

Identification of HPV Variants

John Cason, Jon Bible, and Christine Mant

Summary

The vast majority of anogenital carcinomas are caused by high-risk human papillomaviruses (HPVs), and among Western nations HPV-16 is usually the most predominant cancer-associated type. As a DNA virus, HPV type 16 has a relatively stable genome that is believed to have co-evolved with its host over the millennia. Nevertheless, among the “wild” populations of HPV-16 that are circulating, a large number of variants have been identified, and these may have considerably different pathogenic potentials. In this chapter, methods for screening and characterizing HPV-16 sequence variants are described. In particular, we describe methods for the identification of variation within the HPV-16 E5 open reading frame and for the detection of the nt 131 A→G mutation of the E6 ORF, using restriction fragment length polymorphism assays. In addition, we describe approaches for DNA sequencing and analysis. Such methods are likely to be of particular interest to those involved in epidemiological investigations of virus transmission and pathogenicity studies.

1. Introduction

Cervical cancer is a major cause of female cancer deaths, with some 450,000 incident cases worldwide (1). It is now clearly established that a subgroup of human papillomaviruses (HPVs) are causally associated with this malignancy and are termed *high-risk* (HR) HPVs (2). In the United Kingdom—and in most Western countries—HPV types 16 and 18 are most frequently detected HR-HPVs in cervical malignancies: in our inner-city location, HR-HPV DNA occurs in about 95% of cervical cancers, and 62% are positive for HPV-16 DNA (3). Because the vast majority of HR-HPV infections do not result in carcinoma (4,5), other factors must be involved in malignant progression. Although co-factors for cervical cancer have been sought, no single convincing co-factor has been identified, and the greatest risk for developing cervical cancer remains persistent infection with a HR-HPV and a high viral load.

For a long time, there has been interest in the phylogeny of HPV-16 variants (6), and some studies have sought an association between HPV-16 variants and cervical neoplasia (7–11). However, these reports are largely based on cross-sectional studies of small patient numbers and a limited clinical spectrum of neoplastic lesions, and rarely include infected women with normal cytology. Additional, more detailed longitudinal studies of the association between HPV-16 variants and disease are required. The identification of HPV-16 variants may also be suited for use in studies of virus transmission for epidemiological purposes and in litigation cases of alleged sexual abuse. In this chapter we describe two restriction fragment length polymorphism (RFLP) assays that can be used to rapidly identify the presence of variations within the E5 and E6 open reading frames (ORFs) of HPV-16 variants in large population studies, and a DNA sequencing strategy to rapidly identify and analyze HPV-16 variants.

2. Materials

2.1. Reference Materials and Clinical Sample Preparation

1. Positive controls: reference isolates of HPV-16 such as pAt-16 (12,13), available from Dr. E. M. DeVilliers, DKFZ, Heidelberg, Germany), and DNA from CaSki or SiHa cells (both are HPV-16 DNA positive; obtainable from the American Type Culture Collection [ATCC] Ltd., Rockwell, MD).
2. Negative controls: an HPV-16 DNA negative cell line (e.g., A431, available from ATCC).
3. Cervical brush smears: typical samples for analysis would include cervical brush smears collected with an Axibrush™ (Colgate Medical Ltd.) from women attending local well-woman centers and gynecological outpatient clinics.
4. Dulbecco's phosphate-buffered saline (PBS).
5. Proteinase K (PK) solution: 0.45% v/v NP-40, 0.45% v/v Tween-20, 60 g/L PK (Roche Ltd., UK).

2.2. Setting Up PCR Reactions

1. DNA polymerase (5 U/μL; Promega).
2. 10X polymerase chain reaction (PCR) buffer (Promega).
3. 25 mM MgCl₂ (Promega).
4. E5 PCR primers (see **Table 1** for sequences and conditions).
5. E6 PCR primers and cycling conditions: The first E6 PCR uses the primers E61A (GAGAACTGCAATGTTTCAGG) and E62A (TGATTA CAGCTGGGTTTCTC: 3) which amplifies a 469-bp fragment of the E6 gene (**Fig. 1A**). The second primer set consists of E61B (CCAAAAGAGAACTGCAATGT) and E62B (AATTTTAGAATAAACTTTAAACATT) (**Fig. 1B**).
6. Molecular biology-grade (MBG) DNase-free water.
7. Premixed dNTPs (Cambio, Ltd.).

