SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Phylogenetic and functional analysis of human papillomavirus (HPV) type 31 certain genomic region

by Annamária Ferenczi

Supervisor: György Veress, PhD

UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF PHARMACEUTICAL SCIENCES
DEBRECEN, 2016

Phylogenetic and functional analysis of human papillomavirus (HPV) type 31 certain genomic region

By **Annamária Ferenczi**
molecular biologist MSc (microbiologist)

Supervisor: György Veress, PhD

Doctoral School of Pharmaceutical Sciences, University of Debrecen

Head of the **Examination Committee**:	Árpád Tósaki, PharmD, PhD, DSc
Members of the Examination Committee:	Sándor Biró, PhD, DSc
	János Minárovits, MD, PhD, DSc

The Examination took place at Department of Pharmacology, Faculty of Medicine, University of Debrecen 11:00, 11.06.2015

Head of the **Defense Committee**:	Árpád Tósaki, PharmD, PhD, DSc
Reviewers:	Krisztián Bányai, PhD
	András Mádi, PhD
Members of the Defense Committee:	Sándor Biró, PhD, DSc
	János Minárovits, MD, PhD, DSc

The PhD Defense took place at the Lecture Hall of Bldg. A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen, 13:00, 25.01.2017.

INTRODUCTION

Human papillomavirus (HPV) is one of the most common sexually transmitted infection worldwide. The cervical cancer is the second most frequent cancerous disease of women, caused by persistent HPV infection, this is one of the leading health problems in Hungarian women, too. Close to 200 HPV types have been characterized, about one-third of which infect the anogenital tract. High-risk HPV types (including HPV 16, 18, 31, 33, and others) are mainly found in high-grade cervical lesions and invasive cervical cancers, while low-risk types (such as HPV 6 and 11) are detected in benign lesions, as condyloma acuminatum or genital papillomas.

HPVs belonging to the *Papillomaviridae* family are small, double stranded DNA viruses, 52-55 nm in size and have non-enveloped, icosahedral capsid. The capsid contains two virally encoded, structural proteins: major (L1) and minor (L2) capsid protein. The genome (8000bp) of HPVs consist of the early region (E1-E7) that encodes proteins necessary for viral replication; late region (L1-L2) required for virion assembly and non-coding region (long control region/LCR) which is essential for the replication and transcription of viral DNA.

HPVs have a specific tissue tropism, they infect the proliferating basal cells of stratified epithelium. The life cycle of human papillomavirus strictly follows the differentiation program of the host keratinocyte.HPV infections are not cytolytic, newly synthesized virions are released in parallel with degeneration of desquamating cells.

The E6 and E7 oncoproteins of HPV contribute to oncogenesis by associating with the tumor suppressor protein p53 and pRb, respectively. E6 interacts with and inactivates cellular proteins that have roles in controlling cell cycle, differentiation and apoptosis. E6 protein binds to p53 by formation of a trimolecular complex with E6-AP, resulting in the degradation of p53 by the ubiquitin-proteasome pathway. Similarly, E6 binds to the pro-apoptotic protein Bak causing its degradation. In addition, HPV E6 protein is able to increase the activity of the core promoter of hTERT (human telomerase reverse transcriptase) through association with Myc protein.

Retinoblastoma (Rb) protein is a major target of the E7 oncoprotein of high-risk HPVs. pRb is a member of the pocket protein family, along with p107 and p130. High-risk E7 can induce the release of active E2F transcription factors from pocket proteins, thereby inducing cell cycle entry. Moreover, high risk HPV E7 can induce the degradation of pRb through the proteasome pathway.

The LCR or URR (upstream regulatory region) is a genomic region, which does not have open reading frames, and contains numerous cis-responsive elements that regulate the gene expression and replication of papillomaviruses. The LCR of HPV 31 consists of the following functional domains: a 5' URR domain, an auxiliary enhancer (AE) domain, an epithelial cell specific keratinocyte enhancer (KE) domain, the minimal origin of replication, and the p97 promoter. The LCR contains binding sites for the viral E1 and E2 proteins, as well as for the cellular transcription factors AP1, YY1, NF1, Oct-1, Sp1, TEF-1 and CDP/Cut, some of them function as repressors, while others act as initiators in the course of viral replication.

HPV 16 and 18 are the most prevalent high-risk HPV genotypes worldwide, and are found in about 55 % and 15 % of invasive cervical cancer cases, respectively. In Europe, HPV 31 is the second most prevalent HPV type (after HPV 16) in women with normal cytology, while it is the third most prevalent type (after HPV 16 and 18) in cervical cancer cases. HPV 31 is found in about 4 % of invasive cervical cancers worldwide.

