PERGAMON

European Journal of Cancer 36 (2000) 1402–1410

European Journal of Cancer

www.ejconline.com

Variants of the long control region of human papillomavirus type 16

K. Kurvinen [a], M. Yliskoski [b], S. Saarikoski [b], K. Syrjänen [a], S. Syrjänen [a,c,*]

[a] *MediCity Research Laboratory, Faculty of Medicine, University of Turku, Tykistökatu 6 A, FIN-20520, Turku, Finland*
[b] *Department of Obstetrics and Gynecology, Kuopio University Hospital, PO Box 1777, FIN-70211 Kuopio, Finland*
[c] *Department of Oral Pathology and Oral Radiology, University of Turku, Lemminkäisenkatu 2, FIN-20520 Turku, Finland*

Received 17 September 1999; received in revised form 8 February 2000; accepted 16 February 2000

Abstract

Expression of the human papillomavirus (HPV) *E6* and *E7* oncogenes is regulated on the transcriptional level by specific protein-binding sites contained in the viral long control region (LCR). Sequence changes within the LCR region may have an impact on the transcription of viral oncogenes, possibly resulting in differences in the oncogenic potential of the virus. The present study was designed to determine the sequence variability of the LCR of HPV 16 and to assess whether certain LCR variants do correlate with the clinical outcome of the disease of the uterine cervix. The entire LCR segment of HPV 16 was analysed from 37 cervical biopsy specimens derived from 28 women included in the Kuopio long-term prospective follow-up study. The LCR sequence was identical with the reference sequence in six HPV 16 isolates. Overall, 14 different HPV 16 LCR variants were identified. One of the variants showed sequence variation typical of the Asian–American variant lineage of HPV 16, and all the other variants appeared to belong to the European variant group. The European variants exhibited low genetic diversity, and only five of these LCR variants contained nucleotide changes involving known or proposed binding sites for transcription factors. The variants with changes at nucleotide positions 7193 and 7521 was the most prevalent, accounting for almost 37% of infections. This variant (7193; 7521) has been previously demonstrated to have similar transcriptional activity compared with the reference isolate by Veress and colleagues *J Gen Virol* 1999, **80**, 1035–1043. The reference isolate, variant (7193; 7521) and variants with changes within transcription factor binding sites accounted for most of the infections, and no significant differences were found in the comparison of the distribution of these different LCR isolates in cases where the disease showed progression to severe cervical intraepithelial neoplasia (CIN) or carcinoma *in situ* (CIS). Notably, both the reference isolate and variant (7193; 7521) were also closely associated with infections showing more aggressive behaviour. According to the present findings, in European HPV 16 isolates, intratype genetic variation of the LCR region does not seem to be commonly responsible for differences in the pathogenicity of the virus and thereby for a risk of progressive infections.

Keywords: HPV; LCR; Sequence variation; Viral pathogenicity; Cervical neoplasia

1. Introduction

Human papillomavirus (HPV) has been implicated as a primary aetiological factor in cervical cancer. Over 90% of cervical cancers contain HPV DNA, and the most prevalent virus type appears to be HPV 16. It is now well established that HPV infection plays a significant role in the initiation of a continuum through various stages of cervical intraepithelial neoplasias (CINs) to invasive cancer [1]. The main transforming proteins of HPVs are E6 and E7. The expression of E6 and E7 contributes to the proliferative growth phenotype of HPV-infected cells, and both these viral proteins are required for effective immortalisation of primary human squamous epithelial cells [2,3].

The expression of HPV 16 early genes, including the *E6* and *E7* genes, is directed by the transcriptional promoter P97 located at the *E6*-proximal end of the long control region (LCR). The activity of the promoter is regulated by both viral and cellular proteins, which mediate their transcriptional regulation through interacting with their specific protein-binding sites within the

* Corresponding author. Tel.: +358-2-3338349; fax: +358-2-3337000.
E-mail address: stisyr@utu.fi (S. Syrjänen).

0959-8049/00/$ - see front matter
PII: S0959-8049(00)00121-0

LCR [4]. HPV *E2* gene products can either activate or repress viral transcription, depending on the position of E2 target sites. In the context of the natural HPV 16 promoter, E2 behaves rather as a repressor by interacting with two promoter proximal E2 motifs [4,5]. Stimulation of the promoter activity is conferred by a number of cellular transcription factors, including activator protein-1 (AP-1), nuclear factor-1 (NF-1), transcriptional enhancer factor-1 (TEF-1), SP-1 and glucocorticoid/progesterone receptors [6–9]. By contrast, nuclear factor for interleukin-6 (NF-IL6) and yin-yang 1 (YY1) can participate in the downregulation of the HPV 16 promoter [10,11].

The natural history of HPV infection is still incompletely understood. Persistent infection by high-risk HPVs is considered to be a major risk factor for malignant progression of HPV-induced lesions. Although the highest progression rate of HPV infections has been associated with HPV 16, a significant proportion of HPV 16-positive lesions still persist or spontaneously regress [12,13]. The integration of HPV DNA into the host cell genome has been thought to be a critical step in malignant progression, because of its favourable effects on viral oncogene expression. Integration events commonly disrupt the *E1*/*E2* gene region, resulting in loss of E2 expression [14,15]. Thus, elimination of the E2-mediated transcriptional repression by integration potentially leads to increased expression of the *E6* and *E7* oncogenes, providing a selective growth advantage for the affected cell [16].

