
HSIL	High-grade Squamous Intraepithelial Lesion
i.e.	id est

IF	Immunofluorescence
IRES	Internal Ribosomal Entry Site
lr	low risk
kb	kilo base
LSIL	Low-grade Squamous Intraepithelial Lesion
LV	Lentiviral
mAb	monoclonal Antibody
MCS	Multiple Cloning Site
min	minute(s)
mM	millimolar
mRNA	messenger RNA
MS	Mass Spectrometry
µg	microgram(s)
µL	microliter
ng	nanogram(s)
NIL/M	Negative for Intraepithelial Lesion or Malignancy
No:	Number
nt	nucleotide
ORF	Open Reading Frame
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PEI	Polyethyleneimine
pRb	Retinoblastoma protein
PV	Papillomavirus
q	quantitative
RIN	RNA Integrity Number

RNA	Ribonucleic acid
rpm	rounds per minute
RT	Room Temperature
RT-qPCR	Reverse Transcriptase quantitative PCR
s	second(s)
sj	splice junction
SV40	Simian Virus 40
ubC	ubiquitin C
URR	Upper Regulatory Region
Vs.	Versus
WB	Western blot

8. References

1**.** Bravo IG, et al. ***The clinical importance of understanding the evolution of papillomaviruses.*** Trends in Microbiology 2010; 18(10): 432-438.

2**.** Bernard HU, et al. ***Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments.*** Virology 2010; 401(1): 70–79.

3**.** Doorbar J, et al. ***Human papillomavirus molecular biology and disease association.*** Rev Med Virol. 2015 Mar;25 Suppl 1:2-23.

4. de Villiers, et al. ***Classification of papillomaviruses.*** Virology. 2004 Jun 20; 324(1):17-27

5. Zheng ZM, Baker CC. ***Papillomavirus genome structure, expression, and posttranscriptional regulation.*** Frontiers in Bioscience 2006; 11: 2286–2302.

6. Schiller JT, Lowy DR. ***Understanding and learning from the success of prophylactic human papillomavirus vaccines.*** Nature Reviews. Microbiology 2012; 10(10): 681–692.

7. Chen XS, et al. ***Structure of small virus-like particles assembled from the L1 protein of human papillomavirus 16.*** Molecular Cell 2000; 5(3): 557–567.

8. Rubio I, et al. ***The N-terminal region of the human papillomavirus L2 protein contains overlapping binding sites for neutralizing, cross-neutralizing and non-neutralizing antibodies.*** Virology 2011; 409(2): 348–359.

9. Liu WJ, et al. ***Sequence close to the N terminus of L2 protein is displayed on the surface of bovine papillomavirus type 1 virions.*** Virology 1997; 227(2): 474–483.

10. Kines RC, et al. ***The initial steps leading to papillomavirus infection occur on the basement membrane prior to cell surface binding.*** Proceedings of the National Academy of Sciences of the United States of America 2009; 106(48): 20458–20463.

11. de Villiers EM. ***Cross-roads in the classification of papillomaviruses.*** Virology 2013; 445(1–2): 2–10.

12. Doorbar J, et al. ***The biology and life-cycle of human papillomaviruses.*** Vaccine 2012; 30 (Suppl 5): F55–F70.

13. Doorbar J. ***The E4 protein; structure, function and patterns of expression.*** Virology 2013; 445(1–2): 80–98.

14. Dimaio D, Petti LM. ***The E5 proteins.*** Virology 2013; 445(1–2): 99–114.

15. Graham SV1.***Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies.*** Future Microbiol. 2010 Oct;5(10):1493-506.

16. Patterson NA, Smith JL, Ozbun MA. ***Human papillomavirus type 31b infection of human keratinocytes does not require heparan sulfate.*** J. Virol. 2005; 79(11):6838–47.

17. Shafti-Keramat S, Handisurya A, Kriehuberm E, Slupetzky K, Kirnbauer R. ***Different heparin sulfate proteoglycans serve as cellular receptors for human papillomaviruses.*** J. Virol. 2003; 77(24):13125–35.

18. Evander M, Frazer IH, Payne E, et al. ***Identification of the alpha 6 integrin as a candidate receptor for papillomaviruses.*** J. Virol. 1997; 71(3):2449–56.

19. McMillan NA, Payne E, Frazer IH, Evander M. ***Expression of the α6 integrin confers papillomavirus binding upon receptor-negative B-cells.*** Virol. 1999; 261(2):271–9.

20. Smith JL, Campos SK, Ozbun MA. ***Human papillomavirus type 31 uses a caveolin 1- and dynamin 2-mediated entry pathway for infection of human keratinocytes.*** J. Virol. 2007; 81(18):9922–31.

21. Hamid NA, Brown C, Gaston K. ***The regulation of cell proliferation by the papillomavirus early proteins.*** Cell. Mol. Life Sci. 2009; 66(10):1700–17.

22. zur Hausen, H., 2002. ***Papillomaviruses and cancer: from basic studies to clinical application.*** Nat. Rev. Cancer 2, 342– 350.

23. Smotkin D, Wettstein FO. ***Transcription of human papillomavirus type 16 early genes in a cervical cancer and a cancer-derived cell line and identification of the E7 protein.*** Proc. Natl. Acad. Sci. USA. 1986; 83(13):4680–4.

24. Ozbun MA, Meyers C. ***Temporal usage of multiple promoters during the life cycle of human papillomavirus type 31b.*** J. Virol. 1998; 72(4):2715–22.

25. Thierry F, Howley PM. ***Functional analysis of E2-mediated repression of the HPV18 p105 promoter.*** New Biol. 1991; 3(1):90–100.

26. Thierry F, Heard JJ, Dartmann K, Yaniv M. ***Characterisation of a transcriptional promoter of human papillomavirus 18 and modulation of its expression by simian virus 40 and adenovirus early antigens.*** J. Virol. 1987; 61(1):134–42.

27. Grassman K, Rapp B, Maschek H, Petry KU, Iftner T. ***Identification of a differentiation-inducible promoter in the E7 open reading frame of human papillomavirus type 16 (HPV-16) in raft cultures of a new cell line containing high copy numbers of episomal HPV-16 DNA.*** J. Virol. 1996; 70(4): 2339–49.

28. Carson A, Khan SA. ***Characterisation of transcription factor binding to human papillomavirus type 16 DNA during cellular differentiation.*** J. Virol. 2006; 80(9):4356–62.

29. Woolridge T, Laimins LA. ***Regulation of human papillomavirus type 31 gene expression during the differentiation-dependent life cycle through histone modifications and transcription factor binding.*** Virol. 2008; 374(2):371–80.

30. Milligan SG, Veerapraditsin T, Ahamat B, Mole S, Graham SV. ***Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells.*** Virol. 2007; 360:172–81.

31. Del Mar Pena L, Laimins LA. ***Differentiation-dependent chromatin rearrangement coincides with activation of human papillomavirus type 31 late gene expression.*** J. Virol. 2001; 75(20):10005–13.

32. Zheng Z-M, Baker CC. ***Papillomavirus genome structure, expression, and post-transcriptional regulation.*** Front. Biosci. 2006; 11:2286–302.

33. Ozbun MA, Meyers C. ***Characterisation of late gene transcripts expressed during vegetative replication of human papillomavirus type 31b.*** J. Virol. 1997; 71(7):5161–72.

34. Doorbar J, Parton A, Hartley K, et al. ***Detection of novel splicing patterns in a HPV-16-containing keratinocyte cell line.*** Virol. 1990; 178(1):254–62.

35. Wang Z, Burge CB. ***Splicing regulation: from a parts list of regulatory elements to an integrated splicing code.*** RNA. 2008; 14(5):802–13.

36. Black DL. ***Mechanisms of alternative pre-messenger RNA splicing.*** Annu. Rev. Biochem. 2003; 72:291–336.

37. Long JC, Caceres JF. ***The SR protein family of splicing factors: master regulators of gene expression.*** Biochem. J. 2009; 417(1):15–27.

38. Mole S, McFarlane M, Chuen-Im T, et al. ***RNA splicing factors regulated by HPV16 during cervical tumour progression.*** J. Pathol. 2009; 219:383–91.

39. Mole S, Milligan SG, Graham SV. ***Human papillomavirus type 16 E2 protein transcriptionally activates the promoter of a key cellular splicing factor, SF2/ASF.*** J. Virol. 2009; 83(1):357–67.

40. Stubenrauch F, Colbert AME, Laimins LA. ***Transactivation by the E2 protein of oncogenic human papillomavirus type 31 is not essential for early and late viral functions.*** J. Virol. 1998; 72(10): 8115–23.

41. Sokolowski M, Furneaux H, Schwartz S. ***The inhibitory activity of the AU-rich RNA element in the human papillomavirus type 1 late 3′ untranslated region correlates with its affinity for the elav-like HuR protein.*** J. Virol. 1999; 73(2):1080–91.

42. Sokolowski M, Zhao C, Tan W, Schwartz S. ***AU-rich mRNA instability elements on human papillomavirus type 1 late mRNAs and c-fos mRNAs interact with the same cellular factors.*** Oncogene. 1997; 15(19):2303–19.

43. Sokolowski M, Schwartz S. ***Heterogeneous nuclear ribonucleoprotein C binds exclusively to the functionally important UUUUU-motifs in the human papillomavirus type-1 AU-rich inhibitory element.*** Virus Res. 2001; 73(2):163–75.

44. Tan W, Schwartz S. ***The rev protein of human immunodeficiency virus type 1 counteracts the effect of an AU-rich negative element in the human papillomavirus type 1 late 3′ untranslated region.*** J. Virol. 1995; 69(5):2932–

45. Zhao C, Tan W, Sokolowski M, Schwartz S. ***Identification of nuclear and cytoplasmic proteins that interact specifically with an AU-rich, cis-acting inhibitory sequence in the 3′ untranslated region of human papillomavirus type 1 late mRNAs.*** J. Virol. 1996; 70(6):3659–67.

46. Cumming SA, McPhillips MG, Veerapraditsin T, Milligan SG, Graham SV. ***Activity of the human papillomavirus type 16 late negative regulatory element is partly due to four weak consensus 5'splice sites that bind a U1 snRNP-like complex.*** J. Virol. 2003; 77(9):5167–77.

47. Cumming SA, Repellin CE, McPhillips M, et al. ***The human papillomavirus type 31 late 3' untranslated region contains a complex bipartite negative regulatory element.*** J. Virol. 2002; 76(12):5993–6003.

48. Chuen-Im T, Zhang J, Milligan SG, McPhillips MG, Graham SV. ***The alternative splicing factor hnRNP A1 is up-regulated during virus-infected epithelial cell differentiation and binds the human papillomavirus type 16 late regulatory element.*** Virus Res. 2008; 131:189–98.

49. Cumming SA, Chuen-Im T, Zhang J, Graham SV. ***The RNA stability regulator HuR regulates L1 protein expression in vivo in differentiating cervical epithelial cells.*** Virol. 2009; 383:142–9.

50. Koffa MD, Graham SV, Takagaki Y, Manley JL, Clements JB. ***The human papillomavirus type 16 negative regulatory element interacts with three proteins that act at different posttranscriptional levels.*** Proc. Natl. Acad. Sci. USA. 2000; 97(9):4677–82.

51. Dietrich-Goetz W, Kennedy IM, Levins B, Stanley MA, Clements JB. ***A cellular 65kDa protein recognizes the negative regulatory element of human papillomavirus late mRNA.*** Proc. Natl. Acad. Sci. USA. 1997; 94(1):163–8.

52. Kennedy IM, Haddow JK, Clements JB. ***A negative regulatory element in the human papillomavirus type 16 genome acts at the level of late mRNA stability.*** J. Virol. 1991; 65(4): 2093–7.

53. Brennan CM, Steitz JA. ***HuR and mRNA stability.*** Cell. Mol. Life Sci. 2001; 58(2):266–77.

54. Wiklund L, Sokolowski M, Carlsson A, Rush M, Schwartz S. ***Inhibition of translation by UAUUUAU and UAUUUUUAU motifs of the AU-rich RNA instability element in the HPV-1 late 3′ untranslated region.*** J. Biol. Chem. 2002; 277(43):40462–71.

55. Newbury SF. ***Control of mRNA stability in eukaryotes.*** Biochem. Soc Trans. 2006; 34(1):30–4.

56. Zhou J, Lui WJ, Peng SW, Sun XY, Frazer IH. ***Papillomavirus capsid protein expression levels depends on the match between codon usage and tRNA availability.*** J. Virol. 1999; 73(6):4972–82.

57. Zhao KN, Lui WJ, Frazer IH. ***Codon usage bias and A+T content variation in human papillomavirus genomes.*** Virus Res. 2003; 98(2):95–104.

58. Bergvall M, Melendy T, Archambault J. ***The E1 proteins.*** Virology. 2013 Oct;445(1-2):35-56. doi: 10.1016/j.virol.2013.07.020.

59. Loo, Y. M. & Melendy, T. ***Recruitment of replication protein A by the papillomavirus E1 protein and modulation by single-stranded DNA.*** Journal of virology 78, 1605-1615 (2004).

60. Masterson, P. J., Stanley, M. A., Lewis, A. P. & Romanos, M. A. ***A C-terminal helicase domain of the human papillomavirus E1 protein binds E2 and the DNA polymerase alpha primase p68 subunit.*** Journal of virology 72, 7407-7419 (1998).

61. Conger, K. L., Liu, J. S., Kuo, S. R., Chow, L. T. & Wang, T. S. ***Human papillomavirus DNA replication. Interactions between the viral E1 protein and two subunits of human dna polymerase alpha/primase.*** The Journal of biological chemistry 274, 2696-2705 (1999).

62. Han, Y., Loo, Y. M., Militello, K. T. & Melendy, T. ***Interactions of the papovavirus DNA replication initiator proteins, bovine papillomavirus type 1 E1 and simian virus 40 large T antigen, with human replication protein A.*** Journal of virology 73, 4899-4907 (1999).

63. Alloul, N. & Sherman, L. ***Transcription-modulatory activities of differentially spliced cDNAs encoding the E2 protein of human papillomavirus type 16.*** The Journal of general virology 80 (Pt 9), 2461-2470 (1999).

64. Alloul, N. & Sherman, L. ***The E2 protein of human papillomavirus type 16 is translated from a variety of differentially spliced polycistronic mRNAs.*** The Journal of general virology 80 (Pt 1), 29-37 (1999).

65. Bouvard, V., Storey, A., Pim, D. & Banks, L. ***Characterization of the human papillomavirus E2 protein: evidence of trans-activation and trans-repression in cervical keratinocytes.*** The EMBO journal 13, 5451-5459 (1994).