The highly conserved L1 region human papillomaviruses have been used for classification of HPVs. Sequence variants within an HPV type may differ from each other by less than 2 % in the L1 gene and by up to 5 % in the non-coding LCR. Intratypic nucleotide sequence variation has been extensively studied for HPV 16 and 18, which were reported to have different oncogenic potential. Clinical studies indicated differences in the abilities to persist and to cause premalignant lesions in the uterine cervix between HPV 31 variants belonging to different variant lineages (designated *A*, *B* and *C*). Different HPV 16 and HPV 18 intratypic LCR variants were found to have different transcriptional activities, and this may be one possible molecular mechanism that is responsible for the differences in the biological behaviour of intratypic HPV variants. In addition, functional differences found between E6/E7 variants may be also responsible for the differences in the oncogenic potential. HPV31 variants belonging to different variant lineages were found to have differences in persistence and in the ability to cause high grade cervical intraepithelial neoplasia.

For HPV types 16 and 18, intratypic sequence variation results in biological and clinical consequences. The functional significance of sequence variation among HPV 31 variants was studied less intensively. The purpose of the present study was to explore genetic variability and functional effects of the full length LCR and E6/E7 oncoproteins of HPV 31 variants isolated from Hungarian women with cervical premalignant lesions.

AIMS

- Investigation on the sequence polymorphism and phylogenetic relationships of HPV 31 LCR variants in clinical samples.

- Study on the transcriptional activity of LCR variants belonging to different intratypical lineages, try to localize the nucleotide changes which are responsible for differences of transcriptional activity of full length LCR variants by creation of deletion mutants and test their transcriptional activity.

- Investigation on the nucleotide and amino acid sequence polymorphism and phylogenetic relationships of HPV 31 E6 and E7 variants in clinical samples.

- Study on the effects of HPV 31 E6 and E7 intratypical variants on p53 and adenovirus E2 promoter activity.

- Analyses on the effects of E6 and E7 variants belonging to different intratypical groups on cellular p53 and pRb degradation *in vivo*.

MATERIALS AND METHODS

Clinical samples, HPV DNA detection

Forty-one HPV 31 isolates were amplified from cytological samples from women undergoing colposcopical examination for malignant or premalignant disease. The specimens were collected at the Department of Obstetrics and Gynaecology, University of Debrecen (Hungary) from 2005 to 2009. The mean age of the women was 34.9 (21.0–51.0) years. DNA was isolated from the samples with the High Pure Viral Nucleic Acid Kit. Amplification of the human β-globin gene was used to test for sample adequacy (internal control test). The presence of HPV DNA was detected with the MY09-MY11-HMB01 consensus L1 PCR method. Amplified products were typed by restriction fragment length polymorphism (RFLP).

Amplification of HPV 31 LCR, E6 and E7 for molecular variant analysis

Specific primers were designed to amplify the HPV 31 LCR - containing restriction sites for *KpnI* and *HindIII* restriction enzymes – and E6, E7 regions. Amplification reactions were performed with Gene Amp High Fidelity Enzime Mix according to the manufacturers. LCR amplimers were purified and sequenced.

Phylogenetic Analysis

Multiple sequence alignment of sequences and construction of phylogenetic trees were performed using Mega 5 software. Phylogenetic trees were constructed using the maximum likely (ML) method. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches.

Plasmid constructs

Representative HPV 31 LCR variants were selected for cloning, sequencing and functional analysis. Luciferase reporter constructs were made in the pGL2-Basic (promoterless) vector. Using *KpnI* and *HindIII* restriction enzyme recognition sites containing primer pairs for the LCR segments were amplified with PCR. The PCR amplification products of the HPV 31 LCR variants were digested by the restriction enzymes *KpnI* and *HindIII* and were ligated with T4 DNA Ligase upstream of the luciferase reporter gene of the pGL2-Basic vector.

HPV 31 LCR Del 1 and Del 2 deletion mutants were constructed from selected pGL2 reporter constructs containing full-length HPV 31 LCR variants by digestion with *KpnI* and *RsrII* (Del 1), or *KpnI* and *PmeI* restriction enzymes (Del 2). After digestion, the plasmids were gel purified with QIAquick Gel Extraction Kit and religated with T4 DNA Ligase. The Del 3

mutants were constructed by PCR amplification of selected HPV 31 LCR variant constructs with *HindIII* rescticton enzyme recognition sites containing primer pair. The PCR products were purified, digested by *HindIII* restriction enzyme and cloned into the pGL2 basic vector. All reporter constructs were confirmed by sequencing.

Representative E6 and E7 variants were selected and amplified with specific primers. The pcDNA™3.1/V5-His TOPO TA Expression Kit, was used to clone the E6 and E7 variants. All expression constructs were confirmed by sequencing. The p53-Luc reporter plasmid contains some copies of the p53, and the pAdE2Luc reporter construct contains adenovirus E2 promoter.