Analysis of the regulatory regions of episomal HPV 16 DNAs derived from cervical cancers revealed deletions and mutations within target sequences for the transcriptional repressor *YY1*, and notably, these alterations appeared to contribute to increased activity of the promoter [11,17]. These findings suggest that the disruption of *YY1* binding sites within the LCR may represent an important way to escape from cellular transcriptional repression, resulting in the upregulation of E6/E7 expression. These observations also provide evidence for the hypothesis that naturally occurring sequence changes in the LCR may, at least partially, be responsible for differences in the biological properties of HPV 16 variants, possibly resulting in differences in their pathogenicity. Indeed, recent clinical studies have supported the concept that certain LCR variants of HPV 16 may be more closely associated with risk for progression of the disease than others [18,19].

In the present study, the sequence variation of the entire LCR segment of HPV 16 was determined in a series of cervical biopsy specimens derived from 28 women included in the Kuopio long-term prospective follow-up study. Furthermore, the physical state of HPV 16 DNA in the infected cells was evaluated. Special attention was focused on assessing whether certain HPV 16 variants do correlate with the clinical outcome of the disease of the uterine cervix.

2. Materials and methods

2.1. Patients and samples

The study material consisted of cervical biopsy specimens derived from a cohort of more than 1000 women, prospectively monitored since 1981 for genital HPV infections at the Department of Obstetrics and Gynecology, Kuopio University Hospital, Kuopio, Finland. A group of 28 women with HPV 16 infection included in the present study was chosen by the criteria that at least one frozen biopsy specimen was available from the follow-up. The 37 cervical samples studied were obtained from two groups: 24 samples from a follow-up group (16 patients) and 13 samples from a treatment group (12 patients). The mean follow-up time of the patients was approximately 6.5 years.

The design of the follow-up study has been described previously in greater detail [20,21]. Briefly, the patients were regularly examined at 6-monthly intervals, and at each visit, the patients were subjected to a thorough gynaecological examination. In the follow-up group, the patients were followed-up without any kind of treatment, unless progression to CIN III or carcinoma *in situ* (CIS) was established. In this case, the lesions were eradicated by conisation. The clinical course of the infection (compared with the status at the first visit) was classified as either regressed, persistent, progressed or recurred, depending on the presence or absence of the lesion on colposcopy, HPV-induced cytological changes in Papanicolaou (PAP) smears and/or typical morphology in direct punch biopsies. To classify a lesion as regressed (or cured), all three examination techniques showed a negative finding [21]. In the treatment group, the patients with cervical HPV lesions were treated with one of the following modes: conisation, laser, cryotherapy or interferon. The clinical course of each HPV lesion after treatment was classified as either cured, residual disease or recurred, according to the same criteria as described above [20]. Noteworthy, in the present study, also in the treatment group, the clinical course of the disease was categorised as progressed when the disease showed progression to CIN III or CIS despite treatment.

2.2. Isolation of DNA

DNA was extracted from frozen tissues according to the method of Miller and colleagues [22]. Samples were lysed in 1 ml of 10 mM Tris (pH 8.3), 400 mM NaCl, 1% SDS, 2 mM EDTA and 300 μg/ml proteinase K overnight at 37°C. Protein precipitation was carried out by adding 300 μl of saturated NaCl. After centrifugation, the supernatant was removed and DNA was precipitated with 100% ice-cold ethanol. DNA was dissolved in sterile water.

2.3. HPV DNA analysis

The presence of HPV 16 DNA was determined by polymerase chain reaction (PCR), using two primer pairs hybridising to the *E6* and *E7* regions of the HPV 16 genome. The primers are shown in Table 1. PCR was carried out in a 50 μl reaction mixture containing 300 ng of DNA sample, 20 pmol of each primer, 200 μmol/l of each deoxynucleotide triphosphate and 1.3 U AmpliTaq Go1d™ DNA polymerase (Perkin Elmer, Roche Molecular Systems, Inc., Branchburg, NJ, USA) in enzyme-specific GeneAmp® PCR buffer (Perkin Elmer). Template DNA was first denatured at 94°C for 10 min, and then 35 amplification cycles were carried out in a thermal cycler (Perkin Elmer Cetus, Norwalk, CT, USA) as follows: DNA denaturation at 94°C for 1 min, primer annealing at 62°C for 1 min and primer extension at 72°C for 1 min. The final extension step was prolonged for another 10 min at 72°C. The PCR products were analysed on a 2% agarose gel stained with ethidium bromide (SeaKem LE agarose, FMC BioProducts, Rockland, ME, USA).

The physical state of the HPV 16 genome was determined by analysing the samples for the presence of the *E1* and *E2* genes by PCR and Southern blot hybridisation. Table 1 shows the PCR primers targeting the *E1* and *E2* open reading frames (ORFs). PCR amplification with *E1* and *E2* primers was carried out as described above, except that the primer extension step at 72°C was extended to 2 min. HPV 16 plasmid DNA was used as a positive control, and the PCR reaction mixture without DNA served as a negative control. A 25 μl sample of the PCR product was run in a 2% agarose gel (SeaKem LE agarose) and then transferred to a Nylon membrane (GeneScreen Plus, NEN™ Life Science Products, Inc., Boston, MA, USA). Hybridisation was carried out with digoxigenin-labelled oligonucleotide probes (*E1*, nt 1852–1881: 5′-AGT GAA GTG TAT GGA GAC ACG CCA GAA TGG, Tm 60°C; *E2*, nt 3323–3352: 5′-GGT CAG GTA ATA TTA TGT CCT ACA TCT GTG, Tm 55°C) (DIG Oligonucleotide 3′-End Labeling Kit, Boehringer Mannheim, Mannheim, Germany). The hybrids were detected using antidigoxigenin antibody conjugated to alkaline phosphatase and chemiluminescence substrate (CSPD, DIG Luminescent Detection Kit, Boehringer Mannheim).