Table 1
Example of the E5 Polymerase Chain Reaction Conditions

Reaction buffer (10X stock)	MgCl ₂ (25 mM)	dNTPs (2.5 mM each stock)	Amount of each primer (2.5 mM)	DNA polymerase (5 U/mL)	dH ₂ O	Sample
20 µL	20 µL	16 µL	2 µL	1 µL	119 µL	20 µL
Upstream Primer	Downstream Primer	(×1)	Cycling (×40)	(×1)	Size of in product basepairs	
TACAGGATCC	ATTTAGATC	94°C/15 s	94°C/5min	72°C/	273	
TTATGTAATTA	TATATGACA		55°C/15 s	5min		
AAAAGCGTGC	AATCTTGAT		72°C/10 s			
AT	ACTGC					

A = Adenine C = Cytosine G = Guanine T = Thymidine

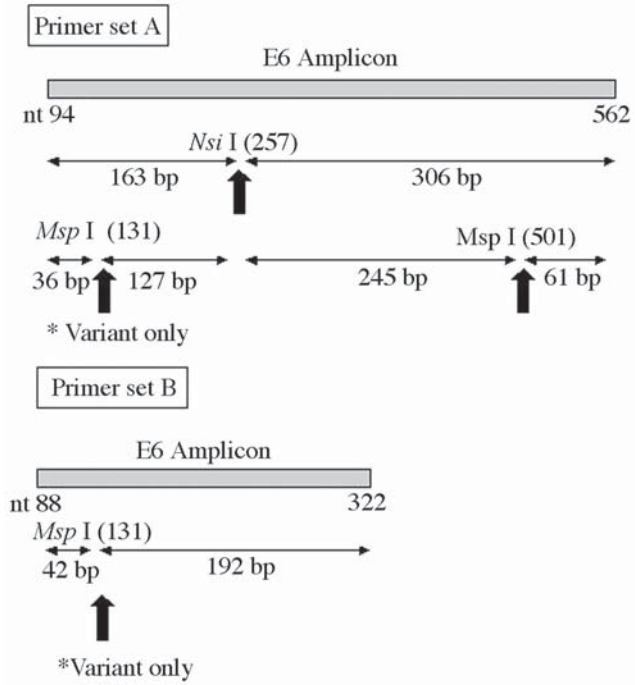


Fig. 1. Restriction fragment length polymorphisms to detect the E6 nt 130 A to G variant.

8. Aerosol-resistant tips.
9. DNase-free plastics (Elkay Laboratory Plastics, Ltd., UK).

2.3. RFLP

1. Phenol:chloroform:isoamylalcohol (50%:48%:2% v/v).
2. Absolute ethanol, and 70% v/v aqueous ethanol.
3. Restriction endonucleases and buffers: *Xcm*1, *Ssp*1, *Nsp*1, *Nsi*1, and *Msp*1.
4. Bovine serum albumin (Sigma).

2.4. Agarose Gel Electrophoresis

1. 10X Orange G loading buffer: 30% (w/v) Ficoll (Sigma), 250 mM ethylenediamine tetraacetic acid (EDTA; disodium salt), 0.25% (w/v) Orange G (BDH, Ltd.).
2. Agarose, UltraPure, electrophoresis grade (Invitrogen).
3. Tris-borate EDTA (TBE) buffer: 0.9 M Tris, 0.9 M boric acid, 2 mM EDTA.
4. Ethidium bromide: 10 mg/mL stock solution.
5. Molecular-weight size marker (e.g., 1 kb Ladder).

2.5. DNA Sequencing

1. Qiagen™ columns for purifying amplicons.
2. *pGem* (Promega).
3. *Escherichia coli* JM109 cells (Promega).
4. Commercially obtainable T7 and SP6 primers (Sequenase™ kit, Pharmacia, Ltd.).

3. Methods

3.1. Preparation of Clinical Samples for PCR

1. Use PBS to resuspend cells from cell-lines or cervical brush smears in four 1-mL aliquots for PCR.
2. Centrifuge a 1-mL sample at 10,000g, resuspend the cell pellet in 200 µL PK solution, and incubate overnight at 55°C.
3. Inactivate the PK (by heating the sample to 90°C for 10 min), then store at -70°C.