Cell Cultures, Transfections

The C-33A HPV negative human cervical cancer cell line was maintained in Dulbecco's modified Eagle's medium (DMEM). Cells (3 x10^5/well) were seeded on 6-well plates and allowed to grow to approximately 70% confluence. C-33A cells were transfected with the calcium-phosphate method using 2 mg of luciferase reporter plasmid and 0.05 mg pRL-TK Control Vector. The transfection mix was added to the cells and incubated for 24 h at 37°C, than the medium was changed to DMEM. Cells were harvested 48 h after transfection by the addition of lysis buffer. The Dual Luciferase Reporter (DLR) Assay system was used to measure the luciferase activity of cell extracts.

The MCF-7 HPV negative human breast cancer cell line was maintained in DMEM. MCF-7 cells (5×10^5/well) were seeded on 6-well plates and allowed to grow to approximately 70% confluence. The MCF-7 cells were transiently co-transfected with Lipofectamine 2000 reagent using 0.75 μg of expression vector (E6 or E7) along with 2 μg of reporter vector (p53 or pAdE2). 48 hours after transfection, cells were lysed. The Luciferase Assay System was used to measure the luciferase activity of the cell extracts. To standardize the protein concentration of the cell extracts, Bradford protein assay was performed. Each transfection experiment was carried out independently at least three times.

For stable transfection, 5×10^5/well MCF-7 cells were seeded on 6-well plates and allowed to grow to approximately 75–80% confluence. The cells were transfected with the E6 or E7 expression vectors as described above (without reporter vectors). At 48 h post transfection, DMEM containing 600 μg/μl Geneticin was added to the cells. This selection medium was changed on the cells every third day for altogether 3 weeks. The selected cell colonies were pooled and cultured in DMEM without Geneticin.

RNA isolation, RT-PCR

Total RNA was isolated from stable transfected MCF-7 cells using TRI reagent. To prepare cDNA, the High Capacity cDNA Reverse Transcription Kit was used following the manufacturers' instructions. The PCR reactions were performed with RedTaq DNA polymerase. The primer pairs used for amplifying HPV16 or HPV 31 E6 and E7 were the same as described above. GAPDH was used as an endogenous control.

Protein isolation, Western blot analysis

Cells stably expressing the HPV E6 or E7 oncoproteins were isolated by using RIPA lysis buffer containing protease and phosphatase inhibitors. Protein extracts were electrophoresed on 10% SDS-polyacrilamide gel and electrotransferred onto nitrocellulose membrane. The membrane was blocked using 5% milk in phosphate-buffered saline (pH 7.2) containing 0.05% Tween20 (PBST). The blots were incubated with primary antibodies overnight at 4 °C. The following primary antibodies were used: mouse monoclonal His-probe antibody (1:2000), mouse monoclonalanti-p53 (1:10,000), mouse monoclonal anti-Rb 1:5000), and rabbit polyclonal anti-actin (1:10,000). After washing in PBST, the blot was incubated with a HRP-conjugated goat anti-mouse antibody or HRP-conjugated goat anti-rabbit antibody at a dilution of 1:20,000 in 5% milk in PBST for 1 h at room temperature. The signals were developed with ChemiDoc MP Imaging System.

Statistics

Analysis of variance (ANOVA) test was used to determine if there were LCR variants with significantly different transcriptional activities in the study group and in the subgroups. Significance of difference between the transcriptional activities in pair wise comparisons was determined by Student’s t-test. Statistical significance was accepted at $p< 0.05$. Significance of differences between the transcriptional activities of different E6 and E7 variants was determined by Student's t-test. The results of the Western blot analysis were analysed using the same method. Significance was accepted at $p <0.05$.

RESULTS

HPV 31 LCR variants

Nucleotide changes in the long control region

Near full-length LCR sequences (from nt 7122 to 179) were amplified from 41 HPV 31 positive clinical samples. A total of 42 single nucleotide changes and 2 deletions were found in the sequenced LCR fragment of 41 clinical isolates. A 6bp deletion was found in a single sample, while a 10bp deletion (50-CTATTTTATA-30 at the positions 7299–7308) was found in altogether 12 samples, all belonging to lineage *B* of HPV 31 variants (see later). There were 35 single nucleotide changes occurring in multiple samples in this study, of which only 2 (G7326A and A7402G) were not reported by others. Of the seven single nucleotide changes found here in single samples, only 2 (A7517G and G7764A) were described in previous studies. Thus, 35 of the 42 nucleotide changes found in this study have been previously reported. Nucleotide changes not reported previously and/or found in single samples only were confirmed by repeated amplification and sequencing of a subset of clinical samples (samples HU 7156, 7404, 7527, 7548, 7762, 8128, and 8265). Several nucleotide changes found in the HPV 31 variants implicated putative binding sites for viral (E2) or cellular transcription factors.