2.4. PCR of the HPV 16 LCR

The entire LCR segment (spanning nucleotides 7110–222) was amplified by PCR, using primers as previously described by Xi and colleagues [23] (Table 1). The PCR reaction mixture was the same as that described above. The PCR cycle conditions were as described in the HPV DNA analysis for *E1* and *E2*.

2.5. Electrophoretic analysis of the LCR amplification products

The purity and fragment size of the amplified LCR fragments was verified by means of electrophoretic high resolution analysis using the PhastSystem electrophoresis apparatus (Pharmacia LKB, Uppsala, Sweden). The PCR products were run on a 20% PhastGel homogeneous polyacrylamide gel with PhastGel native buffer strips (Pharmacia Biotech, Uppsala, Sweden). Separation was carried out at 5°C for 350 volthours (maximum values: 300 V, 20.0 mA, 7.0 W). After electrophoresis, the gels were silver-stained.

Table 1
Genomic positions and sequences of the oligonucleotide primers used

Region	Nucleotide positions for primers (5′ > 3′)[a]	Primer sequence 5′ > 3′
HPV16-E6F[b]	112 > 136	GGA **T**CC ACA GGA GCG ACC CAG AAA G
HPV16-E6R[b]	535 < 556	**CTG** CAG CTG GGT TTC TCT ACG TGT T
HPV16-E7F[b]	577 > 597	**GGA T**CC TAC ATT GCA TGA ATA TAT G
HPV16-E7R[b]	821 < 842	**CTG** CAG ATG GGG CAC ACA ATT CCT A
HPV1 6-E1F	858 > 877	ATC TAC CAT GGC TGA TCC TG
HPV1 6-E1R	2821 < 2840	GGT CAC GTA GGT CTG TAC TA
HPV16-E2F	2739 > 2758	CGA GGA CAA GGA AAA CGA TG
HPV16-E2R	3857 < 3876	TGT GGA TGC AGT ATC AAG AT
HPV16-LCR-F	7110 > 7138	CCT CAT CTA CCT CTA CAA CTG CTA AAC GC
HPV16-LCR-R	192 < 222	CGT CGC AGT AAC TGT TGC TTG CAG TAC ACA C
LCR1-seqprimer	7502 < 7529	GTT TAA ACC ATA GTT GCT GAC ATA GAA C
LCR2-seqprimer	7447 > 7470	GCT TCA ACC GAA TTC GGT TGC ATG

F, forward primer; R, reverse primer.

[a] The nucleotide positions of the primers were numbered according to the HPV16R sequence.

[b] The *E6* and *E7* primers contain restriction sites for *Bam*HI (GGATCC) and *PstI* (CTGCAG). Bold letters in the sequence of these primers indicate nucleotides which have been added or changed to create the specific restriction sites.

2.6. Cloning and sequencing

The LCR amplification products were cloned into a *Sma*1 digested pUC18 vector (SureClone® Ligation Kit, Pharmacia Biotech). Following the transformation of competent *Escherichia coli* JM109 cells, potential insert-bearing clones were selected, and then the structure of these clones was checked by the restriction analysis of plasmid DNA. DNA of the recombinant plasmids was prepared for sequencing using the Wizard *Plus* Minipreps DNA Purification System (Promega, Madison, WI, USA). Sequencing reactions were produced using the ABI PRISM BigDye™ Terminator Cycle Sequencing Kit (PE Applied Biosystems, Perkin Elmer, Norwalk, CT, USA) according to the manufacturer's instructions, and analysis of sequencing results was carried out on the automated ABI PRISM 377 DNA Sequencer (PE Applied Biosystems, Perkin Elmer). Universal forward and reverse M13 primers and two HPV 16 LCR primers were used for sequencing. The HPV 16 LCR primers have been described previously by Xi and colleagues [23] (Table 1).

2.7. Sequence analysis

In the present study, the nucleotide positions of the HPV 16 genome were numbered according to the reference sequence of the HPV 16 genome (HPV16R), which was obtained from an HPV database (hhtp://hpv-web.lanl.gov). Several sequencing errors have been found in the original HPV 16 prototype sequence published by Seedorf and colleagues [24]. The corrections of the errors have been incorporated into the HPV16R sequence. The differences within the LCR between the originally published sequence and the revised sequence are as follows: GC-to-CGG change at positions 7432–7433 and deletion of nucleotide A at position 7861 (nucleotide positions given according to the original sequence). When all errors found in the original HPV 16 sequence have been considered, the numbering of the nucleotide positions of the LCR segment differs from that of the original prototype sequence by two or three additional nucleotides, depending on the genomic region.

Different combinations of sequence changes detected in the LCR of HPV 16 isolates were termed variants. Each sample showing variation was amplified, cloned and sequenced twice, except when a variant was identical with variants previously identified in the study. Sequence changes that were found at least twice in the same sample were counted as variants, whereas differences that were found only once were counted as potential PCR artifacts. Different variants each identified at least twice in one specimen were counted as cases of multiple infection.