3.2. Polymerase Chain Reactions

3.2.1. General PCR Considerations

1. Stringent precautions must be taken to prevent false-positives as a result of contamination, as described in Chapter 23, **Subheading 3.1.**

3.2.2. E5 PCR

1. Use a 20-µL aliquot of the PK-treated cell suspension directly in a 200-µL volume PCR using a proofreading DNA polymerase to amplify nt 3866 to 4077 of the HPV-16 genome (*see Note 1, refs. 14,15, and Fig. 2*).
2. For each batch of 20 clinical samples, two negative controls (molecular biology-grade water and A431 cells) and two positive controls (either pAt-16 plus CaSki

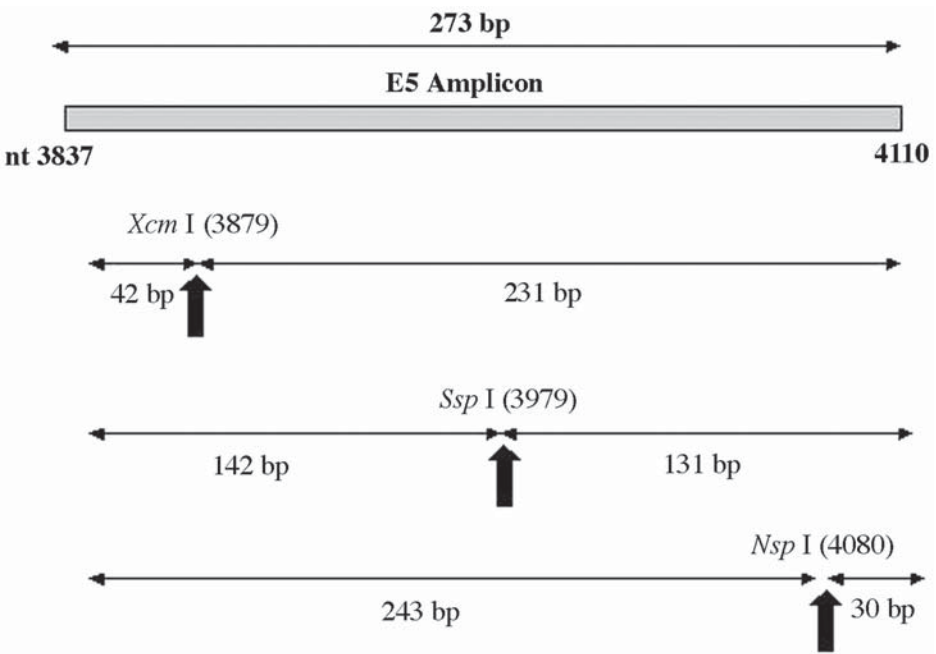


Fig. 2. Restriction polymorphism fragment analysis of human papillomavirus (HPV) type 16 E5.

or pAt-16 plus SiHa DNA) should be included. The PCR is performed using the conditions shown as a guideline (**Table 1**); if problems are encountered, re-optimize conditions (*see Note 2*).

3. After PCR amplification, add 100 μ L of MBG dH₂O and then extract the DNA using phenol:chloroform. Precipitate the DNA with 1 mL of absolute ethanol, wash with 70% v/v aqueous ethanol, and then dry under vacuum for 10 min. Finally, resuspend the pellets in 20 μ L of sterile MBG dH₂O and store at -20°C prior to restriction digests.

3.2.3. E5 RFLP

The E5 PCR amplifies HPV-16 E5 wild-type gene between nucleotides (nt) 3866 to 4077 and has been used previously to identify particular variants most commonly associated with cervical disease (**16**). For the reference HPV-16 sequence, this region contains over 45 restriction endonuclease (RE) cleavage sites, at least three of which (*Xcm* I [³⁸⁷²CCANN NNN↓NNNNTGG³⁸⁸⁶; where *N* = any nt]; *Ssp* I [nt ³⁹⁷⁸AAT↓ATT³⁹⁸³]; and, *Nsp* I [⁴⁰⁷⁷ACATG↓C⁴⁰⁸²]) are disrupted in different reported HPV-16 variants (**6**) (**Fig. 2**). These RE can therefore be used in an RFLP assay to identify eight HPV-16 variants (*see Table 2* and **Note 3**).