Phylogenetic relationships of HPV 31 LCR variants

A phylogenetic tree was constructed based on the maximum likelihood (ML) algorithm representing the HPV 31 reference LCR sequence and the 41 clinical isolates. This tree confirms the presence of three lineages within the HPV 31 type, these lineages are designated A–C. The genetic distance among variants belonging to the same lineage was below 1%, while it was higher than 1% among variants belonging to different lineages. The maximal genetic distance (3.3%) was seen between the variants HU 8265 and HU 8128. Lineage *B* comprised 12 isolates (two different variants), all having a 10bp deletion (nt 7299–7308) and 15 or 17 single nucleotide changes compared to the reference clone. Lineage *C* was represented by altogether 27 isolates (10 different variants) characterized by 22–27 single nucleotide changes compared to the sequence of the reference isolate. Thus, lineage *C* of HPV 31 comprised the highest number of isolates and different variants and also seemed to have the highest intra-lineage variability in this Hungarian population.

Functional analysis of HPV 31 LCR variants

In order to see whether these nucleotide changes resulted in altered transcriptional activity of the HPV 31 LCR variants compared to the prototype selected LCR variants (along with the prototype LCR) were cloned into the promoterless luciferase reporter vector pGL2-Basic. The resultant constructs were confirmed by sequencing and then transiently transfected into the HPV negative cervical cancer cell line C33A. The results of luciferase tests showed that HPV 31 LCR variants belonging to different variant lineages had different transcriptional activities. It was interesting that the activity of the HPV 31 reference LCR construct (representing lineage *A*) was not higher than that of the empty reporter vector (pGL2). All reporter constructs containing the LCR of variants belonging to lineage *B* or *C* had significantly higher activities than the prototype HPV 31 LCR construct ($p < 0.001$, for both comparisons). On the other hand, LCR variants belonging to lineage *B* of HPV 31 had significantly higher transcriptional activities than variants belonging to lineage *C* ($p < 0.001$).

To try to localize the nucleotide alterations responsible for the differences found in transcriptional activities of the different HPV 31 LCR variants, deletion mutants were constructed from selected full-length LCR variants (representing different HPV 31 intratypic lineages) and tested in transient transfection assays. In contrast to the results obtained with the full-length variants, the transcriptional activities of different HPV31 Del1 constructs (with deletion from nucleotide 7122 to 7415) had similar transcriptional activities. This may indicate that at least some of the nucleotide changes that are responsible for the differences in the transcriptional activities of different HPV 31 LCR variants can be localized to the segment of LCR between nucleotides 7122 and 7415. Different HPV31 Del2 constructs (with deletion from nucleotide 7122 to 7792) also had similar transcriptional activities. In contrast, the HPV31 Del3construct representing lineage *C* had significantly lower transcriptional activity than the Del3 constructs representing lineage *A* or *B* ($p < 0.01$). This suggests that the nucleotide changes present in the minimal promoter region of lineage *C* variants compared to lineage *A* or *B* variants result in reduced transcriptional activity.

HPV 31 E6 and E7 oncogenes

Nucleotide/amino acid polymorphisms and phylogenetic relationships of E6 and E7 variants

Altogether 8 nucleotide changes were found in the E6 gene relative to the reference sequence, of which 4 resulted in amino acid changes. In the E7 gene, 5 nucleotide changes were

found, 3 of which resulted in changes in the amino acid sequence relative to the prototype. Each of these nucleotide and amino acid changes were described previously. In order to explore natural nucleotide and amino acid sequence variation in the E6/E7 region of HPV 31, we selected variants representing each of the 3 main variant lineages of HPV 31 on the basis of a previous phylogenetic analysis of HPV 31 LCR variation. Comparative phylogenetic analysis of the E6/E7 and the LCR region showed that nucleotide sequence variation was lower in the E6/E7 region, but either region could be used with high confidence to identify the variant lineage of an isolate.

Inhibition of p53 transcriptional activity by HPV 31 E6 variants

To see whether natural variation in the amino acid sequence of the HPV 31 E6/E7 region results in functional alterations, we initiated the functional analysis of the variant E6 and E7 proteins. The E6 and the E7 coding regions of selected HPV 31 variants were cloned into a eukaryotic expression vector in frame with a C-terminal histidine tag. For comparison, expression constructs containing prototype HPV 31 and HPV 16 E6 and E7 genes were also constructed. The in vivo assays were performed in MCF-7 cells as these cells contain functionally active p53 and pRb proteins. An important function of high-risk HPV E6 proteins is the ability to inhibit the transcriptional activity of the p53 protein. To explore the ability of the HPV 31 E6 variants to inhibit p53 transcriptional activity, MCF-7 cells were transiently co-transfected with expression vectors encoding the E6 variants along with a p53 reporter vector containing multiple binding sites for the p53 protein. For comparison, the HPV 31 and the HPV 16 prototype E6 were also included in the analysis. The variant representing lineage B (called E6V1) was analysed as two separate clones (E6V1a and E6V1b), although these had identical amino acid sequences. The results of luciferase assays showed that the prototype HPV 31 E6 protein (belonging to variant lineage *A*), along with a variant belonging to lineage *C*, was able to inhibit the transcriptional activity of p53. On the contrary, the HPV 31 E6 variant belonging to lineage B (E6V1) showed reduced ability to inhibit p53 transcriptional activity. These results indicate that there may be differences in the functional activities of HPV 31 E6 variants belonging to different variant lineages.