3. Results

3.1. HPV DNA analysis

First, the presence of HPV 16 DNA was confirmed by PCR amplification, using primer pairs targeting the *E6* and *E7* regions of the HPV 16 genome (data not shown). Then, the physical state of HPV 16 DNA was determined by analysing the samples for the presence of the *E1* and *E2* genes by PCR and Southern blot hybridisation. According to previous studies, the presence of intact *E1* and *E2* ORFs correlates with an episomal form of HPV 16 DNA, whereas the *E1*/*E2* gene region is probably disrupted whenever the virus is integrated [25,26]. The *E1* ORF was amplified in all samples, whereas no *E2* amplification was detected in five samples obtained from 2 women, indicating integration of HPV 16 DNA. Table 2 shows the physical state of each HPV 16 isolate. In both cases where HPV 16 DNA showed integration (F6 and F12 in Table 2a), viral DNA was present as an integrated form already in CIN II lesions.

3.2. Sequence analysis of the HPV 16 LCR

Sequence variation of the whole LCR segment of HPV 16 was determined from 37 cervical biopsy specimens derived from 28 women. Fig. 1 shows nucleotide changes found in the LCR sequences of the different HPV 16 isolates compared with the reference sequence. Altogether, 28 different point mutations were detected in 26 nucleotide positions within the LCR segment spanning nucleotides 7110–222. At least 19 of these sequence changes had been identified and reported previously by other authors (Fig. 1) (HPV-16R variants, the HPV compendium 1996 from the HPV database; [18,27,28]. In addition to these earlier described alterations, nine point mutations were observed at nucleotide positions 7139 (A to C), 7168 (A to G), 7172 (T to C), 7174 (A to C), 7310 (C to T), 7435 (G to A), 7456 (G to A), 7496 (T to A) and 7666 (A to C) (Fig. 1). Overall, 14 HPV 16 LCR variants were identified, and the number of changes in individual variants ranged from 1 to 16. Only in six HPV 16 isolates, was the LCR sequence identical with the reference sequence. Two or more samples of the same patient obtained at different follow-up visits were analysed in 5 cases (F6, F7, F11, F12 in Table 2a; T2 in Table 2b). These individuals proved to be infected by the same variant. Double infection with different HPV 16 variants was detected in two specimens (F7 and F16 in Table 2a). At least in one of these cases, the coinfection appeared to be transient (F7).

The variants with changes at positions 7193 and 7521 was the most prevalent, accounting for almost 37% (11 of 30) of HPV infections (including also two coinfection cases). Altogether, combination of mutations at nucleo-

Table 2
Clinical outcome of the disease in relation to HPV 16 long control region (LCR) variants and physical state of HPV DNA

(a) Follow-up patients[a]	Clinical course of the disease[b] (sampling year: grade of lesion)	Clinical outcome of the disease	HPV 16 LCR variants (changed nucleotide positions)	Physical state of HPV DNA
F1	83: CIN I; 85: NCIN; 85: CIS (co) (−91)	Progressed to CIS	7193; 7521	Episomal
F2	84: NCIN; 85: CIN II; 85: CIS (co) (−86)	Progressed to CIS	7193; 7233; 7339; 7394; 7395; 7485; 7489; 7521; 7669; 7689; 7729; 7743; 7764; 7786; 7886; 145	Episomal
F3	84: NCIN; 85: NCIN; 85: normal (−94)	Regressed to normal	7193; 7310; 7521	Episomal
F4	84: CIN II; 84: CIN III; 85: CIS (co) (−92)	Progressed to CIS	Prototype	Episomal
F5	84: CIN I; 84: CIN II; 85: CIN II 85: NCIN; 86: NCIN; 87: normal (−94)	Regressed to normal	7193; 7521	Episomal
F6a–b	86: NCIN; 89: CIN I; 90: CIN I; 90: CIN II; 91: CIN II; 91: CIN III; 91: CIS (co) (−95)	Progressed to CIS	7193; 7521	Integrated
F7a–e	84: CIN II; 85: NCIN; 85: CIN I; 86: CIN II ; 86: CIN II; 87: CIN I; 87: CIN III 87: CIS (co) (−91)	Progressed to CIS	(I) 7193; 7521 (a–e); double infection: (I) 7193; 7521; (II) 7193; 7456; 7521 (transient) (c)	Episomal
F8	85: CIN II; 86: CIS (co) (−86)	Progressed to CIS	7193; 7521	Episomal
F9	86: CIS (co) (−95)	Progressed to CIS	7193; 7521	Episomal
F10	84: CIN II; 85: CIN III (co); 86: CIN I; 87: NCIN; 95: normal (−95)	Progressed to CIN III	7193; 7521	Episomal
F11a–b	85: CIN I; 86: CIN II; 86: CIN I; 87: CIN II; 88: CIN I; 89: NCIN; 89: CIN II; 90: CIN I; 91: CIN II; 91: NCIN; 92: CIS (co) (−94)	Progressed to CIS	7193; 7496; 7521	Episomal
F12a–c	85: CIN I; 85: NCIN; 86: CIN II; 86: CIN II; 87: CIN III; 87: CIS (co) (−95)	Progressed to CIS	7174; 7193; 7496; 7521	Integrated
F13	85: NCIN; 87: NCIN; 88: NCIN; 88: normal (−92)	Regressed to normal	7666	Episomal
F14	85: CIN II; 86: CIN I; 87: NCIN; 87: normal (−94)	Regressed to normal	Prototype	Episomal
F15	87: NCIN; 87: normal (−89)	Regressed to normal	7193; 7496; 7521	Episomal
F16	87: CIN I (−88)	Persistent	Double infection: (I) prototype; (II) 7193; 7521	Episomal
(b) Treatment patients[a]	Clinical course of the disease[b] (sampling year: grade of lesion)	Clinical progress/ treatment response[c]	HPV 16 LCR variants (changed nucleotide positions)	Physical state of HPV DNA
T1	86: CIN I (cr); 87: NCIN; 89: normal (−94)	Cured	7193; 7521	Episomal
T2a–b	87: CIN II (IFN); 88: CIS (co); 88: normal; 91: normal; 92: CIN I; 93: normal (−94)	Progressed to CIS/cured	7193; 7507; 7521	Episomal
T3	87: CIN II (IFN); 88: normal; 89: CIN III (co) (−95)	Progressed to CIN III/ Cured	7193; 7521	Episomal
T4	87: CIN I (cr); 89: normal (−95)	Cured	7168; 7193; 7435; 7521	Episomal
T5	87: CIN III (co) (−95)	Progressed to CIN III/ Cured	7193; 7310; 7521	Episomal
T6	87: CIN III; 87: CIS (co) (−94)	Progressed to CIS/cured	Prototype	Episomal
T7	87: CIS (co) (−92)	Progressed to CIS/cured	7139; 7193; 7310; 7521	Episomal
T8	87: CIN III (co) (−88)	Progressed to CIN III/ Cured	7193; 7233; 7485; 7496; 7521	Episomal
T9	89: CIS (co) (−91)	Progressed to CIS/cured	7193; 7521	Episomal
T10	89: CIS (co) (−95)	Progressed to CIS/cured	Prototype	Episomal
T11	90: CIN III (co) (−95)	Progressed to CIN III/ Cured	Prototype	Episomal
T12	90: CIN II (l) 90: normal (−95)	Cured	7172	Episomal