66. Kosel, S., Burggraf, S., Engelhardt, W. & Olgemoller, B. ***Increased levels of HPV16 E6*I transcripts in high-grade cervical cytology and histology (CIN II+) detected by rapid real time RT-PCR amplification.*** Cytopathology 18, 290-299 (2007).

67. Doorbar, J., Ely, S., Sterling, J., McLean, C. & Crawford, L. ***Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network.*** Nature 352, 824-827 (1991).

68. Crusius, K., Rodriguez, I. & Alonso, A. ***The human papillomavirus type 16 E5 protein modulates ERK1/2 and p38 MAP kinase activation by an EGFR-independent process in stressed human keratinocytes.*** Virus genes 20, 65-69 (2000).

69. Kiselev, F. L. et al. ***[Status of the human DNA papillomavirus in cervical tumors].*** Molekuliarnaia biologiia 35, 470-476 (2001).

70. Huibregtse, J. M., Scheffner, M. & Howley, P. M. ***Cloning and expression of the cDNA for E6-AP, a protein that mediates the interaction of the human papillomavirus E6 oncoprotein with p53.*** Molecular and cellular biology 13, 775-784 (1993).

71. Werness, B. A., Levine, A. J. & Howley, P. M. ***Association of human papillomavirus types 16 and 18 E6 proteins with p53.*** Science 248, 76-79 (1990).

72. Zeitler, J., Hsu, C. P., Dionne, H. & Bilder, D. ***Domains controlling cell polarity and proliferation in the Drosophila tumor suppressor Scribble.*** The Journal of cell biology 167, 1137-1146, doi:10.1083/jcb.200407158 (2004).

73. Sedman, S. A. et al. ***The full-length E6 protein of human papillomavirus type 16 has transforming and trans-activating activities and cooperates with E7 to immortalize keratinocytes in culture.*** J Virol 65, 4860-4866 (1991).

74. Smotkin, D., Prokoph, H. & Wettstein, F. O. ***Oncogenic and non-oncogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms.*** Journal of virology 63,1441-1447 (1989).

75. Tang, S., Tao, M., McCoy, J. P., Jr. & Zheng, Z. M. ***The E7 oncoprotein is translated from spliced E6*I transcripts in high-risk human papillomavirus type 16- or type 18-positive cervical cancer cell lines via translation reinitiation.*** Journal of virology 80, 4249-4263 (2006).

76. Stacey, S. N. et al. ***Leaky scanning is the predominant mechanism for translation of human papillomavirus type 16 E7 oncoprotein from E6/E7 bicistronic mRNA.*** Journal of virology 74, 7284-7297 (2000).

77. Stacey, S. N. et al. ***Translation of the human papillomavirus type 16 E7 oncoprotein from bicistronic mRNA is independent of splicing events within the E6 open reading frame.*** J Virol 69, 7023-7031 (1995).

78. Bodily, J. M., Mehta, K. P., Cruz, L., Meyers, C. & Laimins, L. A. ***The E7 open reading frame acts in cis and in trans to mediate differentiation-dependent activities in the human papillomavirus type 16 life cycle.*** Journal of virology 85, 8852-8862, doi:10.1128/JVI.00664-11 (2011).

79. Crum, C. P., Nuovo, G., Friedman, D. & Silverstein, S. J. ***Accumulation of RNA homologous to human papillomavirus type 16 open reading frames in genital precancers.*** J Virol 62, 84-90 (1988).

80. Stoler, M. H., Wolinsky, S. M., Whitbeck, A., Broker, T. R. & Chow, L. T. ***Differentiation linked human papillomavirus types 6 and 11 transcription in genital condylomata revealed by in situ hybridization with message-specific RNA probes.*** Virology 172, 331-340 (1989).

81. Schmitt, M. ***Detection of nucleic acids from human alpha papillomaviruses in the uterine cervix.*** Ruperto-Carola University of Heidelberg Dissertation (2008).

82. Halec, G. ***Molecular evidence for transforming activity of rare and probable/possible high risk human papillomavirus types in cervical cancer.*** Ruperto-Carola University of Heidelberg Dissertation (2012).

83. ***Cancer fact sheets: Cervical cancer.*** (International Agency of Research on Cancer - http://gco.iarc.fr/today/data/pdf/fact-sheets/cancers/cancer-fact-sheets-16.pdf)

84. Guan, P. et al. ***Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer.*** Int. J. Cancer 131, 2349–2359 (2012).

85. Höfler D, Böhmer G, von Wasielewski R, Neumann H, Halec G, Holzinger D, Dondog B, Gissmann L, Pawlita M, Schmitt M. ***HPV16 RNA patterns defined by novel high-throughput RT-qPCR as triage marker in HPV-based cervical cancer precursor screening.*** Gynecol Oncol. 2015 Sep;138(3):676-82.

86. Schiffman M, Doorbar J, Wentzensen N, de Sanjosé S, Fakhry C, Monk BJ, Stanley MA, Franceschi S. ***Carcinogenic human papillomavirus infection.*** Nat Rev Dis Primers. 2016 Dec 1;2:16086.

87. Plummer, M. et al. ***A 2-year prospective study of human papillomavirus persistence among women with a cytological diagnosis of atypical squamous cells of undetermined significance or low-grade squamous intraepithelial lesion.*** J. Infect. Dis. 195, 1582–1589 (2007).

88. Rodríguez, A. C. et al. ***Rapid clearance of human papillomavirus and implications for clinical focus on persistent infections.*** J. Natl Cancer Inst. 100, 513–517 (2008).

89. de Martel, C. et al. ***Cancers attributable to infections among adults with HIV in the United States.*** AIDS 29, 2173–2181 (2015).

90. International Collaboration of Epidemiological Studies of Cervical Cancer et al. ***Carcinoma of the cervix and tobacco smoking: collaborative reanalysis of individual data on 13,541 women with carcinoma of the cervix and 23,017 women without carcinoma of the cervix from 23 epidemiological studies.*** Int. J. Cancer 118, 1481–1495 (2006).

91. International Collaboration of Epidemiological Studies of Cervical Cancer. ***Cervical carcinoma and reproductive factors: collaborative reanalysis of individual data on 16,563 women with cervical carcinoma and 33,542 women without cervical carcinoma from 25 epidemiological studies. Int. J. Cancer*** 119, 1108–1124 (2006).

92. International Collaboration of Epidemiological Studies of Cervical Cancer et al. ***Cervical cancer and hormonal contraceptives: collaborative reanalysis of individual data for 16,573 women with cervical cancer and 35,509 women without cervical cancer from 24 epidemiological studies.*** Lancet 370, 1609–1621 (2007).

93. Plummer, M., Peto, J., Franceschi, S. & International Collaboration of Epidemiological Studies of Cervical Cancer. ***Time since first sexual intercourse and the risk of cervical cancer.*** Int. J. Cancer 130, 2638–2644 (2012).

94. Chung, S.-H., Franceschi, S. & Lambert, P. F. ***Estrogen and ERα: culprits in cervical cancer?*** Trends Endocrinol. Metab. 21, 504–511 (2010).

95. den Boon, J. A. et al. ***Molecular transitions from papillomavirus infection to cervical precancer and cancer: role of stromal estrogen receptor signaling.*** Proc. Natl Acad. Sci. USA 112, E3255–E3264 (2015).

96. Brake, T. & Lambert, P. F. ***Estrogen contributes to the onset, persistence, and malignant progression of cervical cancer in a human papillomavirus-transgenic mouse model.*** Proc. Natl Acad. Sci. USA 102, 2490–2495 (2005).

97. Spurgeon, M. E., Chung, S.-H. & Lambert, P. F. ***Recurrence of cervical cancer in mice after selective estrogen receptor modulator therapy.*** Am. J. Pathol. 184, 530–540 (2014).

98. Middleton K, et al. ***Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers.*** Journal of Virology 2003; 77(19): 10186–10201.

99. S, Munger K. ***The human papillomavirus type 16 E6 and E7 oncoproteins independently induce numerical and structural chromosome instability.*** Cancer Research 2002; 62(23): 7075–7082.

100. Duensing S, Munger K. ***Human papillomavirus type 16 E7 oncoprotein can induce abnormal centrosome duplication through a mechanism independent of inactivation of retinoblastoma protein family members.*** Journal of Virology 2003;77(22): 12331–12335.

101. Gariglio P, et al. ***The role of retinoid deficiency and estrogens as cofactors in cervical cancer.*** Archives of Medical Research 2009; 40(6): 449–465.

102. Ding DC, et al. ***Methylation of the long control region of HPV16 is related to the severity of cervical neoplasia.*** European Journal of Obstetrics, Gynecology, and Reproductive Biology 2009; 147(2):215–220.

103.Burgers WA, Blanchon L, Pradhan S, de Launoit Y, Kouzarides T, Fuks F. ***Viral oncoproteins target the DNA methyl transferases.*** Oncogene. 2007 Mar 8;26(11):1650-5.

104. D'Costa ZJ, Jolly C, Androphy EJ, Mercer A, Matthews CM, Hibma MH. ***Transcriptional repression of E-cadherin by human papillomavirus type 16 E6.*** D'Costa ZJ, Jolly C, Androphy EJ, Mercer A, Matthews CM, Hibma MH.

105. Warren, C. J. et al. ***APOBEC3A functions as a restriction factor of human papillomavirus.*** J. Virol. 89, 688–702 (2015).

106. Stanley, M. A. ***Epithelial cell responses to infection with human papillomavirus.*** Clin. Microbiol. Rev. 25, 215–222 (2012).

107. Heaton, S. M., Borg, N. A. & Dixit, V. M. ***Ubiquitin in the activation and attenuation of innate antiviral immunity.*** J. Exp. Med. 213, 1–13 (2016).

108. Tummers, B. & Burg, S. H. ***High-risk human papillomavirus targets crossroads in immune signalling.*** Viruses 7, 2485–2506 (2015).

109. Habiger, C., Jäger, G., Walter, M., Iftner, T. & Stubenrauch, F. ***Interferon kappa inhibits human papillomavirus 31 transcription by inducing Sp100 proteins.*** J. Virol. 90, 694–704 (2016).

110. Sperling, T. et al. ***Human papillomavirus type 8 interferes with a novel C/EBPβ-mediated mechanism of keratinocyte CCL20 chemokine expression and Langerhans cell migration.*** PLoS Pathog. 8, e1002833 (2012).

111. Matthews, K. et al. ***Depletion of Langerhans cells in human papillomavirus type 16-infected skin is associated with E6-mediated down regulation of E-cadherin.*** J. Virol. 77, 8378–8385 (2003).

112. Stanley, M., Pinto, L. A. & Trimble, C. ***Human papillomavirus vaccines — immune responses.*** Vaccine 30 (Suppl. 5), F83–F87 (2012).

113. Giannini, S. L. et al. ***Enhanced humoral and memory B cellular immunity using HPV16/18 L1 VLP vaccine formulated with the MPL/aluminium salt combination (AS04) compared to aluminium salt only.*** Vaccine 24, 5937–5949 (2006).

114. Einstein, M. H. et al. ***Comparison of the immunogenicity of the human papillomavirus (HPV)-16/18 vaccine and the HPV-6/11/16/18 vaccine for oncogenic non-vaccine types HPV-31 and HPV-45 in healthy women aged 18–45 years.*** Hum. Vaccin. 7, 1359–1373 (2011).

115. Giuliano, A. R. et al. ***Efficacy of quadrivalent HPV vaccine against HPV infection and disease in males.*** N. Engl. J. Med. 364, 401–411 (2011).

116. Lehtinen, M. & Dillner, J. ***Clinical trials of human papillomavirus vaccines and beyond.*** Nat. Rev. Clin. Oncol. 10, 400–410 (2013).

117. Joura, E. A. et al. ***A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women.*** N. Engl. J. Med. 372, 711–723 (2015).

118. Castellsagué, X. et al. ***Immunogenicity and safety of the 9-valent HPV vaccine in men.*** Vaccine 33, 6892–6901 (2015).

119. [No authors listed.] ***Human papillomavirus vaccines: WHO position paper,*** October 2014. WHO Wkly Epidemiol. Rec. 89, 465–492 (2014).

120. Hanson, C. M., Eckert, L., Bloem, P. & Cernuschi, T. ***Gavi HPV programs: application to implementation.*** Vaccines 3, 408–419 (2015).

121. Palefsky, J. M. et al. ***HPV vaccine against anal HPV infection and anal intraepithelial neoplasia.*** N. Engl. J. Med. 365, 1576–1585 (2011).

122. Giuliano, A. R. et al. ***Seroconversion following anal and genital HPV infection in men: the HIM study.*** Papillomavirus Res. 1, 109–115 (2015).

123. Scherpenisse, M. et al. ***Characteristics of HPV-specific antibody responses induced by infection and vaccination: cross-reactivity, neutralizing activity, avidity and IgG subclasses.*** PLoS ONE 8, e74797 (2013).

124. Wheeler, C. M. et al. ***Cross-protective efficacy of HPV-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by non-vaccine oncogenic HPV types: 4-year end-of-study analysis of the randomised, double-blind PATRICIA trial.*** Lancet Oncol. 13, 100–110 (2012).

125. Arbyn, M. et al. ***Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer.*** Vaccine 30 (Suppl. 5), F88–F99 (2012).

126. Gage, J. C. et al. ***Reassurance against future risk of precancer and cancer conferred by a negative human papillomavirus test.*** J. Natl Cancer Inst. 106, dju153 (2014).

127. Ronco, G. et al. ***Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials.*** Lancet 383, 524–532 (2014).

128. Cuzick, J. et al. ***Comparing the performance of six human papillomavirus tests in a screening population.*** Br. J. Cancer 108, 908–913 (2013).

129. Wentzensen, N., Schiffman, M., Palmer, T. & Arbyn, M. ***Triage of HPV positive women in cervical cancer screening.*** J. Clin. Virol. 76, S49–S55 (2016).

130. Massad, L. S., Jeronimo, J., Schiffman, M. & National Institutes of Health/American Society for Colposcopy and Cervical Pathology (NIH/ASCCP) Research Group. ***Interobserver agreement in the assessment of components of colposcopic grading.*** Obstet. Gynecol. 111, 1279–1284 (2008).

131. Bergeron, C. et al. ***Informed cytology for triaging HPV-positive women: substudy nested in the NTCC randomized controlled trial.*** J. Natl Cancer Inst. 107, dju423 (2015).