Effects of HPV 31 E7 variants on E2F activity

High risk HPV E7 proteins are known to bind to the cellular pRb and related pocket proteins to release transcriptionally active E2F factors resulting in induction of the cell cycle. This function of the HPV 31 E7 prototype and variants were tested in transient transfection experiments using a reporter construct containing the adenovirus E2 promoter (with binding

sites for E2F transcription factors). HPV 31 E7 variants belonging to different variant lineages were able to increase E2F transcriptional activity to the same extent as the prototype HPV 16 E7.

Effects of HPV 31 E6 and E7 variants on the levels of cellular tumour suppressor proteins

A well-known function of high-risk HPV E6 proteins is to induce the degradation of the cellular tumour suppressor protein p53. By using experimental mutagenesis of the HPV 16 E6 protein, it was shown that the ability of E6 to induce p53 degradation and the ability to inhibit the transcriptional activity of p53 may be separable. On the other hand, the high-risk HPV E7 proteins were shown to decrease the stability of the cellular pRb protein. To study the effects of HPV 31 E6 and E7 variants on the level of the p53 protein and the pRb protein, respectively, MCF-7 derived cell lines stably expressing the appropriate variants were constructed. The different HPV 31 E6 and E7 variants were expressed at similar levels both on the mRNA and the protein level. Next, we studied the level of endogenous p53 protein in the MCF-7 derived cell lines stably expressing the different HPV 31 E6 variants. We found that HPV 31 E6 prototype and variants were able to decrease the stability of the p53 protein to a certain extent. On the other hand, there seemed to be no major differences between the HPV 31 E6 variants in the ability to induce p53 degradation. Testing HPV 31 E7 variants for the degradation of endogenous pRb revealed that each of them decreased pRb level to an extent similar to the HPV 16 E7 prototype. There were no significant differences between the HPV 31 E7 variants in the ability to induce pRb degradation.

DISCUSSION

For HPV types 16 and 18, intratypic sequence variation results in biological and clinical consequences. The functional significance of sequence variation among HPV 31 variants was studied less intensively. Clinical studies indicated differences in the abilities to persist and to cause premalignant lesions in the uterine cervix between HPV 31 variants belonging to different variant lineages (designated *A*, *B* and *C*). The underlining mechanisms that are responsible for the clinical differences between the HPV 31 variants are still to be explored. One possible mechanism responsible for these differences may be that HPV 31 LCR variants belonging to different variant lineages display different transcriptional activities. As the LCR regulates transcription of the E6/E7 oncogenes, variation in LCR transcriptional activity may result in altered expression of the oncogenes. Another possible mechanism that could be responsible for the functional differences between HPV 31 variants is that amino acid changes in the E6/E7 proteins may result in some changes in the functional activities of these proteins. HPV 16 E6 natural variants were shown to have variable activities in p53 binding and degradation, inhibition of p53 transactivation, inducing immortalisation, inhibition of keratinocyte differentiation, and modulation of apoptosis.

In the first part of my work, I studied the intratypic HPV 31 nucleotide sequence variation in Hungarian patients. In this study population, HPV 31 was the second most prevalent HPV type (after HPV 16). Although several studies have already explored HPV 31 intratypic sequence variation this is the first report on the functional significance of natural sequence variation in a genomic region of HPV 31. While most of the previous studies reported only partial LCR sequences, one study analysed complete genomic sequences of selected HPV 31 isolates. The current study explored sequence variation in near the whole HPV 31 LCR. The sequences of the HPV 31 LCR variants reported here are in accordance with those described previously.

Accordingly, the phylogenetic tree constructed in this study is very similar to that constructed by using complete genomic sequences of HPV 31 variants, and a tree containing isolates from both studies showed that the isolates from different parts of the world clustered together to form the different variant lineages of HPV 31. Within HPV types 16 and 18, intratypic variants can be classified on the basis of geographical and ethnic origin. For HPV 31 variants, a similar geographical distribution had been suggested previously, but later it was not confirmed. The results of the present study are also against a specific geographical clustering of HPV 31 variants, as the variants isolated in this Hungarian population represented all the

three variant lineages of HPV 31. In contrast, HPV 16 variants previously isolated from Hungary were predominantly European variants, in addition to a few Asian-American variants.