CIN, cervical intraepithelial neoplasia; NCIN, non-CIN; CIS, carcinoma *in situ*.

[a] Lowercase letters after number indicate multiple samples analysed from the same patient.

[b] The cervical specimens analysed in this study have been underlined. The last visit of the follow-up is given in parentheses after the data of clinical course. The lesions were always eradicated by conisation when progression to CIN III or CIS was established. Treatment methods are shown in parentheses and abbreviations of the treatment methods are as follows: co, conisation; cr, cryotherapy; l, laser vaporisation; IFN, interferon.

[c] In the treatment group, the clinical course has been categorised as progressed when the disease showed progression into CIN III or CIS.

	7139	7168	7172	7174	7193	7233	7310	7339	7394	7395	7435	7456	7485	7489	7496	7507	7521	7666	7669	7689	7729	7743	7764	7786	7886	145
Reference sequence	A	A	T	A	G	A	C	A	C	C	G	G	A	G	T	A	G	A	C	C	A	T	C	C	C	G
F4, F14, F16 (I)[a] T6, T10, T11																										
T12			C																							
F13																		C								
F1, F5, F6a–b, F7a–e, F8, F9, F10, F16 (II), T1, T3, T9					T												A									
F3					T		T										A									
F7c (II)					T							A					A									
F11a–b					T										C		A									
F15					T										A		A									
T2a–b					T											C	A									
T5					T		G										A									
F12a–c				C	T										C		A									
T4		G			T						A						A									
T7	C				T		G										A									
T8					T	C							C		C		A									
F2					T	C		T	T	T			C	A			A		T	A	C	G	T	T	G	T
Described earlier					T	C	G	T	T	T			C	A	C	C	A		T	A	C	G	T	T	G	T
Transcription factor binding site							NF-IL6		GRE/1			E2	GRE/2							TEF-1						

[a] Figures in parentheses represent different variants.

Fig. 1. Nucleotide changes in the LCR sequences of different HPV 16 isolates. The first column specifies the isolate code, and the numbers at the top indicate the genomic positions, where sequence changes were detected. Each row represents one LCR variant. The nature of the mutation is indicated in a single letter code, whereas a white field indicates that this position is identical with the reference sequence. Sequence changes identified and reported earlier are marked in grey squares. Nucleotide changes involving known or proposed transcription factor binding sites are indicated in the bottom alignment. NF-IL6, nuclear factor for interleukin-6; GRE, glucocorticoid/progesterone response element; E2, viral E2 protein; TEF-1, transcriptional enhancer factor 1. [a]Figures in parentheses represent different variants in the same specimen; II represents coinfection.

tide positions 7193 and 7521 was present 12 of the 14 variants. Studies of the molecular variants of HPV 16 have often been restricted to the sequence analysis of the LCR segment spanning a region from position 7480 to 7843 [28,29]. These studies indicate that a point mutation at position 7521 is the most frequent sequence change worldwide, and HPV 16 LCR variants identical with the reference clone (prototype) or with only a mutation at 7521 are especially common in Europe. The study of Smits and colleagues [30] showed that sequence changes in both positions 7193 and 7521 are common in the Dutch and Barbadian variants. Most LCR variants identified in this study appeared to belong to the European variant lineage of HPV 16. Only a variant with 16 sequence changes (F2) showed genetic variation typical of the Asian–American lineage; all nucleotide changes within the segment from 7480 to 7843 were identical to those found in Asian–American variants [28,29].

Of the LCR variants identified, six variants contained nucleotide changes within known or proposed binding sites for viral or cellular transcription factors (Fig. 1) [31]. The Asian–American variant (F2) showed nucleotide changes located within two of the three known glucocorticoid response elements (GREs) and within one binding site for TEF-1 [8,9]. Of the European variants, three variants (F3, T5, T7) carried either a C-to-T or C-to-G

change at nucleotide position 7310 located within one of the proposed binding sites for nuclear factor for interleukin-6 (NF-IL6) [10], and one variant (T8) contained a A-to-C change at position 7485 within the GRE [9].

