

Group F Sequences

HPV2a	HPV3
HPV7	HPV10
HPV27	HPV28
HPV29	HPV32
HPV40	HPV42
HPV43	HPV54
HPV57	HPVCP8061

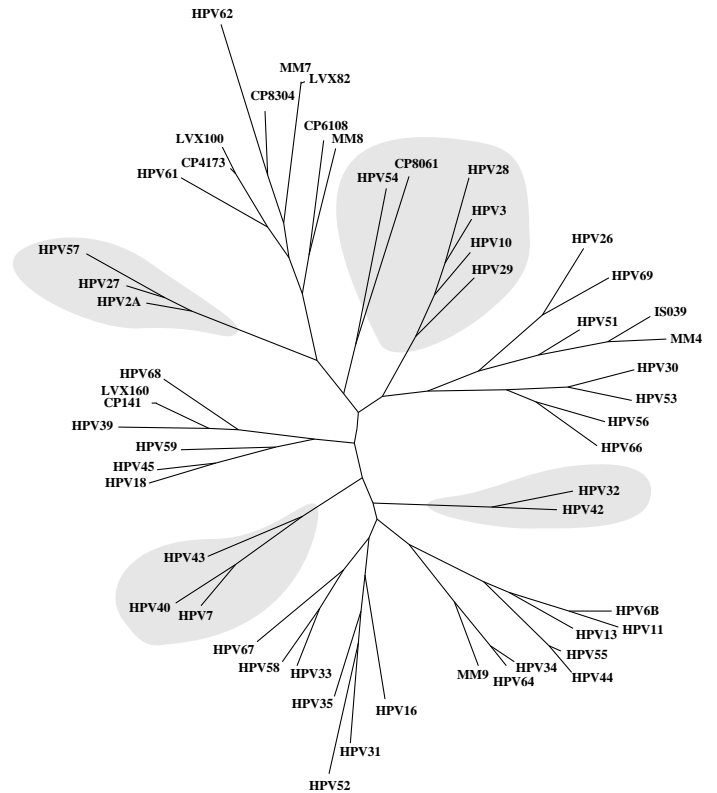
INTRODUCTION

Group F consists of the human papillomavirus types HPV-2a, HPV-3, HPV-7, HPV-10, HPV-27, HPV-28, HPV-29, HPV-32, HPV-40, HPV-42, HPV-43, HPV-54, HPV-57, and the novel virus CP8061. This group is intended to be a collection of types which do not tightly cluster with any of the other major branches in phylogenetic analysis. However, this group does form five small clusters, clusters a - e, which display distinctive clinical and phylogenetic characteristics.

Cluster a: Viruses HPV-2a, HPV-27, and HPV-57 form cluster a. These viruses infect both cutaneous and mucosal tissue. HPV-57 seems to preferentially infect the mucosa, and only rarely infects cutaneous tissues. Conversely, the primary site of infection for HPV-2a is cutaneous tissue. Due to the rare detection of HPV-27, its tissue preference is unknown [1]. HPV-57 has been isolated from a number of oral, upper respiratory and genital lesions, but in only two cases of verrucae vulgares of immunosuppressed patients [2]. HPV-2a frequently causes common and filiform warts, infecting EV patients, immunosuppressed patients and the general population [3]. Hybridization under conditions of low stringency to a tongue carcinoma and cross-hybridization to HPV-18 indicated dual tissue infectivity [4]. Subsequently, HPV-2 has been detected in numerous oral lesions of differing severity and in an anogenital wart [1]. Recently a subtype of HPV-2, HPV-2c, was sequenced over the LCR, E6, and L1 region. It was identical to the sequence of HPV-27. HPV-2c and HPV-2a differ by 55 nucleotides, but qualify as “close types [5]. HPV-27 was isolated from a common hand wart from a renal transplant patient and subsequently from cases of anogenital lesions [1,6].

Cluster b: Viruses HPV-3, HPV-10, HPV-28 and HPV-29, which form cluster b are associated with cutaneous benign lesions. HPV-3, HPV-10, and HPV-28 commonly cause benign flat warts in patients with EV, immunosuppressed patients, and in the general population [1,7,8]. HPV-28 has also been found in intermediate warts, a type which exhibits characteristics of both flat and common warts [8]. HPV-29 was isolated from a skin wart. Prevalence data indicates that HPV-29 occurs very rarely [9].