The LCR of HPV 31 consists of the following functional domains: a 5' URR domain, an auxiliary enhancer (AE) domain, an epithelial cell-specific keratinocyte enhancer (KE) domain, the minimal origin of replication, and the p97 promoter. The LCR contains binding sites for the viral E1 and E2 proteins, as well as for the cellular transcription factors AP1, YY1, NF1, Oct-1, Sp1 and TEF-1. In the HPV 31 variants, nucleotide changes were found in each functional domain of the LCR and encompassed binding sites for viral (E2) and also for some cellular transcription factors. Therefore, it seemed logical to explore the transcriptional activities of different HPV 31 variants in transient transfection assays and luciferase reporter tests. The functional assays conducted in this study with full-length LCR constructs suggest that there are differences in the transcriptional activities between HPV 31 LCR variants belonging to different variant lineages. The full-length reference HPV 31 LCR variant belonging to lineage *A* had very low transcriptional activity in C33A cells compared to the empty reporter vector. The reason for this may be that transcriptional silencer elements counteract the effects of the enhancer regions on the activity of HPV 31 LCR in the reporter construct used in this study. Indeed, a conserved silencing element (overlapping the binding site for the viral E1 protein) is present in genital HPVs between the epithelial-specific enhancer and the E6 promoter, and thus contiguous enhancer-promoter constructs (containing this silencer) have low transcriptional activity in several HPVs. The HPV 31 LCR variants belonging to lineage *B* or *C* all had significantly higher transcriptional activities than the prototype LCR variant, although these also contained the conserved silencing element. This suggests that some of the single nucleotide changes (relative to the sequence of the reference isolate) found in these isolates may significantly alter the binding and/or activity of transcription factors.

On the basis of the experiments conducted here with deletion constructs of HPV 31 variants (representing different intratypic lineages), it is hard to delineate precisely the nucleotide changes that are responsible for the differences in the transcriptional activities of the different LCR variants. These differences might be caused at least partially by nucleotide changes found in the 5'URR domain (between nucleotides 7122 and 7415). On the other hand, nucleotide changes found in the minimal promoter region in lineage *C* variants (compared to lineage *A* or *B* variants) also seemed to modify the transcriptional activity of the LCR. Variants belonging to the same lineage of HPV 31 were found to have similar transcriptional activities in our study. In contrast, there seemed to be significant differences in the transcriptional

activities of variants belonging to different lineages, when these were tested as full-length LCR constructs. This suggests that there may be some correlation between the phylogenetic clustering and some functional characteristics of HPV 31 variants. In accordance, HPV 16 and 18 intratypic LCR variants belonging to different variant lineages also have different transcriptional activities. Asian-American and North-American variants of HPV 16 were found to have higher transcriptional activities than European variants. Similarly, the Asian-Amerindian variants of HPV 18 were reported to have higher transcriptional activities than the European variants. Non-European HPV 16 and 18 intratypic variants were reported to have enhanced oncogenic potential compared to European variants.

In the case of HPV 31, data on the persistence and oncogenic potential of different intratypic variants are still contradictory. The clinical evaluation of HPV 31 variants may be hampered by the relatively small number of HPV 31 positive patients. Nevertheless, the functional differences observed here by using the widely used C33A cervical cancer cell model system suggests that the different variants react differently to one or more transcriptional regulatory mechanisms involved in cervical carcinogenesis. The different transcriptional activities of HPV 31 LCR variants may result in altered oncogenic potential through altered expression level of E6/E7 oncogenes. On the other hand, amino acid changes in the E6/E7 oncogenes could also result in altered oncogenic potential of HPV 31 variants.

Therefore, we extended our study to investigate the phylogenetic and functional analysis of HPV 31 E6 and E7 natural protein variants. The nucleotide and amino acid changes found in the E6/E7 region in our samples were already described in previous publications. As expected, nucleotide variation was higher in the non-coding LCR region than in the protein-coding E6-E7 region. Comparing the phylogenetic trees constructed from LCR and E6/E7 sequences of HPV 31 variants, we can conclude that either region can be used to identify the variant lineage (*A, B* or *C*) of an isolate.

To start the functional analysis of the HPV 31 E7 variants, we tested their ability to release E2F transcription factors from complexes formed with pRb and related pocket proteins. In addition, we also tested the ability of the E7 variants to induce the in vivo degradation of pRB. Our experiments revealed no significant differences between the HPV 31E7 variants in either of these functional assays. However, there might be differences between the E7 variants in other cellular or molecular functions not examined here (such as the ability to bind to different cellular proteins or to cause immortalisation of host cells). Therefore, further studies should be performed with the HPV 31 E7 variants to see if there are any functional differences between them.