Additionally, a variant associated with a transient coinfection (F7c (II)) showed a point mutation within one of the four recognition sites for the viral E2 protein [32].

3.3. Sequence variation of the LCR as related to the clinical outcome of cervical lesions

The material of this study consisted of cervical biopsy specimens derived from women included in the Kuopio cohort study. Of the 37 cervical specimens studied, 24 samples were obtained from the follow-up group (16 patients) and 13 samples from the treatment group (12 patients). The morphological changes of the HPV lesions analysed were: 4 with non-CIN (NCIN; HPV lesion without concomitant CIN), 4 with CIN I, 12 with CIN II, 8 with CIN III and 9 with CIS. The clinical outcome of individual patients was normal cytology in 5 cases, CIN III in 5 cases and CIS in 14 cases. Compared with the status at the beginning of the screening period, the clinical course of HPV lesions showed regression in 5 cases and progression in 19 cases. Because of either the short follow-up time or treatments, the natural history of HPV infections was not known in cases classified as persistent infection (F16 in Table 2a) and cured disease (T1, T4, T12 in Table 2b). Thus, these cases were not included in the comparison of the disease risk of the different HPV 16 variants.

Table 2 shows the clinical course of the disease and disease outcome in relation to HPV 16 LCR variants identified. Because of the presence of multiple different LCR variants amongst the 28 HPV-infected women examined, it was impossible to estimate the disease risk of the individual variant. Reference isolate, the variant with changes at nucleotide positions 7193 and 7521 and variants with changes within transcription factor binding sites accounted for most of the infections (19 of the 24 cases included in the comparison of the disease risk of variants; F16, T1, T4, T12 not included as mentioned above). Table 3 shows the distribution of these LCR isolate groups among different grades of CIN lesions and CIS (disease outcome). The Asian–American variant, reference isolate or the variant with changes at positions 7193 and 7521 were found to be involved in infections showing rapid progression of the disease to CIS (from NCIN/CIN II to CIS within 5–7 months; Fl, F2, F4, F8 in Table 2a). Additionally, the variant with changes at 7193 and 7521 was associated in one of the cases with relapsing infections after conisation (F10, Table 2a).

4. Discussion

The present study was designed to determine the sequence variation of the LCR of HPV 16 and to assess whether certain LCR variants do correlate with the clinical outcome of the disease of the uterine cervix. The entire LCR region was analysed in HPV 16 isolates from a series of cervical biopsy specimens obtained from 28 women included in the Kuopio cohort study. Furthermore, the physical state of HPV 16 DNA in the infected cells was evaluated. On the basis of previous studies of the intratype genetic diversity of HPV 16, one of the HPV 16 isolates showed sequence variation typical of the Asian–American variant lineage, and all the other isolates belonged to the European variant group [28,29]. The LCR sequence appeared to be identical with the reference sequence in six HPV 16 isolates. Of the 14 LCR variants identified, the variant with changes at nucleotide positions 7193 and 7521 was most prevalent, accounting for almost 37% of infections. Double infection with different HPV 16 variants was detected in two specimens. In one of these cases, five specimens of the same patient were analysed over the screening period, and double infection was found only once. One of the variants disappeared during 6 months, suggesting that one major variant predominates over time, whereas coinfection is more transient, as previously reported by Xi and coworkers [33]. HPV 16 DNA was detected as an integrated form in two infections.

Several studies have provided evidence that some sequence changes within the transcriptional regulatory sites of the HPV 16 LCR may result in an enhanced activity of the promoter [11,17,27,34]. These findings

Table 3
Distribution of different HPV 16 LCR isolates in different grades of CIN lesions and CIS

	Disease progression		
	NCIN/CINI	CINII/CINIII/CIS	Total
Reference isolate	1	4	5
Variant with changes at 7193 and 7521	1	8	9
Variants with changes within transcription factor binding sites	1	4	5
Total	3	16	19

CIN, cervical intraepithelial neoplasia, NCIN, non-CIN; CIS, carcinoma *in situ*.

suggest that alterations within the LCR may have an favourable impact on the expression of viral oncogenes, thereby possibly facilitating the potential for malignant progression. Of the LCR variants identified in the present study, five variants contained nucleotide changes within known or proposed binding sites for cellular transcription factors. The Asian–American variant showed several sequence changes, and some of the alterations were within two GREs and within a TEF-l binding site. A recent study by Veress and coworkers [34] showed that this Asian–American variant exhibits an approximately 1.7-fold increased promoter activity in transfection assays compared with the reference LCR isolate. Most of the increased transcriptional activity of the variant appeared to be due to nucleotide changes located in the 3′ half of the LCR (nt 7660–7890), also involving a TEF-l motif [34]. Notably, in the present study, the Asian–American variant was associated with one of the infections showing rapid progression of the disease to CIS. The European variants exhibited only low genetic diversity, and some nucleotide changes found in these variants involved binding sites for NF-IL6 and the glucocorticoid/progesterone receptor. Three variants contained a C-to-T or C-to-G change at nucleotide position 7310 within a proposed binding site for a negative regulator factor, NF-IL6, and both infections with variants having the C-to-G change showed progression to CIN III (T5) or CIS (T7). Interestingly, in a previous study by Chen and associates [27], an HPV 16 LCR isolate from an oral cancer cell line containing nucleotide changes within the same NF-IL6 motif was found to have enhanced transcriptional activity compared with the reference isolate, and one of the observed alterations was the C-to-G change at nt 7310. Thus, one specific aim of our future studies will be to find out by transfection assays whether the LCR variants with the affected NF-IL6 binding site are able to modify the promoter activity.