Table 2
Designation of E5
Restriction Fragment Length Polymorphism Patterns

Variant	<i>XcmI</i>	<i>SspI</i>	<i>NspI</i>
1	+	+	+
2	+	–	+
3	–	+	+
4	+	+	–
5	–	–	+
6	+	–	–
7	–	+	–
8	–	–	–

1. Resuspend DNA pellets in 50 µL of MBG water.
2. Prepare four 5-µL aliquots for each sample, three of which are subjected to an overnight digestion at 37°C with 5 U of one RE (*SspI*, *XcmI*, or *NspI*).
3. Separate the products by agarose gel electrophoresis (**Subheading 3.2.4.**).
4. Spiking experiments indicate that mixed infections can be identified in the RFLP assay only when the least frequent variant exceeds 10% (w/w) of the total HPV-16 E5 DNA present (**Fig. 3**). In our experience, all eight possible RFLP patterns have been detected among samples from patients, with RFLP patterns 1 (28%), 2 (32%), and 9 (36%) being most common among cytologically normal women infected with HPV 16 (**16**). In HPV-16-positive women with cervical neoplasia, pattern 2 was present in 63% of cases and pattern 1 in 21%.

3.2.4. Agarose Gel Electrophoresis of PCR Products/RFLP Digests

1. Prepare a 2% (w/v) agarose gel using TBE buffer and 5 µL ethidium bromide solution per 100 mL agarose gel.
2. Mix PCR amplicons or RFLP digests with 10X Orange G loading buffer and electrophorese at 125 V for approx 1 h. Ensure molecular-weight standards are run in parallel with the molecular-weight markers analyzed on the gel to assess the size of PCR products.
3. Visualize bands on a ultraviolet transilluminator and compare amplicon size against molecular-weight marker. Photographic records should be obtained.

3.3. RFLP Detection of A→G Variation at Position 131 Within the E6 Open Reading Frame

In this section we describe two RFLP assays to detect the E6 variant (A→G at position 131) described by Ellis et al. (**17**), which may be highly associated with women with high-grade CIN and with human leukocyte antigen B7 gene.

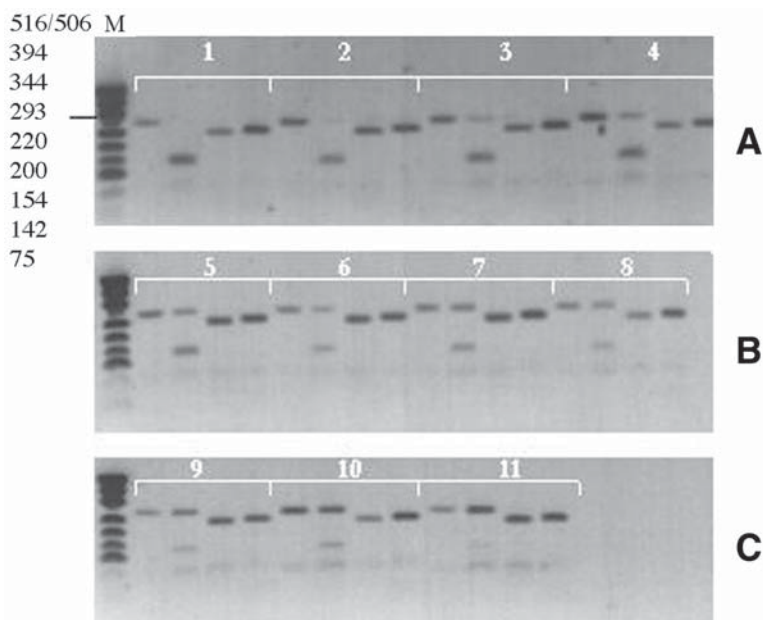


Fig. 3. Sensitivity of the human papillomavirus (HPV) type 16 restriction fragment length polymorphism to mixtures of variants. For this spiking experiment, each group of four lanes corresponds to (left to right) undigested amplicon, *Ssp*I, digest, *Xcm*I digest, and *Nsp*I digest. A1: prototypic HPV16 pAt-16. A2-C10: mixtures of prototypic and HPV16 variants in different ratios from 90% prototypic/10% variant (A2) through to 10% prototypic/90% variant (C10). C11: 100% variant. M: molecular weight marker (sizes in bp).

3.3.1. E6 PCRs

The first E6 PCR uses the primers E61A and E62A (**18**); this amplifies a 469-bp fragment of the E6 gene (**Fig. 1A**). The second primer set consists of E61B and E62B, which amplifies a 234-bp fragment of E6 (**Fig. 1B**).

3.3.2. E6 RFLP

1. Following PCR amplification with primer set A, digest 5 μ L of product with 10 U *Nsi*I (see **Note 3**) in a final dilution of 20 μ L of buffer at 37°C overnight.
2. Further digest 10 μ L of this reaction with *Msp*-I under the same conditions except for the addition of bovine serum albumin at 0.2 g/L.
3. PCR products produced using primer set B are digested with *Msp*I alone.
4. Separate products on a 1.5% agarose gel containing ethidium bromide, and then visualize by trans-illumination with ultraviolet light.