Cluster c: Viruses HPV-7, HPV-40 and HPV-43, which form cluster c are predominantly associated with benign lesions. HPV-40 and HPV-43 infect the genital tract. HPV-7 infects both cutaneous and oral tissue, both with significant frequency. HPV-7 was first isolated from Butcher’s warts, the benign hand warts of meat handlers. Initially, it was almost exclusively detected in individuals with animal contact which suggested the possibility of cross-species transmission. This hypothesis was refuted when 37 bovine tumors were screened for the presence of HPV-7 DNA, and not one positive hybridization was reported [10]. Additionally, HPV-7 was frequently detected in common warts of individuals without known animal contact. Unexpectedly, HPV-7 has been detected in a significant proportion of oral papillomas (9%) and in one case of a tonsillar carcinoma [11,12]. HPV-40 is a rare HPV of the genital tract. It has been detected in benign lesions, low-grade



neoplasias and in bowenoid lesions [1,2]. HPV-43 was classified by Lorincz et al. as a “low-risk” virus [13]. It was originally isolated from vulvar tissue [14].

Cluster d: Viruses HPV-32 and HPV-42 are primarily associated with benign orogenital lesions. HPV-32 was originally isolated from an oral focal epithelial lesion and further detected in several oral papillomas [1]. When 113 benign and malignant tumors of the oro-respiratory system were analyzed for the presence of HPV-32 DNA, 9% of the oral papillomas were positive for HPV-32 [11]. In contrast, HPV-42 is predominantly found in genital lesions. Lorincz et al. have classified HPV-42 as a “low-risk” virus [14]. This placement is justified by the infrequent association of HPV-42 with dysplastic lesions and the absence of HPV-42 in malignant lesions [15,16].

Cluster e: Viruses HPV-54 and the novel virus CP8061 form the cluster e. Both viruses primarily infect genital mucosa. HPV-54 was isolated from a penile Buschke-Lowenstein tumor in conjunction with HPV-6 DNA. Initial prevalence data indicates that it is a rare genital HPV type [17]. CP8061 was isolated from a cervical lavage sample obtained through clinical studies conducted in the state of New Mexico among a tri-ethnic population [18].

Of the members of Group F, complete genomic sequences are available for all except HPV-28, HPV-29, HPV-43, HPV-54 and HPVCP8061; these have all been sequenced only over the My09-My11 region of L1, excepting HPV-43 which has also been sequenced over E6. Within the group, there are several sets of sequences that qualify as “close types”—sequences which are seen as distinct types under the criterion of ten percent dissimilarity at the nucleotide level, but between which most of these changes are in fact “silent”, causing no difference at the amino acid level (Part III). These include HPV-40 and HPV-7, and the sequence cluster comprising HPV-2a, HPV-27 and HPV-57.

-
- [1] de Villiers,E.M. Human pathogenic papillomavirus types: an update. in *Human pathogenic papillomaviruses*, edited by Harald zur Hausen, Springer-Verlag, Heidelberg, pp 1–12 (1994)
 - [2] de Villiers,E.M., Hirsch-Behnam,A., von Knebel-Doeberitz,C., Neumann,C., and zur Hausen, H. Two newly identified human papillomavirus types (HPV 40 and 57) isolated from mucosal lesions. *Virology* **171**: 248–53 (1989)
 - [3] Corley,E., Pueyo,S., Goc,B., Diaz,A., and Zorzopulos,J. Papillomaviruses in human skin warts and their incidence in an Argentine population. *Diagn Microbiol Infect Dis* **10**: 93–101 (1988)
 - [4] Hirsch-Behnam,A., Delius,H., and de Villiers,E.M. A comparative sequence analysis of two human papillomavirus (HPV) types 2a and 57. *Virus Res* **18**: 81–97 (1990)
 - [5] Chan,S.Y., Tan,C.H., Delius,H. and Bernard,H.U. Human papillomavirus type 2c is identical to human papillomavirus type 27. *Virology* **201**: 397–8 (1994)
 - [6] Ostrow,R.S., Zachow,K.R., Shaver,M.K., and Faras,A.J. Human papillomavirus type 27: detection of a novel human papillomavirus in common warts of a renal transplant recipient. *J Virol* **63**: 4904 (1989)
 - [7] Kremsdorf,D., Jablonska,S., Favre,M., and Orth,G. Human papillomaviruses associated with epidermodysplasia verruciformis. II. Molecular cloning and biochemical characterization of human papillomavirus 3a, 8, 10, and 12 genomes. *J Virol* **48**: 340–51 (1983)
 - [8] Favre,M., Obalek,S., Jablonska,S., and Orth,G. Human papillomavirus type 28 (HPV-28), an HPV-3-related type associated with skin warts. *J Virol* **63**: 4905 (1989)
 - [9] Favre,M., Croissant,O., Orth,G. Human papillomavirus type 29 (HPV-29), an HPV type cross-hybridizing with HPV-2 and with HPV-3-related types. *J Virol* **63**: 4906 (1989)
 - [10] Oltersdorf,T., Campo,M.S., Favre,M., Dartmann,K., and Gissmann,L. Molecular cloning and characterization of human papillomavirus type 7 DNA. *Virology* **149**: 247–50 (1986)
 - [11] de Villiers,E.M., Weidauer,H., Le,J.Y., Neumann,C., and zur Hausen,H. Papilloma viruses in benign and malignant tumors of the mouth and upper respiratory tract. *Laryngol Rhinol Otol (Stuttg)* **65**: 177–9 (1986)