132. Richardson, L. A. et al. ***HPV DNA testing with cytology triage in cervical cancer screening: influence of revealing HPV infection status.*** Cancer Cytopathol. 123, 745–754 (2015).

133. Wentzensen, N. et al. ***p16/Ki-67 dual stain cytology for detection of cervical precancer in HPV-positive women.*** J. Natl Cancer Inst. 107, djv257 (2015).

134. Ikenberg, H. et al. ***Screening for cervical cancer precursors with p16/Ki-67 dual-stained cytology: results of the PALMS study.*** J. Natl Cancer Inst. 105, 1550–1557 (2013).

135. Carozzi, F. et al. ***Risk of high-grade cervical intraepithelial neoplasia during follow-up in HPV-positive women according to baseline p16-INK4A results: a prospective analysis of a nested substudy of the NTCC randomised controlled trial.*** Lancet Oncol. 14, 168–176 (2013).

136. Huh, W. K. et al. ***Use of primary high-risk human papillomavirus testing for cervical cancer screening: interim clinical guidance.*** Gynecol. Oncol. 136, 178–182 (2015).

137. De Strooper, L. M. A. et al. ***CADM1, MAL and miR124-2 methylation analysis in cervical scrapes to detect cervical and endometrial cancer.*** J. Clin. Pathol. 67, 1067–1071 (2014).

138. Vasiljevic´, N., Scibior-Bentkowska, D., Brentnall, A. R., Cuzick, J. & Lorincz, A. T. ***Credentialing of DNA methylation assays for human genes as diagnostic biomarkers of cervical intraepithelial neoplasia in high-risk HPV positive women.*** Gynecol. Oncol. 132, 709–714 (2014).

139. Mirabello, L. et al. ***Methylation of human papillomavirus type 16 genome and risk of cervical precancer in a Costa Rican population.*** J. Natl Cancer Inst. 104, 556–565 (2012).

140. Vasiljevic´, N., Scibior-Bentkowska, D., Brentnall, A. R., Cuzick, J. & Lorincz, A. ***A comparison of methylation levels in HPV18, HPV31 and HPV33 genomes reveals similar associations with cervical precancers.*** J. Clin. Virol. 59, 161–166 (2014).

141. Wentzensen, N. et al. ***Methylation of HPV18, HPV31, and HPV45 genomes and cervical intraepithelial neoplasia grade 3.*** J. Natl Cancer Inst. 104, 1738–1749 (2012).

142. McCredie, M. R. E. et al. ***Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study.*** Lancet Oncol. 9, 425–434 (2008).

143. Castle, P. E., Schiffman, M., Wheeler, C. M. & Solomon, D. ***Evidence for frequent regression of cervical intraepithelial neoplasia-grade 2.*** Obstet. Gynecol. 113, 18–25 (2009).

144. Trimble C. L. et al. ***Safety, efficacy, and immunogenicity of VGX-3100, a therapeutic synthetic DNA vaccine targeting human papillomavirus 16 and 18 E6 and E7 proteins for cervical intraepithelial neoplasia 2/3: a randomised, double-blind, placebo-controlled phase 2b trial.*** Lancet. 2015 Nov 21;386(10008):2078-2088.

145. Roden RBS, Stern PL. ***Opportunities and challenges for human papillomavirus vaccination in cancer.*** Nat Rev Cancer. 2018 Apr;18(4):240-254.

146. Baker, C., and C. Calef. ***Maps of papillomavirus mRNA transcripts.*** (1996).

147. Hoefler, D. ***HPV16 RNA patterns as diagnostic marker for cervical cancer precursor lesions. Validation by newly developed high-throughput RT-qPCR.*** Ruperto-Carola University of Heidelberg Dissertation (2013).

148. Molden T, Kraus I, Karlsen F, Skomedal H, Nygard JF, Hagmar B. ***Comparison of human papillomavirus messenger RNA and DNA detection: a cross-sectional study of 4,136 women >30 years of age with a 2-year follow-up of high-grade squamous intraepithelial lesion.*** Cancer Epidemiol Biomarkers Prev 2005;14:367–72.

149. Castle PE, Dockter J, Giachetti C, et al. ***A cross-sectional study of a prototype carcinogenic human papillomavirus E6/E7 messenger RNA assay for detection of cervical precancer and cancer.*** Clin Cancer Res 2007;13:2599–605.

150. Cuschieri KS, Whitley MJ, Cubie HA. ***Human papillomavirus type specific DNA and RNA persistence-implications for cervical disease progression and monitoring.*** J Med Virol 2004;73:65–70.

151. Molden T, Kraus I, Karlsen F, Skomedal H, Hagmar B. ***Human papillomavirus E6/E7 mRNA expression in women younger than 30 years of age.*** Gynecol Oncol 2006;100:95–100.

152. Schmitt M, Dalstein V, Waterboer T, Clavel C, Gissmann L, Pawlita M. ***Diagnosing cervical cancer and high-grade precursors by HPV16 transcription patterns.*** Cancer Res. 2010 Jan 1;70(1):249-56.

153. Schmitt M, Bravo IG, Snijders PJ, Gissmann L, Pawlita M, Waterboer T. ***Bead-based multiplex genotyping of human papillomaviruses.*** J Clin Microbiol 2006;44:504–12.

154. Clavel C, Cucherousset J, Lorenzato M, et al. ***Negative human papillomavirus testing in normal smears selects a population at low risk for developing high-grade cervical lesions.*** Br J Cancer 2004;90: 1803–8.

155. Clavel C, Masure M, Bory JP, et al. ***Human papillomavirus testing in primary screening for the detection of high-grade cervical lesions: a study of 7932 women.*** Br J Cancer 2001;84: 1616–23.

156. Darragh, T. M. et al. ***The Lower Anogenital Squamous Terminology Standardization project for HPV-associated lesions: background and consensus recommendations from the College of American Pathologists and the American Society for Colposcopy and Cervical Pathology.*** Int. J. Gynecol. Pathol. 32, 76–115 (2013).

157. Alloul, N. & Sherman, L. (1999). ***The E2 protein of human papillomavirus type 16 is translated from a variety of differentially spliced polycistronic mRNAs.*** Journal of General Virology 80, 29±37.

158. Compton J. ***Nucleic acid sequence-based amplification.*** Nature 1991;350:91–2.

159. Smits HL, van Gemen B, Schukkink R, et al. ***Application of the NASBA nucleic acid amplification method for the detection of human papillomavirus type 16 E6-7 transcripts.*** J Virol Methods 1995;54:75–81.

160. Kraus I, Molden T, Erno LE, Skomedal H, Karlsen F, Hagmar B. ***Humanpapillomavirus oncogenic expression in the dysplastic portio; an investigation of biopsies from 190 cervical cones.*** Br J Cancer 2004; 90:1407–13.

161. Gray E, Pett MR, Ward D, Winder DM, Stanley MA, Roberts I, Scarpini CG, Coleman N. ***In vitro progression of human papillomavirus 16 episome-associated cervical neoplasia displays fundamental similarities to integrant-associated carcinogenesis.*** Cancer Res. 2010 May 15;70(10):4081-91.

162. Stanley MA, Browne HM, Appleby M, Minson AC. ***Properties of a non-tumorigenic human cervical keratinocyte cell line. International Journal of Cancer.*** 1989; 43:672–6.

163. Jeon S, Allen-Hoffmann BL, Lambert PF. ***Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells.*** J Virol. 1995; 69(5):2989–97.

164. C C Baker, W C Phelps, V Lindgren, M J Braun, M A Gonda, P M Howley. ***Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines.*** J Virol. 1987 Apr; 61(4): 962–971.

165. Bo Xu, Sasithorn Chotewutmontri, Stephan Wolf, Ursula Klos, Martina Schmitz, Matthias Dürst, Elisabeth Schwarz. ***Multiplex Identification of Human Papillomavirus 16 DNA Integration Sites in Cervical Carcinomas.*** PLoS One. 2013; 8(6): e66693. Published online 2013 Jun 18. doi: 10.1371/journal.pone.0066693

166. Tang S, Tao M, McCoy JP Jr. Zheng ZM. ***The E7 oncoprotein is translated from spliced E6*I transcripts in high-risk human papillomavirus type 16- or type 18-positive cervical cancer cell lines via translation reinitiation.*** J Virol. 2006; 80(9):4249–63.

167. Sotlar K, Diemer D, Dethleffs A, Hack Y, Stubner A, Vollmer N, Menton S, Menton M, Dietz K, Wallwiener D, Kandolf R, Bultmann B. ***Detection and typing of human papillomavirus by e6 nested multiplex PCR.*** J Clin Microbiol. 2004; 42(7):3176–84.

168. Pim D, Tomaic V, Banks L. ***The human papillomavirus (HPV) E6* proteins from high-risk, mucosal HPVs can direct degradation of cellular proteins in the absence of full-length E6 protein.*** J Virol. 2009; 83(19):9863–74.

169. Filippova M, Johnson MM, Bautista M, Filippov V, Fodor N, Tungteakkhun SS, Williams K, Duerksen-Hughes PJ. ***The large and small isoforms of human papillomavirus type 16 E6 bind to and differentially affect procaspase 8 stability and activity.*** J Virol. 2007; 81(8):4116–29.

170. Zheng ZM. ***Viral oncogenes, noncoding RNAs, and RNA splicing in human tumor viruses.*** Int J Biol Sci. 2010; 6(7):730–55.

171. Davy CE, Jackson DJ, Wang Q, Raj K, Masterson PJ, Fenner NF, Southern S, Cuthill S, Millar JB, Doorbar J. ***Identification of a G(2) arrest domain in the E1 wedge E4 protein of human papillomavirus type 16.*** J Virol. 2002; 76(19):9806–18.

172. Doorbar, J., S. Ely, J. Sterling, C. McLean, and L. Crawford. 1991. ***Specific interaction between HPV16 16E1^E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network.*** Nature 352:824–827.

173. Roberts, S., I. Ashmole, G. D. Johnson, J. W. Kreider, and P. H. Gallimore. 1993. ***Cutaneous and mucosal human papillomavirus E4 proteins form intermediate filament-like structures in epithelial cells.*** Virology 197:176–187.

174. Doorbar, J., R. Elston, S. Napthine, K. Raj, E. Medcalf, D. Jackson, N. Coleman, H. Griffin, P. Masterson, S. Stacey, Y. Mengitsu, and J. Dunlop. 2000. ***The E1∧E4 protein of human papillomavirus type 16 associates with a putative RNA helicase through sequences in its C terminus.*** J. Virol. 74: 10081–10095.

175. Filippova M, Evans W, Aragon R, Filippov V, Williams VM, Hong L, Reeves ME, Duerksen-Hughes P. ***The small splice variant of HPV16 E6, E6, reduces tumor formation in cervical carcinoma xenografts.*** Virology. 2014 Feb;450-451:153-164.

176. Stubenrauch F, Hummel M, Iftner T, Laimins LA. ***The E8E2C protein, a negative regulator of viral transcription and replication, is required for extrachromosomal maintenance of human papillomavirus type 31 in keratinocytes.*** J Virol. 2000; 74(3):1178–86.

177. Stubenrauch F, Zobel T, Iftner T. ***The E8 domain confers a novel long-distance transcriptional repression activity on the E8E2C protein of high-risk human papillomavirus type 31.*** J Virol. 2001; 75(9):4139–49.

178. Kisselev AF1, Goldberg AL. ***Proteasome inhibitors: from research tools to drug candidates.*** Chem Biol. 2001 Aug;8(8):739-58.

179. K.L. Rock, C. Gramm, L. Rothstein, K. Clark, R. Stein, L. Dick, D. Hwang, A.L. Goldberg. ***Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class 1 molecules.*** Cell 78 (1994) 761-771.

180. A. Craiu, M. Gaczynska, T. Akopian, C.F. Gramm, G. Fenteany, A.L. Goldberg, K.L. Rock. ***Lactacystin and clasto-lactacystin L-lactone modify multiple proteasome L-subunits and inhibit intracellular protein degradation and major histocompatibility complex class I antigen presentation.*** J. Biol. Chem. 272 (1997) 13437-13445.

181. Nunes AT, Annunziata CM. ***Proteasome inhibitors: structure and function.*** Semin Oncol. 2017 Dec;44(6):377-380.

182. K.K. Wang, P.W. Yuen. ***Calpain inhibition : an overview of its therapeutic potential.*** Trends Pharmacol. Sci. 15 (1994) 412-419.

183. V.J. Palombella, O.J. Rando, A.L. Goldberg, T. Maniatis. ***The ubiquitin- proteasome pathway is required for processing the NF-kappa- B1 precursor protein and the activation of NF-kappa-B.*** Cell 78 (1994) 773-785.

184. Gorman, C. M., Moffat, L. F. & Howard, B. H. (1982). ***Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells.*** Molecular and Cellular Biology 2, 1044±1051.

185. Cornel Mülhardt, E.W. Beese M.D., in Molecular Biology and Genomics, 2007 ***Reporter Genes for Quanititative Evidence: Chloramphenicol Acetyltransferase.***

https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/chloramphenicol-acetyltransferase

186. Schenker A, Straub E, Iftner T, Stubenrauch F. ***Cell-type-dependent activities of regulatory regions and E2 proteins derived from carcinogenic and non-carcinogenic human alphapapillomaviruses.*** J Gen Virol. 2013 Jun;94(Pt 6):1343-50.

187. Naoko Kajitani and Stefan Schwartz. ***RNA Binding Proteins that Control Human Papillomavirus Gene Expression.*** Biomolecules. 2015 Jun; 5(2): 758–774.

188. Bell, I.; Martin, A.; Roberts, S. ***The E1^E4 protein of human papillomavirus interacts with the serine-arginine-specific protein kinase SRPK1.*** J. Virol. 2007, 81, 5437–5448.

189. Bodaghi, S.; Jia, R.; Zheng, Z.M. ***Human papillomavirus type 16 E2 and E6 are RNA-binding proteins and inhibit in vitro splicing of pre-mRNAs with suboptimal splice sites.*** Virology 2009, 386, 32–43.

190. Johansson, C.; Somberg, M.; Li, X.; Backström Winquist, E.; Fay, J.; Ryan, F.; Pim, D.; Banks, L.; Schwartz, S. ***HPV-16 E2 contributes to induction of HPV-16 late gene expression by inhibiting early polyadenylation.*** EMBO J. 2012, 13, 3212–3227.

191. Cumming, S.A.; Repellin, C.E.; McPhilips, M.; Redford, J.C.; Clements, J.B.; Graham, S.V. ***The human papillomavirus type 31 untranslated region contains a complex bipartite negative regulatory element.*** J. Virol. 2002, 76, 5993–6003.