5. Variants can be identified by the presence of the unique *MspI* site. Thus, in the case of the PCR using primer set A, the 469-bp product is digested with *NsiI* to produce fragments of 163 and 306 bp. These products are then digested with *MspI*, which cuts the 306-bp fragment into 245- and 61-bp products and cuts the 163-bp fragment into 127 and 36 bp only when the variant is present. The variant can be seen after electrophoresis, where the 163-bp fragment of the prototypic sequence is replaced by a 127-bp fragment, which indicates the variant is present. In the case of amplicons produced using primer set B, *MspI* digestion results in a band of 234 bp when the prototypic sequence is present, or a band of 192 bp when the nt 131 E6 variant is present.

3.3.3. DNA Sequencing

There are two main approaches for DNA sequencing analysis. First, one can perform direct/bulk sequencing of PCR amplicons, which will provide a consensus of any HPV-16 sequences present; second, a more accurate approach is to clone the PCR products into a plasmid and then sequence >20 clones.

1. Purify E5 PCR amplicons using Qiagen columns and blunt end clone into *pGem*.
2. Add transformed plasmids into *E. coli* JM109 cells and grow overnight at 37°C on agar plates.
3. Select white colonies, grow midi-cultures, and purify plasmid DNA using Qiagen columns.
4. Sequence the inserts in both orientations (with T7 and SP6 primers that recognize sequences in *pGem* which flank the inserted E5 DNA) using Sequenase kit.
5. Analyze resulting E5 products representing HPV-16 E5 DNA sequence between nucleotides 3866 and 4077 on a DNA sequencer.

3.3.4. DNA Sequence Analyses

A variety of free-to-use software is available on the net for DNA sequence analyses (**Table 3**). DNA sequence data can be easily arranged into a FASTA format:

>DNA sequence title (hard return)

ACCGGGGGGGGTGCTCAG . . . (but containing no “hard returns”)

This file can be saved as a normal text “.doc” file and manipulated for sequence editing (e.g., use the “Find” option in the “Edit” function of Microsoft Word to rapidly identify the primer sites, and the “Replace” function to remove hard returns). Such FASTA files can then be cut and pasted into an alignment program such as CLUSTAL-W (<http://www.ebi.ac.uk/clustalw/index.html> or <http://clustalw.genome.ad.jp/>). These programs will also permit “gap stripping” so that only like sequences of DNA are compared, the production of homology measurements, bootstrap analyses (which indicate how many times a given

Table 3
Useful Web Sites for DNA Analyses

http://cti.itc.virginia.edu/~cmg/Demo/wheel/wheelApp.html
http://www.ncbi.nlm.nih.gov/BLAST/
http://searchlauncher.bcm.tmc.edu/multi-align/multi-align.html
http://ca.expasy.org/
http://www.ddbj.nig.ac.jp/E-mail/clustalw-e.html
http://www.ualberta.ca/%7Estothard/javascript/rev_comp.html
http://pbil.univ-lyon1.fr/alignment.html
http://www.cbs.dtu.dk/databases/PhosphoBase/predict/predict.html
http://www.ch.embnet.org/software/TCoffee.html
http://www.ebi.ac.uk/clustalw/index.html
http://www.ebi.ac.uk/services/
http://clustalw.genome.ad.jp/

branch occurs when trees are produced with a random input of the sequence files), and the drawing of phylogenetic trees. The trees can then be pasted into TreeView (<http://taxonomy.zoology.gla.ac.uk/rod/treeview.html>) in order to obtain phylogenetic trees that will indicate the amount of variation.

4. Notes

1. Phenol/chloroform extraction of target DNA often results in significant loss of target material. For this reason, we habitually do not use this method and just treat samples with proteinase K. To check that no significant carry-over of potential PCR inhibitors has occurred, all samples should also be tested by PCR for a “housekeeping” gene such as β -globin.
2. For all PCRs, it is recommended that “checkerboard” analyses (titrations of differing concentrations of different PCR components) be performed to optimize the PCRs by checking the optimal concentrations of primers, dNTPs, and magnesium ions.
3. There can be considerable variation in the efficacy of purportedly the same enzymes obtained from different suppliers; however, in our experience NEB, Ltd., produces highly active RE. Prior to any RFLP analyses of clinical samples, the activities of any RE should be determined by digesting a known section of DNA containing the appropriate RE cut sites.