- [12] Snijders,P.J.F., van den Brule, A.J.C., Meijer,C.J.L.M., and Walboomers, J.M.M. Papillomaviruses and cancer of the upper digestive and respiratory tracts. in *Human Papillomaviruses*, edited by Harald zur Hausen, Springer-Verlag, Heidelberg, 1994, pp 177–197
- [13] Lorincz,A.T., Reid,R., Jenson,A.B., Greenberg,M.D., Lancaster,WD, and Kurman,R.J. Human papillomavirus infection of the cervix: relative risk associations of 15 common anogenital types. *Obestet Gynecol* **79**: 328–337
- [14] Lorincz,A.T, Quinn,A.P., Goldsborough,M.D., Schmidt,B.J., and Temple, G.F. Cloning and partial DNA sequencing of two new human papillomavirus types associated with condylomas and low-grade cervical neoplasia. *J Virol* **63**: 2829–34 (1989)
- [15] Beaudenon,S., Kremsdorf,D., Obalek,S., Jablonska,S., Pehau-Arnaudet,G. Croissant,O., and Orth, G. Plurality of genital human papillomaviruses: characterization of two new types with distinct biological properties. *Virology* **161**: 374–84 (1987)
- [16] Philipp,W., Honore,N., Sapp,M., Cole,S.T., and Streeck,R.E. Human papillomavirus type 42: new sequences, conserved genome organization. *Virology* **186**: 331–4 (1992)
- [17] Favre,M., Kremsdorf,D., Jablonska,S., Obalek,S., Pehau-Arnaudet, G., Croissant, O., and Orth,G. Two new human papillomavirus types (HPV54 and 55) characterized from genital tumours illustrate the plurality of genital HPVs. *Int J Cancer* **45**: 40–46 (1990)
- [18] Peyton,C.L. and Wheeler,C.M. Identification of five novel human papillomaviruses in the New Mexico triethnic population. *J. Infect. Dis.* (1994) In press

HPV2a

LOCUS HPV2a 7860 bp ds-DNA VRL 21-JUN-1991
 DEFINITION Human papillomavirus type 2a (HPV-2a), complete genome.
 ACCESSION X55964
 SOURCE Human papillomavirus type 2a DNA.
 REFERENCE 1 (bases 1 to 7860)
 AUTHORS Delius,H.
 JOURNAL Unpublished (1990)
 COMMENT HPV-2a was originally isolated from a verruca vulgaris lesion, and was thereafter detected quite frequently in similar types of lesions. It was thought for some time to be strictly a "cutaneous HPV" but was later discovered in a tongue carcinoma and in other mucosal lesions, thus demonstrating some ambiguity in its tissue specificity. Its preferential tropism, however, seems to be toward cutaneous tissue. It shows strong similarity to HPV-27 and HPV-57, and a much weaker similarity to HPV-3. Like the latter, it is often present in the cutaneous lesions of epidermodysplasia verruciformis patients, as well as those of immunosuppressed patients (Hirsch-Behnam et al. Virus Res. 18, 81-98).

The ORFs of the HPV-2a genome are generally homologous to those of all other sequenced papillomaviruses. Hirsch-Behnam et al. report that none of its ORFs show any definite homology to the E5 ORFs of other papillomaviruses, but propose several possible candidates in the region following E2 and preceding L2. In addition, they note the presence of the following potential regulatory elements: polyadenylation signals following both the early and late sets of genes; a number of direct repeats both in the LCR and in the non-coding region between the E2 and L2 genes; a palindromic sequence in the LCR from nt 7720 to nt 7754; E2-binding sites located in the LCR; NF-1 binding sites located in the LCR on both strands of the genome. They also note that the glucocorticoid response element, found in the LCRs of several other papillomaviruses, is absent in HPV-2a.