192. Polycarpou-Schwarz M, Groß M, Mestdagh P, Schott J, Grund SE, Hildenbrand C, Rom J, Aulmann S, Sinn HP, Vandesompele J, Diederichs S. ***The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation.*** Oncogene. 2018 Aug;37(34):4750-4768.

193. Mirian Galliote Morale, Walason da Silva Abjaude, Aline Montenegro Silva, Luisa Lina Villa, Enrique Boccardo. ***HPV-transformed cells exhibit altered HMGB1-TLR4/MyD88-SARM1 signaling axis.*** Sci Rep. 2018; 8: 3476.

194. Anne Gershenson. ***Deciphering Protein Stability in Cells.*** J Mol Biol. 2014 Jan 9; 426(1): 4–6.

195. Marc C. Deller, Leopold Kong, Bernhard Rupp. ***Protein stability: a crystallographer's perspective.*** Acta Crystallogr F Struct Biol Commun. 2016 Feb 1; 72(Pt 2): 72–95.

196. Mittal S., Saluja D. (2015) ***Protein Post-translational Modifications: Role in Protein Structure, Function and Stability.*** In: Singh L.R., Dar T.A., Ahmad P. (eds) Proteostasis and Chaperone Surveillance. Springer, New Delhi.

197. Naoko Kajitani, Stefan Schwartz. ***RNA Binding Proteins that Control Human Papillomavirus Gene Expression.*** Biomolecules. 2015 Jun; 5(2): 758–774.

198. Monika Bergvall, David Gagnon, Steve Titolo, Michaël Lehoux, Claudia M. D'Abramo, Thomas Melendy, Jacques Archambault. ***Requirement for the E1 Helicase C-Terminal Domain in Papillomavirus DNA Replication In Vivo.*** J Virol. 2016 Mar 15; 90(6): 3198–3211.

199. Whelan F, Stead JA, Shkumatov AV, Svergun DI, Sanders CM, Antson AA. 2012. ***A flexible brace maintains the assembly of a hexameric replicative helicase during DNA unwinding.*** Nucleic Acids Res 40:2271–2283. http://dx.doi.org/10.1093/nar/gkr906.

200. Schuck S, Stenlund A. 2015. ***A conserved regulatory module at the C terminus of the papillomavirus E1 helicase domain controls E1 helicase assembly.*** J Virol 89:1129–1142. http://dx.doi.org/10.1128/JVI.01903-14.

201. Haller F, Bieg M et al. ***Enhancer hijacking activates oncogenic transcription factor NR4A3 in acinic cell carcinomas of the salivary glands.*** Nat Commun. 2019 Jan 21;10(1):368.

202. Alison A. McBride. ***The Papillomavirus E2 Proteins.*** Virology. 2013 Oct; 445(0): 57–79.

203. Ziad Sahab, Sawali R. Sudarshan, Xuefeng Liu, YiYu Zhang, Alexander Kirilyuk, Christopher M. Kamonjoh, Vera Simic, Yuhai Dai, Stephen W. Byers, John Doorbar, Frank A. Suprynowicz, Richard Schlegel. ***Quantitative Measurement of Human Papillomavirus Type 16 E5 Oncoprotein Levels in Epithelial Cell Lines by Mass Spectrometry.*** J Virol. 2012 Sep; 86(17): 9465–9473.

204. PaVE: Papillomavirus Episteme. https://pave.niaid.nih.gov

205. Pauling, L., Corey, R. B. & Branson, H. R. (1951). ***The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain.*** Proc. Natl Acad. Sci. USA, 37, 205–211.

206. Brunori, M. (2014). ***Variations on the theme: allosteric control in hemoglobin.*** FEBS J. 281, 633–643.

207. Monod, J., Wyman, J. & Changeux, J.-P. (1965). ***On the nature of allosteric transitions: a plausible model.*** J. Mol. Biol. 12, 88–118.

208. Ronda, L., Bruno, S. & Bettati, S. (2013). ***Tertiary and quaternary effects in the allosteric regulation of animal hemoglobins.*** Biochim. Biophys. Acta, 1834, 1860–1872**.**

209. Parsell DA, Sauer RT. ***The structural stability of a protein is an important determinant of its proteolytic susceptibility in Escherichia coli.*** J Biol Chem. 1989 May 5;264(13):7590-5.

210. Goldberg, A. L., and St. John, A. C. (1976) Annu. Rev. Biochem. 45,747-803

211. Hershko, A., and Ciechanover A. (1982) Annu. Rev. Biochem. 51, 335-364

212. Rechsteiner, M., Rogers, S., and Rote, K. (1987) Trends Biochem. Sci. 12,390-394

213. Pace, C. N., and Barrett, A. J. (1984) Biochem. J. 219, 411-417

214. Thomas M, Tomaić V, Pim D, Myers MP, Tommasino M, Banks L. ***Interactions between E6AP and E6 proteins from alpha and beta HPV types.*** Virology. 2013 Jan 20;435(2):357-62.

215. Chang SW, Liu WC, Liao KY, Tsao YP, Hsu PH, Chen SL. ***Phosphorylation of HPV-16 E2 at serine 243 enables binding to Brd4 and mitotic chromosomes.*** PLoS One. 2014 Oct 23;9(10):e110882.

216. Selvey LA, Dunn LA, Tindle RW, Park DS, Frazer IH. ***Human papillomavirus (HPV) type 18 E7 protein is a short-lived steroid-inducible phosphoprotein in HPV-transformed cell lines.*** J Gen Virol. 1994 Jul;75 (Pt 7):1647-53.

217. Pim D, Tomaic V, Banks L. ***The human papillomavirus (HPV) E6* proteins from high-risk, mucosal HPVs can direct degradation of cellular proteins in the absence of full-length E6 protein.*** J Virol. 2009 Oct;83(19):9863-74.

218. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J. ***The Perseus computational platform for comprehensive analysis of (prote)omics data.*** Nat Methods. 2016 Sep;13(9):731-40.

219. McArthur H, Walter G. ***Monoclonal antibodies specific for the carboxy terminus of simian virus 40 large T antigen.*** J Virol. 1984;145:483–491.

220. Wolfgang Meschede, Klaus Zumbach, Joris Braspenning, Martin Scheffner, Luis Benitez-Bribiesca, Jeff Luande, Lutz Gissmann, Michael Pawlita. ***Antibodies against Early Proteins of Human Papillomaviruses as Diagnostic Markers for Invasive Cervical Cancer.*** J Clin Microbiol. 1998 Feb; 36(2): 475–480.

221. Lee D, Norby K, Hayes M, Chiu YF, Sugden B, Lambert PF. ***Using Organotypic Epithelial Tissue Culture to Study the Human Papillomavirus Life Cycle.*** Curr Protoc Microbiol. 2016 May 6;41:14B.8.1-14B.8.19.

222. https://www.hpvcenter.se/human_reference_clones

223. Schmitt M, Pawlita M. ***The HPV transcriptome in HPV16 positive cell lines.*** Mol Cell Probes. 2011 Apr-Jun;25(2-3):108-13.

224.https://www.promega.com/-/media/files/resources/protocols/technical-manuals/0/dual-luciferase-reporter-assay-system-protocol.pdf – ***Dual-Luciferase® Reporter Assay System, Technical Manual, Promega***.

225. De-Castro Arce J, Göckel-Krzikalla E, Rösl F. ***Silencing of multi-copy HPV16 by viral self-methylation and chromatin occlusion: a model for epigenetic virus-host interaction.*** Hum Mol Genet. 2012 Apr 15;21(8):1693-705.

226. Hoppe-Seyler F1, Butz K, zur Hausen H. ***Repression of the human papillomavirus type 18 enhancer by the cellular transcription factor Oct-1.*** J Virol. 1991 Oct;65(10):5613-8.

227. de Wet JR, Wood KV, DeLuca M, Helinski DR, Subramani S. ***Firefly luciferase gene: structure and expression in mammalian cells***. Mol Cell Biol. 1987 Feb;7(2):725-37.

228. Sehr P, Zumbach K, Pawlita M. ***A generic capture ELISA for recombinant proteins fused to glutathione S-transferase: validation for HPV serology.*** J Immunol Methods. 2001 Jul 1;253(1-2):153-62.

229. Eunyoug Lee and Do Hee Lee. ***Emerging roles of protein disulfide isomerase in cancer.*** BMB Rep. 2017 Aug; 50(8): 401–410.

230. Ramos FS, Serino LT, Carvalho CM, Lima RS, Urban CA2, Cavalli IJ1, Ribeiro EM3. ***PDIA3 and PDIA6 gene expression as an aggressiveness marker in primary ductal breast cancer.*** Genet Mol Res. 2015 Jun 26;14(2):6960-7.

231. Brychtova V, Vojtesek B, Hrstka R. ***Anterior gradient 2: a novel player in tumor cell biology.*** Cancer Lett. 2011;304:1–7. doi: 10.1016/j.canlet.2010.12.023.

232. Xu S, Sankar S, Neamati N. ***Protein disulfide isomerase: a promising target for cancer therapy.*** Drug Discov Today. 2014 Mar;19(3):222-40.

233. Zhang Y, Li T, Zhang L, Shangguan F, Shi G, Wu X, Cui Y, Wang X, Wang X, Liu Y, Lu B, Wei T, Wang CC, Wang L. ***Targeting the functional interplay between endoplasmic reticulum oxidoreductin-1α and protein disulfide isomerase suppresses the progression of cervical cancer.*** EBioMedicine. 2019 Mar;41:408-419.

9. Supplementary Information

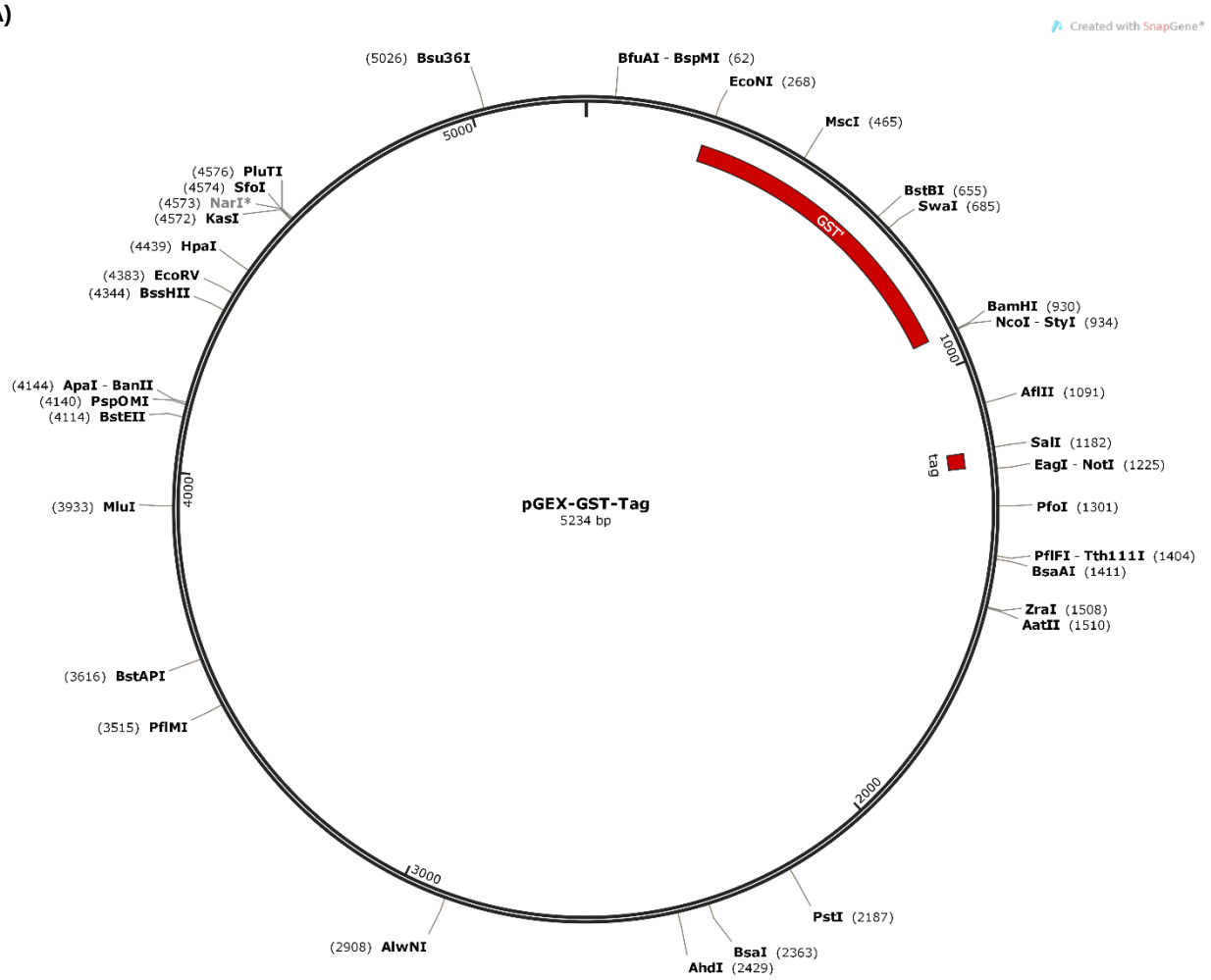
Table 9.1. **Details of plasmids containing spliced HPV16 transcripts[147].**

Name	Accession number	Cloned region from reference genome	Cloning vector
E6*I	HPV16R[a]	83-795	Bluescript M13-KS
E1^E4	HPV16R	83-795	Bluescript M13-KS
E1C	HPV16R	83-4210	Bluescript M13-KS
HPV16 fl	HPV16R	1-7904	Bluescript M13-KS