Reference isolate, the variant with changes at nucleotide positions 7193 and 7521 and variants with changes within transcription factor binding sites accounted for most of the infections, and no significant differences were found in the comparison of the distribution of these different LCR isolates in cases where the disease showed progression to CIN III or CIS. In a study by Veress and colleagues [34], analysis of the functional consequences of sequence variation in the HPV 16 LCR demonstrated that the variant with changes at 7193 and 7521 had similar transcriptional activity compared with the reference isolate. Notably, in the present study, either the reference isolate or variant (7193; 7521) was found to be associated with 4 of the 6 cases where HPV 16 infections showed more aggressive behaviour, including the infections showing rapid progression of the disease to CIS (F1, F2, F4, F8) and the infections showing relapsing disease (F10, T2). Furthermore, viral DNA was found to be in an episomal form in all of these cases. Thus, neither of the proposed mechanisms (sequence changes of the LCR or integration of HPV DNA) for increased transcription of the *E6* and *E7* oncogenes appeared to be responsible for more aggressive behaviour of the infections. Overall, on the basis of the present findings, it can be concluded that sequence variation of the HPV 16 LCR is not commonly associated with a risk of progression of the disease to severe CIN or CIS.

HPV 16 DNA is often maintained as an episomal form in benign and premalignant cervical lesions, whereas viral DNA is frequently integrated into the host cell genome in cervical cancers and cancer-derived cell lines [14,15,35]. Thus, the integration of HPV DNA has been thought to be a critical step in malignant progression. In accordance with previous studies [25,36], analysis of the physical state of HPV DNA in the present study showed that integration of viral DNA occurs infrequently in precancerous lesions, including also CIS cases. The integrated forms of HPV 16 DNA were detected in biopsy specimens obtained from two women (F6 and F12). In both cases, viral DNA was already totally integrated in the CIN II lesion, and the patient's clinical status showed progression from CIN II to CIS within 13 months. When compared with lesions containing episomal viral DNA, no significant differences in clinical progress were found. However, the integration of HPV DNA is considered as an important hallmark for the transition from premalignant lesions to invasive cancer, and thereby, the different physical status of viral DNA in HPV-associated lesions may be related to differences in the progression of the disease into the invasive state.

In conclusion, the European variants of HPV 16 appeared to exhibit only low genetic diversity in the LCR region, and sequence changes infrequently involved known or proposed binding sites for transcription factors. These findings support the concept that sequence variation in the HPV 16 LCR is not a common strategy to promote the transcription of viral oncogenes and to facilitate the oncogenic potential of the virus. There exists evidence that variants of the HPV 16 E6 protein may have differences in their oncogenic properties [37,38]. Thus, it is possible that intratype genetic variation of the E6-coding region may possess more functional significance in the pathogenicity of HPV 16 than sequence variation of the regulatory region.

Acknowledgements

This study was supported in part by research grants from the Medical Research Council of the Academy of Finland, the Finnish Cancer Foundation, the Emil Aaltonen Foundation, the Research and Science Foundation of Farmos, and the University of Turku.