The amino acid changes found in the HPV 31 E6 variants encompassed both the N-terminal and the C-terminal zinc-finger region of the protein. In the closely related HPV 16 E6 protein, several functions map to these two zinc-finger domains. Thus, it seemed promising to initiate a functional analysis of the HPV 31 E6 variants. To this end, we tested two well-established activities of high-risk E6 proteins: the ability to induce the degradation of p53 protein and the ability to inhibit the trans-activation function of p53. In the case of HPV 16 E6, the functional analysis of engineered mutants showed that these two activities of E6 could be more or less separated from each other. We studied the p53 degradation activities of HPV 31 E6 variants in vivo (in stable transfected MCF-7 cell lines) and found that the prototype HPV 31 E6 had weaker activity compared to HPV 16 E6 in inducing the degradation of p53. This finding is in agreement with the results of a previous study reporting the p53 degradation activities of several different HPV E6 proteins. We found no significant differences in the in vivo p53 degradation activities of HPV 31 E6 variants belonging to different variant lineages. To confirm these results, further studies should be performed with higher number of variants. Furthermore, the in vitro p53 degradation activity of the E6 variants could be also studied.

We also explored the HPV 31 E6 variants for their ability to inhibit the transcriptional trans-activation function of p53. We found that, although the E6 variants had comparable abilities to induce the *in vivo* degradation of p53, there were differences between them in the ability to inhibit the trans-activation function of p53. Namely, the prototype HPV 31 E6 (belonging to variant lineage *A*), along with the variant belonging to lineage *C* was active in this function and comparable to the HPV 16 E6 protein, while the variant belonging to lineage *B* showed reduced ability to inhibit the trans-activation function of p53. On the other hand, we recently found that HPV 31 LCR variants belonging to lineage *B* display higher transcriptional activities than variants belonging to lineage *A* or *C*.

These functional results are not perfectly accordant with the results of epidemiological reports studying the clinical behaviour of HPV 31 variants. In these studies, HPV 31 variants belonging to lineage *A* or *B* were shown to have higher potential to cause cervical premalignant lesions than variants belonging to lineage *C*. Regarding this discrepancy between molecular and epidemiological data, it would be important to extend the functional analysis of HPV 31 E6 and E7 variants, and to study further activities of the variant proteins.

SUMMARY

About one-third of human papillomavirus (HPV) types infect the anogenital tract. High-risk genital HPV types (such as HPV 16, 18, 31, 33, and 35) are linked causally to the development of cervical cancer. Nucleotide changes in the LCR and/or E6 and E7 oncogenes may lead to different transcriptional activities and/or different oncogenic potential and this may be one molecular mechanism that is responsible for the differences in the behaviour of intratypic HPV variants.

The purpose of present study was to explore genetic variability and functional differences in the complete LCR and E6, E7 regions of the less studied HPV 31 natural variants isolated from Hungarian women with cervical premalignant lesions. According to the phylogenetical tree of the LCR, E6 and E7 variants three major intratypical lineages (*A, B, C*) were identified. Comparative phylogenetic analysis of the E6/E7 and the LCR region showed that nucleotide sequence variation was lower in the E6/E7 region, but either region could be used with high confidence to identify the variant lineage of an isolate. We found significant differences between the transcriptional activities of the HPV 31 LCR variants belonging to the different variant lineages. On the basis of the experiments with deletion mutants of HPV 31 LCR variants the nucleotide changes causing the functional differences are localized at least partially in the 5' URR domain, on the other hand in the minimal promoter region.

We began the functional analysis of the HPV 31 E6 and E7 variants with testing their effects on p53 and adenovirus E2 promoters. Than we checked the ability of the E6 and E7 variants to induce the degradation of p53 and pRB in vivo. The different HPV 31 E7 sequence variants were able to increase E2F transcriptional activity to the same extent as the prototype HPV 16 E7. The HPV 31 E6 variants belonging to different intratypic lineage had comparable abilities to induce the degradation of p53 *in vivo*, there were significant differences between them in the ability to inhibit the trans-activation function of p53. The prototype HPV 31 E6 belonging to variant lineage *A*, along with the variant belonging to lineage *C* was active to inhibit the trans-activation function of p53, while the variant belonging to lineage *B* showed reduced ability in this function. In the in vivo p53 and pRb degradation activities of HPV 31 E6 and E7 variants belonging to different variant lineages no significant differences were found.