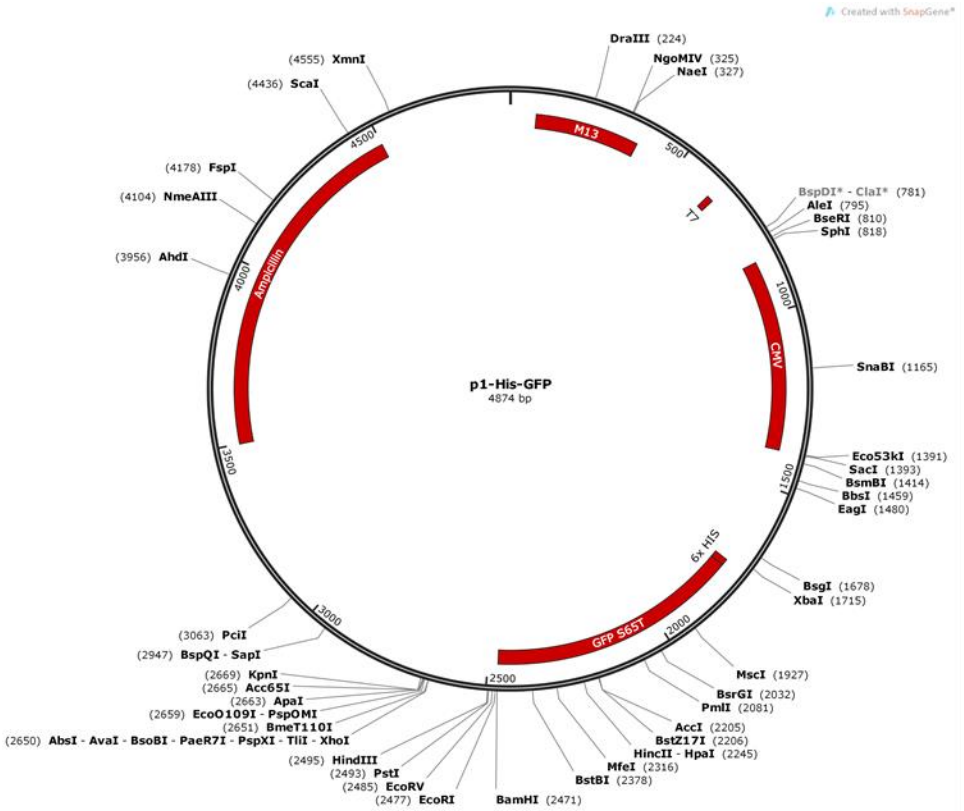
a - HPV16 sequence and base positions are numbered according to the 1996 sequence database (Los Alamos National Laboratory). http://hpv-web.lanl.gov/COMPENDIUM_PDF/95PDF/1/A1-9.pdf

9.2. Vector maps

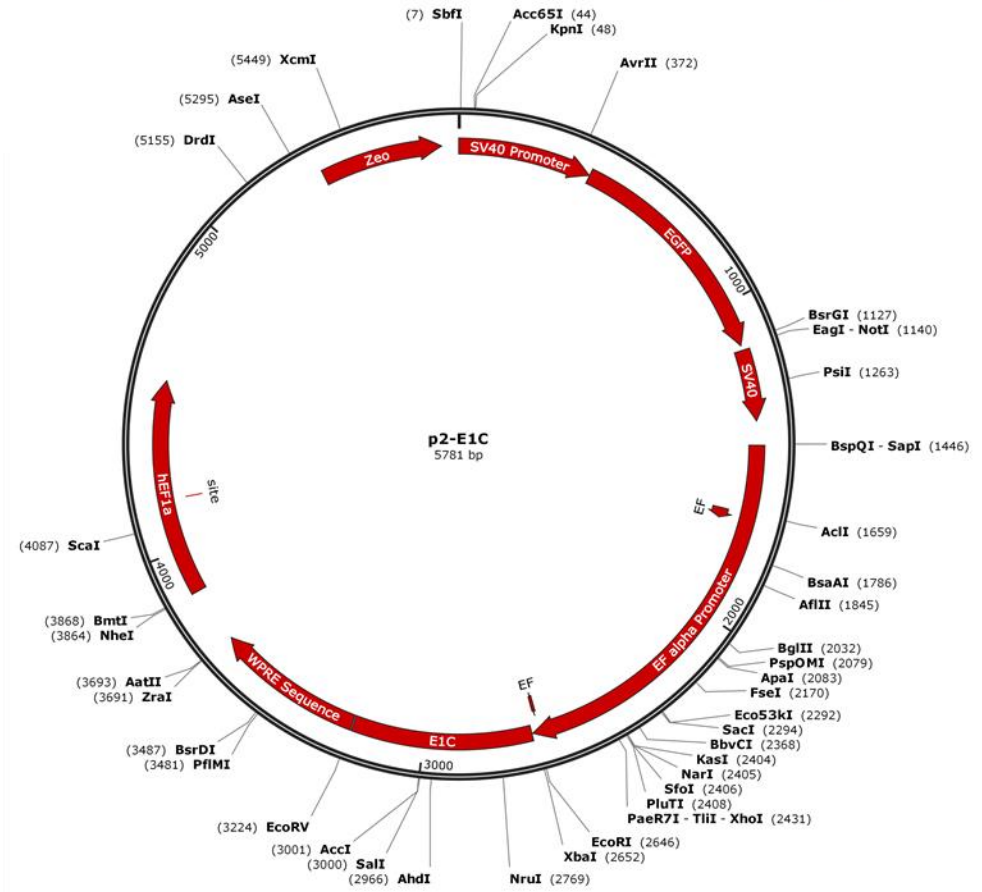
A)



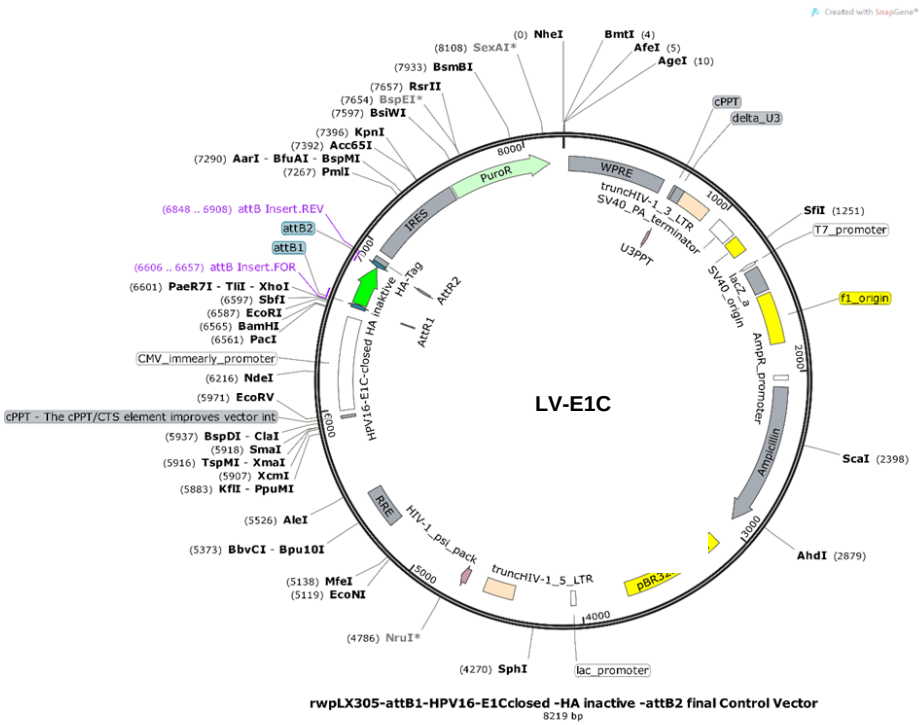
B)


Created with SnapGene®
p1-His-GFP
4874 bp
M13
T7
CMV
6x HIS
GFP S65T
Ampicillin
DraIII (224)
NgoMIV (325)
NaeI (327)
BspDI* - ClaI* (781)
AleI (795)
BseRI (810)
SphI (818)
SnaBI (1165)
Eco53kI (1391)
SacI (1393)
BsmBI (1414)
BbsI (1459)
EagI (1480)
BsgI (1678)
XbaI (1715)
MscI (1927)
BsrGI (2032)
PmlI (2081)
AccI (2205)
BstZ17I (2206)
HincII - HpaI (2245)
MfeI (2316)
BstBI (2378)
BamHI (2471)
(2477) EcoRI
(2485) EcoRV
(2493) PstI
(2495) HindIII
(2650) AbsI - AvaI - BsoBI - PaeR7I - PspXI - TliI - XhoI
(2651) BmeT110I
(2659) EcoO109I - PspOMI
(2663) ApaI
(2665) Acc65I
(2669) KpnI
(2947) BspQI - SapI
(3063) PciI
(3956) AhdI
(4104) NmeAIII
(4178) FspI
(4436) ScaI
(4555) XmnI

C)


Created with SnapGene®
p2-E1C
5781 bp
Zeo
SV40 Promoter
EGFP
SV40
EF
EF alpha Promoter
E1C
WPRE Sequence
hEF1a
site
(7) SbfI
Acc65I (44)
KpnI (48)
AvrII (372)
BsrGI (1127)
EagI - NotI (1140)
PsiI (1263)
BspQI - SapI (1446)
AclI (1659)
BsaAI (1786)
AflII (1845)
BglII (2032)
PspOMI (2079)
ApaI (2083)
FseI (2170)
Eco53kI (2292)
SacI (2294)
BbvCI (2368)
KasI (2404)
NarI (2405)
SfoI (2406)
PluTI (2408)
PaeR7I - TliI - XhoI (2431)
EcoRI (2646)
XbaI (2652)
NruI (2769)
(2966) AhdI
(3000) SalI
(3001) AccI
(3224) EcoRV
(3481) PflMI
(3487) BsrDI
(3691) ZraI
(3693) AatII
(3864) NheI
(3868) BmtI
(4087) ScaI
(5155) DrdI
(5295) AseI
(5449) XcmI

D)



E)

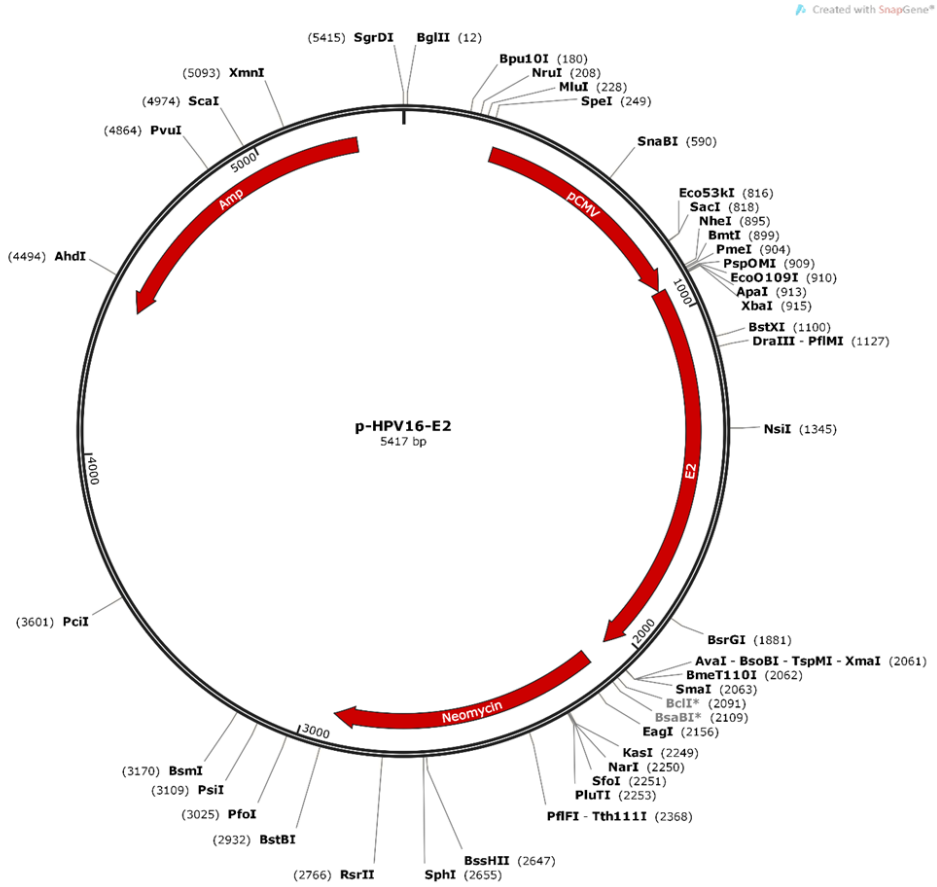


Figure 9.2. **Vector maps of A) pGEX-GST-Tag B) p1-His-GFP C) p2-E1C D) LV-E1C E) p-HPV16-E2**

9.3. DNA sequence information

a) HPV16 E1C

5'- ATGGCTGATCCTGCAGATTCTAGGTGGCCTTATTTACATAATAGATTGGTGGTGTTTACATTTCCTAATGAGTTTCCAT
TTGACGAAAACGGAAATCCAGTGTATGAGCTTAATGATAAGAACTGGAAATCCTTTTTCTCAAGGACGTGGTCCAGATTA
AGTTTGCACGAGGACGAGGACAAGGAAAACGATGGAGACTCTTTGCCAACGTTTAAATGTGTGTCAGGACAAAATACTA
ACACATTATGA-3'

b) GST-E1C-Tag

5'-ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATATCTTGAAGAAAAA
TATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTC
CTTATTATATTGATGGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAACATGTTGGGTGGT
TGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTGCATATAGTA
AAGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAAACA
TATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCT
GGATGCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATA
TAGCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGATCTGGTTCCGCGTGATGGC
TGATCCTGCAGATTCTAGGTGGCCTTATTTACATAATAGATTGGTGGTGTTTACATTTCCTAATGAGTTTCCATTTGACGAAAACG
GAAATCCAGTGTATGAGCTTAATGATAAGAACTGGAAATCCTTTTTCTCAAGGACGTGGTCCAGATTAAGTTTGCACGAGGACG
AGGACAAGGAAAACGATGGAGACTCTTTGCCAACGTTTAAATGTGTGTCAGGACAAAATACTAACACATTATGAAAACCTCCCAC
ACCTCCCCCTGAACCTGAAACATAA-3'

Color code: Red – GST sequence, Black – E1C sequence, Green – Tag sequence

c) His-GFP-E1C

5'-ACATCACCATCACCATCACATGAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGAT
GTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTA
CTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCAAGATACCCAGATCATATG
AAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACA
AGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAA
CATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTT
AACTTCAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCC
CTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTCC
TTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAGGATGGCTGATCCTGCAGATTCTAGGTGG
CCTTATTTACATAATAGATTGGTGGTGTTTACATTTCCTAATGAGTTTCCATTTGACGAAAACGGAAATCCAGTGTATGAGCTTAA
TGATAAGAACTGGAAATCCTTTTTCTCAAGGACGTGGTCCAGATTAAGTTTGCACGAGGACGAGGACAAGGAAAACGATGGAGA
CTCTTTGCCAACGTTTAAATGTGTGTCAGGACAAAATACTAACACATTATGA-3'