References

1. zur Hausen H. Human papillomaviruses in the pathogenesis of anogenital cancer. *Virology* 1991, **184**, 9–13.
2. Münger K, Phelps WC, Bubb V, Howley PM, Schlegel R. The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes. *J Virol* 1989, **63**, 4417–4421.
3. Hudson JIB, Bedell MA, McCance Di, Laimins LA. Immortalization and altered differentiation of human keratinocytes *in vitro* by the E6 and E7 open reading frames of human papillomavirus type 18. *J Virol* 1990, **64**, 519–526.
4. Turek LP, Smith EM. The genetic program of genital human papillomaviruses in infection and cancer. *Obstet Gynecol Clin N Am* 1996, **23**, 735–758.
5. Tan S-H, Leong LE-C, Walker PA, Bernard H-U. The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID. *J Virol* 1994, **68**, 6411–6420.
6. Gloss B, Bernard H-U. The E6/E7 promoter of human papillomavirus type 16 is activated in the absence of E2 proteins by a sequence-aberrant Sp1 distal element. *J Virol* 1990, **64**, 5577–5584.
7. Chong T, Apt D, Gloss B, Isa M, Bernard H-U. The enhancer of human papillomavirus type 16: binding sites for the ubiquitous transcription factors oct-1, NFA, TEF-2, NFl, and AP-l participate in epithelial cell-specific transcription. *J Virol* 1991, **65**, 5933–5943.
8. Ishiji T, Lace MJ, Parkkinen S, *et al*. Transcriptional enhancer factor (TEF)-1 and its cell-specific co-activator activate human papillomavirus-16 *E6* and *E7* oncogene transcription in keratinocytes and cervical carcinoma cells. *EMBO J* 1992, **11**, 2271–2281.
9. Mittal R, Tsutsumi K, Pater A, Pater MM. Human papillomavirus type 16 expression in cervical keratinocytes: role of progesterone and glucocorticoid hormones. *Obstet Gynecol* 1993, **81**, 5–12.
10. Kyo S, Inoue M, Nishio Y, *et al*. NF-IL6 represses early gene expression of human papillomavirus type 16 through binding to the noncoding region. *J Virol* 1993, **67**, 1058–1066.
11. May M, Dong X-P, Beyer-Finkler E, Stubenrauch F, Fuchs PG, Pfister H. The E6/E7 promoter of extrachromosomal HPV 16 DNA in cervical cancers escapes from cellular repression by mutation of target sequences for YY1. *EMBO J* 1994, **13**, 1460–1466.
12. Lorincz AT, Reid R, Jenson AB, Greenberg MD, Lancaster W, Kurman RJ. Human papillomavirus infection of the cervix: relative risk associations of 15 common anogenital types. *Obstet Gynecol* 1992, **79**, 328–337.
13. Londesborough P, Ho L, Terry G, Cuzick J, Wheeler C, Singer A. Human papillomavirus genotype as a predictor of persistence and development of high-grade lesions in women with minor cervical abnormalities. *Int J Cancer* 1996, **69**, 364–368.
14. Schwarz E, Freese UK, Gissmann L, *et al*. Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. *Nature* 1985, **314**, 111–114.
15. 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**, 4680–4684.
16. 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**, 2989–2997.
17. Dong X-P, Stubenrauch F, Bayer-Finkler E, Pfister H. Prevalence of deletions of YY1-binding sites in episomal HPV 16 DNA from cervical cancers. *Int J Cancer* 1994, **58**, 803–808.
18. Xi LF, Koutsky LA, Galloway DA, *et al*. Genomic variation of human papillomavirus type 16 and risk for high grade cervical intraepithelial neoplasia. *J Natl Cancer Inst* 1997, **89**, 796–802.
19. Xi LF, Critchlow CW, Wheeler CM, *et al*. Risk of anal carcinoma *in situ* in relation to human papillomavirus type 16 variants. *Cancer Res* 1998, **58**, 3839–3844.
20. Yliskoski M, Saarikoski S, Syrjänen K, Syrjänen S, Castrén O. Cryotherapy and CO_2-laser vaporization in the treatment of cervical and vaginal human papillomavirus (HPV) infections. *Acta Obstet Gynecol Scand* 1989, **68**, 619–625.
21. Kataja V, Syrjänen S, Mäntyjärvi R, Yliskoski M, Saarikoski S, Syrjänen K. Prognostic factors in cervical human papillomavirus infections. *Sex Transm Dis* 1992, **19**, 154–160.
22. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988, **16**, 1215.
23. Xi LF, Demers GW, Kiviat NB, Kuypers J, Beckmann AM, Galloway DA. Sequence variation in the noncoding region of human papillomavirus type 16 detected by single-strand conformation polymorphism analysis. *J Infect Dis* 1993, **168**, 610–617.
24. Seedorf K, Krämmer G, Dürst M, Suhai S, Röwekamp WG. Human papillomavirus type 16 DNA sequence. *Virology* 1985, **145**, 181–185.
25. Das BC, Sharma JK, Gopalakrishna V, Luthra UK. Analysis by polymerase chain reaction of the physical state of human papillomavirus type 16 DNA in cervical preneoplastic and neoplastic lesions. *J Gen Virol* 1992, **73**, 2327–2336.
26. Vernon SD, Unger ER, Miller DL, Lee DR, Reeves WC. Association of human papillomavirus type 16 integration in the *E2* gene with poor disease-free survival from cervical cancer. *Int J Cancer* 1997, **74**, 50–56.
27. Chen Z, Storthz KA, Shillitoe EJ. Mutations in the long control region of human papillomavirus DNA in oral cancer cells, and their functional consequences. *Cancer Res* 1997, **57**, 1614–1619.
28. Yamada T, Manos MM, Peto J, *et al*. Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. *J Virol* 1997, **71**, 2463–2472.
29. Ho L, Chan S-Y, Burk RD, *et al*. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J Virol* 1993, **67**, 6413–6423.
30. Smits HL, Traanberg KF, Krul MRL, *et al*. Identification of a unique group of human papillomavirus type 16 sequence variants among clinical isolates from Barbados. *J Gen Virol* 1994, **75**, 2457–2462.
31. O'Connor M, Chan S-Y, Bernard H-U. Transcription factor binding sites in the long control region of genital HPVs. In The HPV compendium 1995, the HPV database, hhtp://hpv-web.lanl.gov.
32. Romanczuk H, Thierry F, Howley PM. Mutational analysis of *cis* elements involved in E2 modulation of human papillomavirus type 16 P_{97} and type 18 P_{105} promoters. *J Virol* 1990, **64**, 2849–2859.
33. Xi LF, Demers GW, Koutsky LA, *et al*. Analysis of human papillomavirus type 16 variants indicates establishment of persistent infection. *J Infect Dis* 1995, **172**, 747–755.
34. Veress G, Szarka K, Dong X-P, Gergely L, Pfister H. Functional significance of sequence variation in the E2 gene and the long control region of human papillomavirus type 16. *J Gen Virol* 1999, **80**, 1035–1043.
35. Dürst M, Kleinheimz A, Hotz M, Gissmann L. The physical state of human papillomavirus type 16 DNA in benign and malignant genital tumours. *I Gen Virol* 1985, **66**, 1515–1522.
36. Cullen AP, Reid R, Campion M, Lörincz AT. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. *J Virol* 1991, **65**, 606–612.
37. Stöppler MC, Ching K, Stöppler H, Clancy K, Schlegel R, Icenogle J. Natural variants of the human papillomavirus type 16 E6 protein differ in their abilities to alter keratinocyte differentiation and to induce p53 degradation. *J Virol* 1996, **70**, 6987–6993.
38. Zehbe I, Wilander E, Delius H, Tommasino M. Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. *Cancer Res* 1998, **58**, 829–833.