Registry number: DEENK/234/2016.PL
Subject: PhD Publikációs Lista

Candidate: Annamária Ferenczi
Neptun ID: Q1ZTA0
Doctoral School: Doctoral School of Pharmacy

List of publications related to the dissertation

1. **Ferenczi, A.**, Gyöngyösi, E., Szalmás, A., László, B., Kónya, J., Veress, G.: Phylogenetic and functional analysis of sequence variation of human papillomavirus type 31 E6 and E7 oncoproteins.
   Infection, Genetics and Evolution. 43, 94-100, 2016.
   DOI: http://dx.doi.org/10.1016/j.meegid.2016.05.020
   IF: 2.591 (2015)
2. **Ferenczi, A.**, Gyöngyösi, E., Szalmás, A., Hernádi, Z., Tóth, Z., Kónya, J., Veress, G.: Sequence variation of human papillomavirus Type 31 long control region: phylogenetic and functional implications.
   J. Med. Virol. 85 (5), 852-859, 2013.
   DOI: http://dx.doi.org/10.1002/jmv.23542
   IF: 2.217

List of other publications

3. László, B., **Ferenczi, A.**, Madar, L., Gyöngyösi, E., Szalmás, A., Szakács, L., Veress, G., Kónya, J.: CpG methylation in human papillomavirus (HPV) type 31 long control region (LCR) in cervical infections associated with cytological abnormalities.
   Virus Genes. 52 (4), 552-555, 2016.
   DOI: http://dx.doi.org/10.1007/s11262-016-1338-6
   IF: 1.285 (2015)

DEBRECENI EGYETEM KENÉZY ÉLETTUDOMÁNYI KÖNYVTÁRA EGYETEMI ÉS NEMZETI KÖNYVTÁR

4. Gall-Debreceni, A., Lázár, J., Kádas, J., Balogh, A., **Ferenczi, A.**, Sós, E., Takács, L., Kurucz, I.: Specific detection and quantitation of bovine IgG in bioreactor derived mouse mAb preparations.
   J. Immunol. Methods. [Epub ahead of print], 2016.
   DOI: http://dx.doi.org/10.1016/j.jim.2016.08.005
   IF: 1.858 (2015)

5. Gyöngyösi, E., Szalmás, A., **Ferenczi, A.**, Póliska, S., Kónya, J., Veress, G.: Transcriptional regulation of genes involved in keratinocyte differentiation by human papillomavirus 16 oncoproteins.
   Arch. Virol. 160 (2), 389-398, 2015.
   DOI: http://dx.doi.org/10.1007/s00705-014-2305-y
   IF: 2.255

6. Szalmás, A., Gyöngyösi, E., **Ferenczi, A.**, László, B., Karosi, T., Csomor, P., Gergely, L., Veress, G., Kónya, J.: Activation of Src, Fyn and Yes non-receptor tyrosine kinases in keratinocytes expressing human papillomavirus (HPV) type 16 E7 oncoprotein.
   Virol. J. 10 (1), 79, 2013.
   DOI: http://dx.doi.org/10.1186/1743-422X-10-79
   IF: 2.089

7. Gyöngyösi, E., Szalmás, A., **Ferenczi, A.**, Kónya, J., Gergely, L., Veress, G.: Effects of human papillomavirus (HPV) type 16 oncoproteins on the expression of involucrin in human keratinocytes.
   Virol. J. 9, 36, 2012.
   DOI: http://dx.doi.org/10.1186/1743-422X-9-36
   IF: 2.092

Total IF of journals (all publications): 14,387
Total IF of journals (publications related to the dissertation): 4,808

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

21 September, 2016

List of major presentations

Ferenczi A., Gyöngyösi E., Kónya J., Veress G.: Functional analysis of HPV 31 LCR (LONG CONTROL REGION) sequence variation, Annual Meeting of the Hungarian Society for Microbiology, October 12-15, 2010, Hotel Helikon, Keszthely, Hungary.

Ferenczi A., Gyöngyösi E., Kónya J., Veress G.: Human papillomavirus 31 LCR (Long Control Region) sequence variation: phylogenetic and functional analysis, 16th International Congress of the Hungarian Society for Microbiology, July 20-22, 2011. Eötvös Loránd University Convention Centre, Budapest, Hungary.

Ferenczi A., Gyöngyösi E., Szalmas A.; Kónya J., Veress G.: Sequence variation of human papillomavirus (HPV) type 31 E6 and E7 oncoproteins: phylogenetic and functional analysis, Annual Meeting of the Hungarian Society for Microbiology, October 24-26, 2012, Hotel Helikon, Keszthely, Hungary.

Ferenczi A., Gyöngyösi E., Szalmas A.; Kónya J., Veress G.: Phylogenetic and functional analysis of sequence variation of human papillomavirus (HPV) type 31 E6 and E7 oncoproteins,4th Central European Forum for Microbiology, October 16-18, 2013, Hotel Helikon, Keszthely, Hungary.

Ferenczi A., Gyöngyösi E., Szalmas A.; Kónya J., Veress G.: Phylogenetic and functional analysis of sequence variation of human papillomavirus (HPV) type 31 E6 and E7 oncoproteins, Annual Meeting of the Hungarian Society for Microbiology, October 15-17, 2014, Hotel Helikon, Keszthely, Hungary