Color code: Purple – 'His' tag sequence, Brown – GFP sequence, Black – E1C sequence

d) HPV16-URR-FLuc

5'-TGCAGACCTAGATCAGTTTCCTTTAGGACGCAAATTTTTACTACAAGCAGGATTGAAGGCCAAACCAAAATTTACATTAGG
AAAACGAAAAGCTACACCCACCACCTCATCTACCTCTACAACTGCTAAACGCAAAAAACGTAAGCTGTAAGTATTGTATGTATGT
TGAATTAGTGTTGTTTGTTGTTTATATGTTTGTATGTGCTTGTATGTGCTTGTAAATATTAAGTTGTATGTGTGTTTGTATGTATGG
TATAATAAACACGTGTGTATGTGTTTTTAAATGCTTGTGTAACTATTGTGTCATGCAACATAAATAAACTTATTGTTTCAACACCTA
CTAATTGTGTTGTGGTTATTCATTGTATATAAACTATATTTGCTACATCCTGTTTTTGTTTTATATATACTATATTTTGTAGCGCCAG
CGGCCATTTTGTAGCTTCAACCGAATTCGGTTGCATGCTTTTTGGCACAAAATGTGTTTTTTTAAATAGTTCTATGTCAGCAACTA
TAGTTTAAACTTGTACGTTTCCTGCTTGCCATGCGTGCCAAATCCCTGTTTTCCTGACCTGCACTGCTTGCCAACCATTCCATTG
TTTTTTACACTGCACTATGTGCAACTACTGAATCACTATGTACATTGTGTCATATAAAATAAATCACTATGCGCCAACGCCTTACA
TACCGCTGTTAGGCACATATTTTTGGCTTGTTTTAACTAACCTAATTGCATATTTGGCATAAGGTTTAAACTTCTAAGGCCAACTA
AATGTCACCCTAGTTCATACATGAACTGTGTAAAGGTTAGTCATACATTGTTCATTTGTAAAACTGCACATGGGTGTGTGCAAAC
CGTTTTGGGTTACACATTTACAAGCAACTTATATAATAATACTAAACTACAATAATTCATGTATAAAACTAAGGGCGTAACCGAAA
TCGGTTGAACCGAAACCGGTTAGTATAAAAGCAGACATTTTATGCACCAAAAGAGAACTGCATGGAAGACGCCAAAAACATAAA
GAAAGGCCCGGCGCCATTCTATCCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGCCCTG
GTTCCTGGAACAATTGCTTTTGTGAGTATTTCTGTCTGATTTCTTTCGAGTTAACGAAATGTTCTTATGTTTCTTTAGACAGATGCA
CATATCGAGGTGAACATCACGTACGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCTGAAT
ACAAATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCCGGTGTTGGGCGCGTTATTTATCGGAGTTGCAG
TTGCGCCCGCGAACGACATTTATAATGAACGTAAGCACCCTCGCCATCAGACCAAAGGGAATGACGTATTTAATTTTTAAGGTG

AATTGCTCAACAGTATGAACATTTCGCAGCCTACCGTAGTGTTTGTTTCCAAAAAGGGGTTGCAAAAAATTTTGAACGTGCAAAA
AAAATTACCAATAATCCAGAAAATTATTATCATGGATTCTAAAACGGATTACCAGGGATTTCAGTCGATGTACACGTTCGTCACAT
CTCATCTACCTCCCGGTTTTAATGAATACGATTTTGTACCAGAGTCCTTTGATCGTGACAAAACAATTGCACTGATAATGAATTCC
TCTGGATCTACTGGGTTACCTAAGGGTGTGGCCCTTCCGCATAGAACTGCCTGCGTCAGATTCTCGCATGCCAGGTATGTCGTA
TAACAAGAGATTAAGTAATGTTGCTACACACATTGTAGAGATCCTATTTTTGGCAATCAAATCATTCCGGATACTGCGATTTTAAG
TGTTGTTCCATTCCATCACGGTTTTGGAATGTTTACTACACTCGGATATTTGATATGTGGATTTCGAGTCGTCTTAATGTATAGAT
TTGAAGAAGAGCTGTTTTTACGATCCCTTCAGGATTACAAAATTCAAAGTGCGTTGCTAGTACCAACCCTATTTTCATTCTTCGCC
AAAAGCACTCTGATTGACAAATACGATTTATCTAATTTACACGAAATTGCTTCTGGGGGCGCACCTCTTTCGAAAGAAGTCGGGG
AAGCGGTTGCAAAACGGTGAGTTAAGCGCATTGCTAGTATTTCAAGGCTCTAAAACGGCGCGTAGCTTCCATCTTCCAGGGATA
CGACAAGGATATGGGCTCACTGAGACTACATCAGCTATTCTGATTACACCCGAGGGGGATGATAAACCGGGCGCGGTCGGTAA
AGTTGTTCCATTTTTTGAAGCGAAGGTTGTGGATCTGGATACCGGGAAAACGCTGGGCGTTAATCAGAGAGGCGAATTATGTGT
CAGAGGACCTATGATTATGTCCGGTTATGTAAACAATCCGGAAGCGACCAACGCCTTGATTGACAAGGATGGATGGCTACATTC
TGGAGACATAGCTTACTGGGACGAAGACGAACACTTCTTCATAGTTGACCGCTTGAAGTCTTTAATTAAATACAAAGGATATCAG
GTAATGAAGATTTTTACATGCACACACGCTACAATACCTGTAGGTGGCCCCCGCTGAATTGGAATCGATATTGTTACAACACCCC
AACATCTTCGACGCGGGCGTGGCAGGTCTTCCCGACGATGACGCCGGTGAACTTCCCGCCGCCGTTGTTGTTTTGGAGCACG
GAAAGACGATGACGGAAAAAGAGATCGTGGATTACGTCGCCAGTAAATGAATTCGTTTTACGTTACTCGTACTACAATTCTTTTC
ATAGGTCAAGTAACAACCGCGAAAAAGTTGCGCGGAGGAGTTGTGTTTGTGGACGAAGTACCGAAAGGTCTTACCGGAAAACT
CGACGCAAGAAAAATCAGAGAGATCCTCATAAAGGCCAAGAAGGGCGGAAAGTCCAAATTGTAA-3’

Color code: Blue – HPV16-URR sequence, Black – FLuc sequence

e) HPV16-E2

5’-ATGGAGACTCTTTGCCAACGTTTAAATGTGTGTCAGGACAAAATACTAACACATTATGAAAATGATAGTACAGACCTACGTGA
CCATATAGACTATTGGAAACACATGCGCCTAGAATGTGCTATTTATTACAAGGCCAGAGAAATGGGATTTAAACATATTAACCAC
CAAGTGGTGCCAACACTGGCTGTATCAAAGAATAAAGCATTACAAGCAATTGAACTGCAACTAACGTTAGAAACAATATATAACT
CACAATATAGTAATGAAAAGTGGACATTACAAGACGTTAGCCTTGAAGTGTATTTAACTGCACCAACAGGATGTATAAAAAAACA
TGGATATACAGTGGAAGTGCAGTTTGATGGAGACATATGCAATACAATGCATTATACAAACTGGACACATATATATATTTGTGAA
GAAGCATCAGTAACTGTGGTAGAGGGTCAAGTTGACTATTATGGTTTATATTATGTTCATGAAGGAATACGAACATATTTTGTGC
AGTTTAAAGATGATGCAGAAAAATATAGTAAAAATAAAGTATGGGAAGTTCATGCGGGTGGTCAGGTAATATTATGTCCTACATC
TGTGTTTAGCAGCAACGAAGTATCCTCTCCTGAAATTATTAGGCAGCACTTGGCCAACCACCCCGCCGCGACCCATACCAAAGC
CGTCGCCTTGGGCACCGAAGAAACACAGACGACTATCCAGCGACCAAGATCAGAGCCAGACACCGGAAACCCCTGCCACACC
ACTAAGTTGTTGCACAGAGACTCAGTGGACAGTGCTCCAATCCTCACTGCATTTAACAGCTCACACAAAGGACGGATTAACTGT
AATAGTAACACTACACCCATAGTACATTTAAAAGGTGATGCTAATACTTTAAAATGTTTAAGATATAGATTTAAAAAGCATTGTACA
TTGTATACTGCAGTGTCGTCTACATGGCATTGGACAGGACATAATGTAAAACATAAAAGTGCAATTGTTACACTTACATATGATA
GTGAATGGCAACGTGACCAATTTTTGTCTCAAGTTAAAATACCAAAAACTATTACAGTGTCTACTGGATTTATGTCTATATGA-3’

f) HPV16-E1C-RNA mut

5’-GATTAATGGTATAACAGGATTAAGAAACCAATACAAAGGCTACATCCTCACTTAGATGAAAGCAAACGCAGAATAATGATTA
CTTTTTCGATACGTGAAACATATCCCATGGTAGTCCAAAGACTTGAAAGTCTATCACCTCTAGGGCCCTTTTCTGGATATAAACG
CCAAGTTGAATCCGTATTTGGAGGTACGATGGATCAGTCTGGATGAGACGTGCTTCATTTATATCGTAAGTAGGGTCGACCAAG
AACCCAAATC-3’

g) HPV16-E1C-Pro mut

5’-GATTAATGATAGTAACCTGCAGATTCTAGGTGGCCTTATTTACATAATAGATTGGTGGTGTTTACATTTCCTAATGAGTTT
CCATTTGACGAAAACGGAAATCCAGTGTATGAGCTTAATGATAAGAACTGGAAATCCTTTTTCTCAAGGACGTGGTCCAGATTAA
GTTTGCACGAGGACGAGGACAAGGAAAACGATGGAGACTCTTTGCCAACGTTTAAATGTGTGTCAGGACAAAATACTAACACAT
TATGACAAATC-3’

Color code: In red, green and blue are the three stop codons

9.4. Materials used for W12 cell culture

3T3 Cell medium

Ingredients	Quantity
DMEM with high glucose	500 mL
Fetal Bovine Serum	50 mL
Penicillin/streptomycin	5 mL

W12 Medium

Ingredients	Quantity
DMEM with high glucose	375 mL
F12 medium	125 mL
Fetal Bovine Serum	25 mL
100X Hydrocortisone	5 mL
100X Insulin	5 mL
100X Cholera Toxin	5 mL
100X Adenine	5 mL
Penicillin/streptomycin	5 mL
100X Epidermal Growth Factor	5 mL

Reagents used for W12 cell culture

Reagents	Composition
0.02% EDTA Solution	1L 1X PBS 0.2 g EDTA Autoclave the solution
Ham's F-12 nutrient mixture	
Earle's Salts	For 2L add: 128 g NaCl 8 g KCl 74 g NaHCO3 2.5 g NaH2PO4.H2O 4 g MgSO4.7H2O 0.002 g Fe(NO3)3.9H2O 0.1 g Phenol Red Bring to 2L with sterile water and filter-sterilize
HEPES Buffered Earle's Salts (HBES)	25 mL 1M HEPES buffer 100 mL Earle's Salts Bring to 2L with sterile water and filter-sterilize
50X Mitomycin C	2mL HBES 2mg mitomycin C 8 mL 3T3 cell medium Filter-sterilize and store at -20°C as 1 mL aliquots.
100X Hydrocortisone	5 mL cold 100% Ethanol

	25 mg Hydrocortisone This is 5 mg/mL solution. - Add 0.8 mL of 5 mg/mL hydrocortisone solution to 100 mL HBES containing 5% FBS. - Filter-sterilize and store at -20°C as 10 mL aliquots.
100X Cholera toxin	1.2mL autoclaved Bidest H2O 1mg Cholera toxin This is 10 µM solution. - Dilute 50 µL of 10 µM solution into 50 mL HBES containing 0.1% Bovine Serum Albumin. - Filter-sterilize and store at 4°C as 10 mL aliquots.
100X Insulin	Prepare immediately before use and never freeze. - Dissolve 12.5 mg insulin in 25 mL of 0.005M HCL. - Filter-sterilize with a syringe filter prewet with FBS.
100X Adenine	- Dissolve 121 mg adenine in 50 mL of 0.05M HCl by stirring one hour. - Filter-sterilize and store at -20°C in 10 mL aliquots.
100X Epidermal Growth Factor	- Dissolve 100 □g vial in 10 mL sterile water. - Add 90 mL HBES containing 0.1% bovine serum albumin. - Filter-sterilize and store at -20°C in 10 mL aliquots.

9.5. Complete transduction summary

In the below tables, details of all transductions including transduction no:, date on which transduction was performed, batch of lentiviral* particles preparation used for different transductions of different cell lines, cell lines used, lentiviral vector transduced, total no: of cells harvested after completion of puromycin selection, amount of cells used for RNA extraction, no: of RNA extractions performed, total amount of RNA extracted, no: of PCRs performed, copies/cell analyzed for different HPV16 transcripts (E6*I, E1^E4, E7 and ubC), Ratios of these transcript copies/cell to ubC (E6*I/ubC, E1^E4/ubC and E7/ubC) and finally, each of these ratios in LV-E1C transductions normalized to LV transductions

to analyze the changes in the expression of the corresponding transcripts (E6*I LV-E1C/LV, E1^E4 LV-E1C/LV and E7 LV-E1C/LV) are given.