References

1. zur Hausen, H. (2002) Papillomaviruses and cancer: from basic studies to clinical application. *Nat. Rev. Cancer* **2**, 342–350.
2. zur Hausen, H. (2001) Oncogenic DNA viruses. *Oncogene* **20**, 7820–7823.
3. Cavuslu, S., Goodlad, J., Connor, A., et al. (1997) Relationship between human papillomavirus infection and overexpression of p53 protein in cervical carcinomas and lymph node metastases. *J. Med. Virol.* **53**, 111–117.

4. IARC (1995) Epidemiology of infection. In, *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Human Papillomaviruses*, Lyon: International Agency for Research on Cancer, World Health Organisation, pp. 60–65.
5. Conrad-Stoppler, M., Ching, K., Stoppler, H., Clancy, K., Schlegel, R., and Icenogle, J. (1996) 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.* **70**, 6987–6993.
6. Chan S-Y., Ho, L., Ong, C-K., et al. (1992) Molecular variants of human papillomavirus type 16 from four continents suggest ancient pandemic spread of the virus and its coevolution with mankind. *J. Virol.* **66**, 2057–2066.
7. Fujinaga, Y., Okazawa, K., Nishikawa, A., et al. (1994) Sequence variation of human papillomavirus type 16 E7 in preinvasive and invasive cervical neoplasias. *Virus Genes* **9**, 85–92.
8. Hecht, J. L., Kadish, A. S., Jiang, G., and Burk, R. D. (1995) Genetic characterization of the human papillomavirus (HPV) 18 E2 in clinical specimens suggests the presence of a subtype with decreased oncogenic potential. *Int. J. Cancer* **60**, 369–376.
9. Londesborough, P., Ho, L., Terry, G., Cuzick, J., Wheeler, C., and Singer, A. (1996) Human papillomavirus genotype as a predictor of persistence and development of high-grade lesions in women with minor cervical abnormalities. *Int. J. Cancer* **69**, 364–368.
10. Xi, L-F., Koutsky, L. A., Galloway, D. A., et al. (1997) Genomic variation of human papillomavirus type 16 and risk for high grade cervical intraepithelial neoplasia. *J. Natl. Cancer Inst.* **89**, 796–802.
11. Zehbe, I., Wilander, E., Delius, H., and Tommasino, M. (1998) Human papillomavirus 16 E6 variants are more prevalent in cervical carcinoma than the prototype. *Cancer Res.* **58**, 829–833.
12. Seedorf, K., Kraemmer, G., Duerst, M., Suhai, S., and Rowekamp, W. G. (1985) Human papillomavirus type 16 DNA sequence. *Virology* **145**, 181–185.
13. Halbert, C. L. and Galloway, D. A. (1998) Identification of the E5 open reading frame of human papillomavirus type 16. *J. Virol.* **62**, 1071–1075.
14. Cavuslu, S., Starkey, W. G., Kaye, J. N., et al. (1996) Detection of human papillomavirus type-16 (HPV-16) DNA utilising microtitre-plate based amplification reactions and a solid-phase enzyme-immunoassay detection system. *J. Virol. Methods* **58**, 59–69.
15. Mant, C., Kell, B., Best, J. M., and Cason, J. (1997) Polymerase chain reaction protocols for the detection of DNA from mucosal human papillomavirus types - 6, -11, -16, -18, -31 & -33. *J. Virol. Methods* **66**, 169–178.
16. Bible, J. M., Mant, C., Best, J. M., et al. (2000) Cervical lesions are associated with human papillomavirus type 16 intratypic variants that have high transcriptional activity and increased usage of common mammalian codons. *J. Gen. Virol.* **81**, 1517–1527.

17. Ellis, J. R. M., Keating, P. J., Baird, J., et al. (1995) The association of an HPV 16 oncogene variant with HLA B7 has implications for vaccine design in cervical cancer. *Nat. Med.* **1**, 464–470.
18. Luxton, J., Mant, C., Greenwood, B., et al. (2000) HPV16 E6 oncogene variants in women with cervical intraepithelial neoplasia. *J. Med. Virol.* **60**, 337–341.

Human Papillomaviruses

Methods and Protocols

Davy, C.; Doorbar, J. (Eds.)

2006, XVI, 496 p. 153 illus., Hardcover

ISBN: 978-1-58829-373-2

A product of Humana Press