***- LV stands for ‘Lentiviral’ in the tables.**

Table 1. **Details of transduction 1.**

Transduction	Date	LV batch	Cell lines	LV vector transduced	No: of cells harvested (N)	Amount of cells used for RNA extraction	RNA extraction	Total amount of RNA extracted (ng)	PCR	ubC copies per cell	E6*I copies per cell	E6*I/ubC	E6*I LV-E1C/LV	E1^E4 copies per cell	E1^E4/ubC	E1^E4 LV-E1C/LV	E7 copies per cell	E7/ubC	E7 LV-E1C/LV
1	21-Sep-17	1	W12-epi	LV	3.38E+06	5.00E+05	1	9630	1	202	695	3.4		482	2.4		970	4.8	
									2	168	632	3.8		417	2.5				
				LV-E1C	4.50E+06	5.00E+05	1	5924	1	133	480	3.6	1.0	208	1.6	0.7	737	5.5	1.2
									2	92	398	4.3	1.1	124	1.3	0.5			
			W12-int	LV	1.88E+06	5.00E+05	1	9630	1	160	214	1.3		2799	17.5				
									2	78	214	2.7		4567	58.6				
									3	117	234	2.0		3003	25.7				
									4	74	188	2.5		3186	43.1				
									5	141	306	2.2		9229	65.5				
									6	124	276	2.2		11503	92.8				
							2	12155	1	169	110	0.7		1748	10.3				
									2	164	101	0.6		1388	8.5				
									3	188	123	0.7		1032	5.5		43	0.2	
									4	190	132	0.7		1768	9.3				
				LV-E1C	3.00E+06	5.00E+05	1	28734	1	413	394	1.0	0.7	4313	10.4	0.6			
									2	235	347	1.5	0.5	6372	27.1	0.5			
									3	440	400	0.9	0.5	4472	10.2	0.4			
									4	286	351	1.2	0.5	6396	22.4	0.5			
									5	367	436	1.2	0.5	11816	32.2	0.5			
									6	314	418	1.3	0.6	14578	46.4	0.5			
							2	38834	1	121	222	1.8	2.8	3596	29.7	2.9			
									2	122	215	1.8	2.9	3370	27.6	3.3			
									3	134	229	1.7	2.6	1839	13.7	2.5	110	0.8	3.6
									4	134	215	1.6	2.3	3614	27.0	2.9			

Table 2. **Details of transduction 2.**

Transduction	Date	LV batch	Cell lines	LV vector transduced	No: of cells harvested (N)	Amount of cells used for RNA extraction	RNA extraction	Total amount of RNA extracted (ng)	PCR	ubC copies per cell	E6*I copies per cell	E6*I/ubC	E6*I LV-E1C/LV	E1^E4 copies per cell	E1^E4/ubC	E1^E4 LV-E1C/LV	E7 copies per cell	E7/ubC	E7 LV-E1C/LV
2	7-Dec-17	2	W12-epi	LV	9.75E+05	1.63E+05	1	3165	1	247	38	0.2		336	1.4		18	0.07	
									2	261	36	0.1		452	1.7				
							2	4575	1	392	30	0.1		790	2.0		21	0.05	
									2	311	30	0.1		830	2.7				
				LV-E1C	8.25E+05	1.38E+05	1	5045	1	387	43	0.1	0.7	382	1.0	0.7	18	0.05	0.6
									2	408	41	0.1	0.7	521	1.3	0.7			
							2	5265	1	504	27	0.1	0.7	490	1.0	0.5	15	0.03	0.6
									2	522	31	0.1	0.6	688	1.3	0.5			
			W12-int	LV	1.28E+06	2.13E+05	1	9468	1	651	111	0.2		902	1.4		46	0.07	
									2	629	106	0.2		1350	2.1				
							2	6189	1	814	86	0.1		1651	2.0		36	0.04	
									2	815	101	0.1		2356	2.9				
				LV-E1C	7.50E+05	1.25E+05	1	9018	1	675	123	0.2	1.1	890	1.3	1.0	49	0.07	1.0
									2	764	128	0.2	1.0	1587	2.1	1.0			
							2	10332	1	845	94	0.1	1.1	1692	2.0	1.0	38	0.04	1.0
									2	754	101	0.1	1.1	2031	2.7	0.9			

Tables 3 and 4. **Details of transductions 3 and 4.**

Transduction	Date	LV batch	Cell lines	LV vector transduced	No: of cells harvested (N)	Amount of cells used for RNA extraction	RNA extraction	Total amount of RNA extracted (ng)	PCR	ubC copies per cell	E6*I copies per cell	E6*I/ubC	E6*I LV-E1C/LV	E1^E4 copies per cell	E1^E4/ubC	E1^E4 LV-E1C/LV	E7 copies per cell	E7/ubC	E7 LV-E1C/LV	Comments
3	8-Feb-18	3	W12-epi	LV	5.00E+05	5.00E+04	1	469	1	23	12	0.5		85	3.7		1	0.04		applied 16ng RNA to PCR
							2	362	1	10	6	0.6		43	4.3		8	0.8		applied 16ng RNA to PCR
				LV-E1C	5.25E+06	5.00E+04	1	287	1	91	66	0.7	1.4	702	7.7	2.1	105	1.2	26.5	applied 16ng RNA to PCR
							2	297	1	95	79	0.8	1.4	754	7.9	1.8	97	1.0	1.3	applied 16ng RNA to PCR
			W12-int	LV	1.43E+07	5.00E+05	1	8398	1	229	137	0.6		1798	7.9		131	0.6		
									2	111	62	0.6		470	4.2		59	0.5		
				LV-E1C	6.75E+06	5.00E+05	1	8315	1	167	172	1.0	1.7	2321	13.9	1.8	153	0.9	1.6	
									2	75	79	1.1	1.9	626	8.3	2.0	78	1.0	2.0	
			CaSki	LV	1.00E+07	5.00E+05	1	7245	1	1058	188	0.2		2952	2.8		219	0.2		
				LV-E1C	5.50E+06	5.00E+05	1	11212	1	1019	151	0.1	0.8	2338	2.3	0.8	202	0.2	1.0	
			SiHa	LV	2.85E+07	5.00E+05	1	3714	1	246	78	0.3					160	0.7		
				LV-E1C	4.00E+06	5.00E+05	1	3827	1	338	116	0.3	1.1				146	0.4	0.7	

Transduction	Date	LV batch	Cell lines	LV vector transduced	No: of cells harvested (N)	Amount of cells used for RNA extraction	RNA extraction	Total amount of RNA extracted (ng)	PCR	ubC copies per cell	E6*I copies per cell	E6*I/ubC	E6*I LV-E1C/LV	E1^E4 copies per cell	E1^E4/ubC	E1^E4 LV-E1C/LV	E7 copies per cell	E7/ubC	E7 LV-E1C/LV	Comments
4	8-Feb-18	3	W12-epi	LV	7.50E+05	5.00E+04	1	734	1	41	22	0.5		205	5.0		16	0.4		applied 16ng RNA to PCR
							2	497	1	36	21	0.6		183	5.1		30	0.8		applied 16ng RNA to PCR
				LV-E1C	7.75E+06	5.00E+04	1	684	1	101	52	0.5	1.0	648	6.4	1.3				
							2	479	1	78	52	0.7	1.1	540	6.9	1.4	76	1.0	1.2	applied 16ng RNA to PCR
			W12-int	LV	1.35E+07	5.00E+05	1	8805	1	321	199	0.6		2815	8.8		206	0.6		
									2	166	104	0.6		925	5.6		89	0.5		
				LV-E1C	1.35E+07	5.00E+05	1	5051	1	161	129	0.8	1.3	1734	10.8	1.2	130	0.8	1.3	
									2	75	90	1.2	1.9	441	5.9	1.1	47	0.6	1.2	

Table 5. **Details of transduction 5.**

Transduction	Date	LV batch	Cell lines	LV vector transduced	No: of cells harvested (N)	Amount of cells used for RNA extraction	RNA extraction	Total amount of RNA extracted (ng)	PCR	ubC copies per cell	E6*I copies per cell	E6*I/ubC	E6*I LV-E1C/LV	E1^E4 copies per cell	E1^E4/ubC	E1^E4 LV-E1C/LV	E7 copies per cell	E7/ubC	E7 LV-E1C/LV	Comments
5	15-Nov-18	4	W12-epi	LV	1.05E+07	5.00E+05	1	2724	1	88	207	2.4		570	6.5		7	0.08		
									2											sample did not use!
									3	53	171	3.2		241	4.5					roughly 2 cp values higher in 4th PCR
									4	27	353	13.1		89	3.3					for all standards!
				LV-E1C	transduction did not work!															
			W12-int	LV	1.45E+07	5.00E+05	1	14433	1	1091	5535	5.1		23896	21.9		85	0.08		
									2	1111	27714	24.9		7586	6.8					
									3	534	5915	11.1		14864	27.8					roughly 2 cp values higher in 4th PCR
									4	1029	11502	11.2		9429	9.2					for all standards!
				LV-E1C	2.00E+07	5.00E+05	1	11318	1	984	4989	5.1	1.0	22615	23.0	1.0	64	0.07	0.8	
									2	1333	27831	20.9	0.8	9654	7.2	1.1				
									3	606	6186	10.2	0.9	14530	24.0	0.9				roughly 2 cp values higher in 4th PCR
									4	793	10381	13.1	1.2	4497	5.7	0.6				for all standards!
			CaSki	LV	1.28E+07	5.00E+05	1	7977	1	24186	512	0.02		2564	0.1		22	0.001		
									2	45514	2251	0.05		1398	0.03					
									3	23862	656	0.03		2038	0.1					roughly 2 cp values higher in 4th PCR
									4	127633	1568	0.01		1778	0.01					for all standards!
				LV-E1C	1.15E+07	5.00E+05	1	9240	1	23760	553	0.02	1.1	3013	0.1	1.2	28	0.001	1.3	
									2	42475	2328	0.05	1.1	1593	0.04	1.2				
									3	23539	678	0.03	1.0	2132	0.1	1.1				roughly 2 cp values higher in 4th PCR
									4	110694	1452	0.01	1.1	1145	0.01	0.7				for all standards!
			SiHa	LV	2.58E+07	5.00E+05	1	3983	1	4300	1161	0.27					15	0.003		
									2	7496	6430	0.86								
									3	4401	1456	0.33								roughly 2 cp values higher in 4th PCR
									4	17457	3409	0.20								for all standards!
				LV-E1C	1.48E+07	5.00E+05	1	2702	1	5674	1131	0.20	0.7				32	0.006	1.6	
									2	12292	8100	0.66	0.8							
									3	5270	1268	0.24	0.7							roughly 2 cp values higher in 4th PCR
									4	21924	3251	0.15	0.8							for all standards!

Table 6. **Details of transduction 6.**

Transduction	Date	LV batch	Cell lines	LV vector transduced	No: of cells harvested (N)	Amount of cells used for RNA extraction	RNA extraction	Total amount of RNA extracted (ng)	PCR	ubC copies per cell	E6*I copies per cell	E6*I/ubC	E6*I LV-E1C/LV	E1^E4 copies per cell	E1^E4/ubC	E1^E4 LV-E1C/LV	E7 copies per cell	E7/ubC	E7 LV-E1C/LV	Comments
6	15-Nov-18	4	W12-epi	LV	1.60E+07	5.00E+05	1	4344	1	536	646	1.2		1432	2.7		22	0.04		
									2	588	3020	5.1		560	1.0					
									3	300	577	1.9		934	3.1					roughly 2 cp values higher in 4th PCR for all standards!
									4	509	1245	2.4		510	1.0					
				LV-E1C	8.00E+06	5.00E+05	1	7311	1	496	1166	2.4	2.0	969	2.0	0.7	32	0.06	1.6	
									2	529	5413	10.2	2.0	219	0.4	0.4				
									3	326	1088	3.3	1.7	548	1.7	0.5				roughly 2 cp values higher in 4th PCR for all standards!
									4	351	2056	5.9	2.4	175	0.5	0.5				
			W12-int	LV	2.28E+07	5.00E+05	1	10134	1	1130	6663	5.9		27535	24.4		87	0.08		
									2	1438	42349	29.4		15193	10.6					
									3	574	5806	10.1		13832	24.1					roughly 2 cp values higher in 4th PCR for all standards!
									4	1080	11672	10.8		6861	6.4					
				LV-E1C	2.25E+07	5.00E+05	1	9696	1	1110	4934	4.4	0.8	21486	19.4	0.8	93	0.08	1.1	
									2	1262	31832	25.2	0.9	10736	8.5	0.8				
									3	560	4983	8.9	0.9	12162	21.7	0.9				roughly 2 cp values higher in 4th PCR for all standards!
									4	980	9233	9.4	0.9	5840	6.0	0.9				
			CaSki	LV	1.35E+07	5.00E+05	1	9333	1	22130	86	0.004		405	0.02		5	0.0002		
									2	41408	277	0.007		181	0.004					
									3	19926	90	0.005		271	0.01					roughly 2 cp values higher in 4th PCR for all standards!
									4	91057	208	0.002		136	0.001					
				LV-E1C	7.25E+06	5.00E+05	1	13305	1	21740	393	0.018	4.7	2156	0.1	5.4	10	0.0005	2.0	
									2	40959	1650	0.040	6.0	1187	0.03	6.6				
									3	20685	442	0.021	4.7	1380	0.1	4.9				roughly 2 cp values higher in 4th PCR for all standards!
									4	100396	1022	0.010	4.5	954	0.01	6.4				
			SiHa	LV	1.73E+07	5.00E+05	1	4319	1	5594	1456	0.260					40	0.007		
									2	9529	8447	0.886								
									3	3894	1109	0.285								roughly 2 cp values higher in 4th PCR for all standards!
									4	16423	3088	0.188								
				LV-E1C	2.23E+07	5.00E+05	1	3744	1	5674	1494	0.263	1.0				28	0.005	0.7	
									2	9426	7866	0.835	0.9							
									3	5342	1653	0.309	1.1							roughly 2 cp values higher in 4th PCR for all standards!
									4	23494	3933	0.167	0.9							

Table 7. **Details of transduction 7.**

Transduction	Date	LV batch	Cell lines	LV vector transduced	No: of cells harvested (N)	Amount of cells used for RNA extraction	RNA extraction	Total amount of RNA extracted (ng)	PCR	ubC copies per cell	E6*I copies per cell	E6*I/ubC	E6*I LV-E1C/LV	E1^E4 copies per cell	E1^E4/ubC	E1^E4 LV-E1C/LV	E7 copies per cell	E7/ubC	E7 LV-E1C/LV	Comments
7	11-Apr-19	5	W12-epi	LV	4.38E+06	5.00E+05	1	15980	1	177	259	1.5		200	1.1					
									2	998	381	0.4		874	0.9		29	0.03		
									3	832	314	0.4		578	0.7		30	0.04		
							2	10800	1	609	290	0.5		554	0.9					no Cp value detected!
									2	857	309	0.4		678	0.8					
				LV-E1C	3.50E+06	5.00E+05	1	13364	1	311	408	1.3	0.9	282	0.9	0.8				
									2	959	469	0.5	1.3	775	0.8	0.9	27	0.03	1.0	
									3	882	342	0.4	1.0	480	0.5	0.8	29	0.03	0.9	
							2	16322	1	949	347	0.4	0.8	550	0.6	0.6	37	0.04		
									2	905	397	0.4	1.2	663	0.7	0.9				
			W12-int	LV	2.13E+06	5.00E+05	1	16818	1	1325	2623	2.0		4125	3.1					
									2	3531	2121	0.6		9483	2.7		129	0.04		
									3	3531	1609	0.5		4254	1.2		136	0.04		
							2	14133	1	3689	1439	0.4		6728	1.8		163	0.04		
									2		1564			9960						no Cp value detected!
				LV-E1C	2.88E+06	5.00E+05	1	22115	1	1678	2644	1.6	0.8	4862	2.9	0.9				
									2	3715	1780	0.5	0.8	7135	1.9	0.7	94	0.03	0.7	
									3		1464			3915			96			no Cp value detected!
							2	20673	1	2386	1236	0.5	1.3	5275	2.2	1.2	117	0.05		
									2	2179	1305	0.6		7320	3.4					
			CaSki	LV	7.88E+06	5.00E+05	1	11829	1	21420	556	0.03		1077	0.1					
									2	20105	474	0.02		1820	0.1		41	0.002		
									3	22095	385	0.02		1126	0.1		44	0.002		
							2	10178	1	23417	399	0.02		1772	0.1		60	0.00		
									2	21076	449	0.02		2403	0.1					
				LV-E1C	5.00E+06	5.00E+05	1	16587	1	26001	492	0.02	0.7	1015	0.04	0.8				
									2	23079	367	0.02	0.7	1406	0.1	0.7	35	0.002	0.7	
									3	20251	292	0.01	0.8	881	0.04	0.9	44	0.002	1.1	
							2	17483	1	23673	312	0.01	0.8	1320	0.1	0.7	48	0.00		
									2	22582	374	0.02	0.8	1850	0.1	0.7				
			SiHa	LV	1.29E+07	5.00E+05	1	5868	1	4688	673	0.14								
									2	5246	764	0.15					62	0.01		
									3	4914	586	0.12					70	0.01		
							2	3377	1	6764	639	0.09					120	0.02		
									2	6913	744	0.11								
				LV-E1C	5.25E+06	5.00E+05	1	8100	1	4562	960	0.21	1.5							
									2	4932	749	0.15	1.0				67	0.01	1.1	
									3	3909	670	0.17	1.4				70	0.02	1.3	
							2	6645	1	4312	633	0.15	1.6				125	0.03		
									2	5480	760	0.14	1.3							

Table 8. **Details of transduction 8.**

Transduction	Date	LV batch	Cell lines	LV vector transduced	No: of cells harvested (N)	Amount of cells used for RNA extraction	RNA extraction	Total amount of RNA extracted (ng)	PCR	ubC copies per cell	E6*I copies per cell	E6*I/ubC	E6*I LV-E1C/LV	E1^E4 copies per cell	E1^E4/ubC	E1^E4 LV-E1C/LV	E7 copies per cell	E7/ubC	E7 LV-E1C/LV	Comments
8	11-Apr-19	5	W12-epi	LV	3.13E+06	5.00E+05	1	20786	1	349	470	1.3		522	1.5					
									2	845	411	0.5		911	1.1		26	0.03		
									3	709	304	0.4		509	0.7		29	0.04		
							2	22044	1	809	312	0.4		614	0.8		38	0.05		
									2	895	378	0.4		873	1.0					
				LV-E1C	4.75E+06	5.00E+05	1	17097	1	343	561	1.6	1.2	350	1.0	0.7				
									2	794	506	0.6	1.3	631	0.8	0.7	34	0.04	1.4	
									3	785	377	0.5	1.1	381	0.5	0.7	35	0.04	1.1	
							2	7625	1	851	372	0.4	1.1	478	0.6	0.7	46	0.05	1.2	
									2	845	407	0.5	1.1	531	0.6	0.6				
			W12-int	LV	3.75E+06	5.00E+05	1	14531	1	1443	3594	2.5		6835	4.7					
									2	3032	2487	0.8		11607	3.8		139	0.05		
									3		2092			6219			152			no Cp value detected!
							2	14265	1	2660	1727	0.6		8515	3.2		198	0.07		
									2		1707			10477						no Cp value detected!
				LV-E1C	2.00E+06	5.00E+05	1	17288	1	1250	747	0.6	0.2	1027	0.8	0.2				
									2	3190	570	0.2	0.2	1780	0.6	0.1	38	0.01	0.3	
									3	2830	446	0.2		1051	0.4		38	0.01		
							2	14195	1	2779	430	0.2	0.2	1400	0.5	0.2	49	0.02	0.2	
									2	2830	474	0.2		1751	0.6					
			CaSki	LV	8.00E+06	5.00E+05	1	10313	1	22879	463	0.02		882	0.0					
									2	19108	377	0.02		1354	0.1		37	0.002		
									3	18293	300	0.02		869	0.0		42	0.002		
							2	11484	1	20924	327	0.02		1336	0.1		47	0.002		
									2	22176	382	0.02		1796	0.1					
				LV-E1C	6.88E+06	5.00E+05	1	11916	1	27988	495	0.02	0.9	1056	0.04	1.0				
									2	25457	443	0.02	0.9	1702	0.1	0.9	38	0.001	0.8	
									3	27474	360	0.01	0.8	1110	0.04	0.9	42	0.002	0.7	
							2	12089	1	26207	346	0.01	0.8	1550	0.1	0.9	50	0.002	0.8	
									2	28489	412	0.01	0.8	2255	0.1	1.0				
			SiHa	LV	1.08E+07	5.00E+05	1	7721	1	5432	980	0.18								
									2	4843	779	0.16					68	0.01		
									3	4932	635	0.13					72	0.01		
							2	6248	1	4455	631	0.14					106	0.02		
									2	4704	708	0.15								
				LV-E1C	7.00E+06	5.00E+05	1	9386	1	5734	1270	0.22	1.2							
									2	5189	969	0.19	1.2				75	0.01	1.0	
									3	5022	807	0.16	1.2				87	0.02	1.2	
							2	7398	1	4826	694	0.14	1.0				118	0.02	1.0	
									2	4520	768	0.17	1.1							

Table 9. **Details of transduction protocol of transductions 1 – 4.**

														Experimenter			
Transduction	Date	LV batch	Cell lines	Passage number	Passage at seeding	Passage at harvest	No: of cells seeded	LV vector transduced	Start of puromycin selection	Splitting in between	Harvest date	Duration of puromycin selection	No: of days post transduction - harvest	LV-txn	Puromycin selection and cells harvest	RNA extraction	PCR
1	21-Sep-17	1	W12-epi	11	13	14	3.00E+05	LV	26-Sep-17	2-Oct-17	11-Oct-17	14 days	19 days	DKFZ Genomics & Proteomics core facility	Christy Susan Varghese	Christy Susan Varghese	Christy Susan Varghese
								LV-E1C	24-Sep-17	2-Oct-17	16-Oct-17	19 days	24 days				
			W12-int	7	9	10	1.50E+05	LV	26-Sep-17	2-Oct-17	7-Oct-17	10 days	15 days				
								LV-E1C	24-Sep-17	2-Oct-17	7-Oct-17						
2	7-Dec-17	2	W12-epi	11	12	13	8.30E+04	LV	13-Dec-17	no	20-Dec-17	6 days	12 days	DKFZ Genomics & Proteomics core facility	Christy Susan Varghese	Christy Susan Varghese	Christy Susan Varghese
								LV-E1C	13-Dec-17	no	20-Dec-17						
			W12-int	7	8	10	1.25E+05	LV	11-Dec-17	13-Dec-17	20-Dec-17	8 days	12 days				
								LV-E1C	11-Dec-17	13-Dec-17	20-Dec-17						
3	8-Feb-18	3	W12-epi	11	13	14	4.00E+05	LV	11-Feb-18	18-Feb-18	27-Feb-18	15 days	18 days	DKFZ Genomics & Proteomics core facility	Christy Susan Varghese	Yvonne Alt	Yvonne Alt
								LV-E1C	11-Feb-18	18-Feb-18	27-Feb-18						
			W12-int	7	9	10	4.00E+05	LV	11-Feb-18	18-Feb-18	24-Feb-18	12 days	15 days				
								LV-E1C	11-Feb-18	18-Feb-18	24-Feb-18						
			CaSki				4.00E+05	LV	9-Feb-18	3 splits	24-Feb-18	14 days	15 days				
								LV-E1C	9-Feb-18	2 splits	24-Feb-18						
			SiHa				4.00E+05	LV	9-Feb-18	1 split	27-Feb-18	17 days	18 days				
								LV-E1C	9-Feb-18		27-Feb-18						
4	8-Feb-18	3	W12-epi	11	13	14	4.00E+05	LV	11-Feb-18	18-Feb-18	27-Feb-18	15 days	18 days	DKFZ Genomics & Proteomics core facility	Christy Susan Varghese	Yvonne Alt	Yvonne Alt
								LV-E1C	11-Feb-18	18-Feb-18	27-Feb-18						
			W12-int	7	9	10	4.00E+05	LV	11-Feb-18	18-Feb-18	24-Feb-18	12 days	15 days				
								LV-E1C	11-Feb-18	18-Feb-18	24-Feb-18						
			CaSki				4.00E+05	LV	9-Feb-18	3 splits	24-Feb-18	14 days	15 days				
								LV-E1C	9-Feb-18	2 splits	24-Feb-18						
			SiHa				4.00E+05	LV	9-Feb-18	1 split	27-Feb-18	17 days	18 days				
								LV-E1C	9-Feb-18		27-Feb-18						

Table 10. **Details of transduction protocol of transductions 5 – 8.**

5	15-Nov-18	4	W12-epi	18	13	14	4.00E+05	LV	19-Nov-18	1 split	5-Dec-18	15 days	19 days	DKFZ Genomics & Proteomics core facility	Christy Susan Varghese	Maria Boulougouri	Maria Boulougouri
								LV-E1C	19-Nov-18	1 split	5-Dec-18						
			W12-int	8	9	10	4.00E+05	LV	19-Nov-18	1 split	5-Dec-18	15 days	19 days				
								LV-E1C	19-Nov-18	1 split	5-Dec-18						
			CaSki				3.00E+05	LV	16-Nov-18	2 splits	3-Dec-18	16 days	17 days				
								LV-E1C	16-Nov-18	2 splits	3-Dec-18						
			SiHa				4.00E+05	LV	16-Nov-18	1 split	29-Nov-18	12 days	13 days				
								LV-E1C	16-Nov-18	1 split	29-Nov-18						
6	15-Nov-18	4	W12-epi	18	13	14	4.00E+05	LV	19-Nov-18	1 split	5-Dec-18	15 days	19 days	DKFZ Genomics & Proteomics core facility	Christy Susan Varghese	Maria Boulougouri	Maria Boulougouri
								LV-E1C	19-Nov-18	1 split	5-Dec-18						
			W12-int	8	9	10	4.00E+05	LV	19-Nov-18	1 split	5-Dec-18	12 days	19 days				
								LV-E1C	19-Nov-18	1 split	5-Dec-18						
			CaSki				3.00E+05	LV	16-Nov-18	2 splits	3-Dec-18	16 days	17 days				
								LV-E1C	16-Nov-18	2 splits	3-Dec-18						
			SiHa				4.00E+05	LV	16-Nov-18	1 split	29-Nov-18	12 days	13 days				
								LV-E1C	16-Nov-18	1 split	29-Nov-18						
7	11-Apr-19	3	W12-epi	11	13	14	4.00E+05	LV	15-Apr-19	1 split	23-Apr-19	7 days	11 days	DKFZ Genomics & Proteomics core facility	Christy Susan Varghese	Maria Boulougouri	Maria Boulougouri
								LV-E1C	15-Apr-19	1 split	23-Apr-19						
			W12-int	7	9	10	4.00E+05	LV	15-Apr-19	1 split	23-Apr-19	7 days	11 days				
								LV-E1C	15-Apr-19	1 split	23-Apr-19						
			CaSki				4.00E+05	LV	15-Apr-19	2 splits	24-Apr-19	8 days	12 days				
								LV-E1C	15-Apr-19	2 splits	24-Apr-19						
			SiHa				4.00E+05	LV	15-Apr-19	1 split	25-Apr-19	9 days	13 days				
								LV-E1C	15-Apr-19	1 split	25-Feb-19						
8	11-Apr-19	3	W12-epi	11	13	14	4.00E+05	LV	15-Apr-19	1 split	23-Apr-19	7 days	11 days	DKFZ Genomics & Proteomics core facility	Christy Susan Varghese	Maria Boulougouri	Maria Boulougouri
								LV-E1C	15-Apr-19	1 split	23-Apr-19						
			W12-int	7	9	10	4.00E+05	LV	15-Apr-19	1 split	23-Apr-19	7 days	11 days				
								LV-E1C	15-Apr-19	1 split	23-Apr-19						
			CaSki				4.00E+05	LV	15-Apr-19	2 splits	24-Apr-19	8 days	12 days				
								LV-E1C	15-Apr-19	2 splits	24-Apr-19						
			SiHa				4.00E+05	LV	15-Apr-19	1 split	25-Apr-19	9 days	13 days				
								LV-E1C	15-Apr-19	1 split	25-Feb-19						

9.6. Table showing the firefly and renilla luciferase signals of all the transfections.

Firefly luciferase signal								
p-EV	p-HPV16-E2	p-HPV16-E1C	p-HPV16-E1C-RNA mut	p-HPV16-E1C-Pro mut	p-EV + p-HPV16-E2	p-HPV16-E2 + p-HPV16-E1C	p-HPV16-E2 + p-HPV16-E1C-RNA mut	p-HPV16-E2 + p-HPV16-E1C-Pro mut
4464	2933							
5255	2768							
6349	5434							
5617	4924							
5407	1854							
5904	1436							
5858	1317							
6555	1461							
6140	4616							
6335	4789							
2347	1277							
2498	1573							
3246	2201	4505			1255	1678		
2996	2328	4472			1444	1795		
5531	3829	7599			3235	3052		
4661	3711	7288			3236	2814		
2725	2248	4455	4052	5049	1199	1344	1680	2488
3251	2417	4208	3782	5591	955	1835	1728	2510
4037	3627	4325	4733	6027	1154	1519	1903	2639
3511	3717	5640	4618	6314	1406	1646	1457	2299

Renilla luciferase signal								
p-EV	p-HPV16-E2	p-HPV16-E1C	p-HPV16-E1C-RNA mut	p-HPV16-E1C-Pro mut	p-EV + p-HPV16-E2	p-HPV16-E2 + p-HPV16-E1C	p-HPV16-E2 + p-HPV16-E1C-RNA mut	p-HPV16-E2 + p-HPV16-E1C-Pro mut
150	75							
170	83							
239	217							
242	193							
1326	365							
1391	343							
1339	361							
6519	1474							
2201	423							
2133	426							
1328	843							
1418	970							
398	109	500			63	65		
376	133	540			43	66		
645	102	523			78	94		
619	96	481			68	84		
184	59	149	129	384	32	44	32	59
148	51	150	124	365	20	40	37	49
1552	282	1104	869	1111	209	226	210	234
1530	301	1250	887	1184	200	229	234	261