BASE COUNT 2010 a 1788 c 2016 g 2046 t
 ORIGIN 88 bp upstream from beginning of E6 cds
 1 ATAAtgtata actataatcc ttattttaa aataggggtgt gACCGAAAAC GGTcagACCG
 E6 orf start -> -> E2-bind -> E2-bind
 61 AATTCGGTtg TATATAAaca gaagcaggAT Gcacacaagg gcagggatgt ctgaggagaa
 signal -> E6 cds ->
 121 tccatgccct aggaacatct ttttgctttg caaagagtat ggtttgagc tagaggattt
 181 gcgattgctc tgtgtatggt gcaaacggcc gttatcagag gctgacatat gggcatttgc
 241 aataaaagaa ctgtttgtag tgtggagaaa gggctccca tttggagcct gcgaaaaatg
 301 cctgattgca gcaggaaaac ttagacaata cagacattgg cattactcat gctacggaga
 361 cacagtggag actgagacag gaatacccat ACCTCAGCTG TTtatgagat gctatatttg
 -> E2-bind
 421 ccaTAAGccc ctgagctggg aggagaagga ggcattacta gttggaaaca agcgtttcca
 E7 orf start ->
 481 caacatatca ggccgggtgga cgggacattg catgaactgc gggatcatcAT Gcaccggaac
 E7 cds ->
 541 cgaccagcc tcaaggacat tacacTAATA ttggatgaaa taccgaaat tgttgaccta
 <- E6 end
 601 cattgcgagc agcaatttga cagctcagaa gaagagaata accatcaact gacagaacca
 661 gatgtgcagg cctacggggt ggtaactacc tgctgtaagt gtggcagaac cgtccggctg
 721 gtggttagt gcggacaagc agacctaaga gagctggaac agctgttctt gaagacgctg
 781 actcTAGtgt gccctcactg gcctTAGcgt tATGgaggat tccgaaggta ccgacgggac
 E1 orf start -> E1 cds ->
 <- E7 end
 841 cgaggaggac gggtgccggg caggggggtg gtttcatgtg gaggccatta taacacacgg
 901 ccagaggcag gtatccagtg acgaggacga ggacgaaaca gagacagggg aggatttaga
 961 ctttatagac aataggggttc ccggagatgg gcaggaaatt cccttgagc tatatgcaca
 1021 acaaacggct caggatgacg aagcaacagt gcaggcccta aaacgaaagt ttgtggccag
 1081 tcctttgtct gcatgctcat gcatagagaa tgatttaagt ccagattag atgcaatctc
 1141 cctaacaga aagtcagaaa aggcgaagag gcgcttattc gagacagaac caccagacag
 1201 tgggtatggc aatacgcaga tggttgttgg aacgccagag gaggtaacgg gggatgagga

```

1261 aagccaaggg gggcggccgg tggaggatca ggaggaggag cgtcaagggg gagacggaga
1321 ggcagatcta actgtacaca ctccacagtc aggaacagat gcggcgggta gcgtgctgac
1381 ctactaaga agtagcaatc tgaaggcgac gttgctgagt aagtttaagg acctgtttgg
1441 ggtgggattc tatgaactgg tcagacagtt caaaagcagc aagacagcat gtgcagactg
1501 ggtcgtctgc gcctatgggtg tgtattatgc tgtagcggag ggtctaaaga aattaataca
1561 gccacataca caatatgcac atatacaggt acagaccagc tegtggggca tgggtgctott
1621 tatgctctg cagatacaact gtgcaaaaaa cagggaactca gtgtccaaga acatgagcat
1681 gctgctaaac attcccgaaa agcatatgct catagaacca ccaaaactga gaagtacccc
1741 tgccgcctta tactggtaca agacggccat gggcaacgga agtgaggtat atggggaaac
1801 accagaatgg attgttagac agacgttggg aggacatagc atggaagacg aacagttcag
1861 actgtcagtt atggtacagt atgcatatga ccatgacatt gtagaggaaa gtgtgcttgc
1921 atttgagat gcacaactag cagatgtgga tgccaatgca gcagcatttc taacacgtaa
1981 ctgtcaggcc aagtagctga aggacgcagt gacaatgtgc aggcactata agcgtgcaga
2041 gagagaacag atgagtatgt cacagtggat aacattcaga ggaaataagg tatcagagga
2101 aggggactgg aagcccatag tcaggtttct aagacatcaa ggggtagagt ttgtgctggt
2161 cctagctgcc tttaaattgt tcctaaaagg cgtgccaaag aaaaattgta tagtgttcta
2221 tggacctgca gacacaggca aatcatatct ttgcatgagc ttgttgtagt tcctaggcgg
2281 cgctgttatc tcatatgcta attctagcag ccatttttgg cttcaacctt tatcagatag
2341 taagataggg ttactggacg acgcaacacc ccagtgttgg agttacatag atatatattt
2401 aagaaatctt ttggatggac acccagtgag catagacaga aagcacaana ctttctgca
2461 gcttaagtgt ccaccctaa tgataacaac caacaccaat cctctagagg aggacagatg
2521 gaaatatttg cgcagcaggc tgacagtgtt tacatttaag aatcatttc cttttgcaag
2581 tccgggagag cccctgtacc cgataaataa tgcaaaactgg aatgctttt tccaaaggtc
2641 gtggctccgc ttagaccTAA acagtccaga ggagcaggac gacaATGgaa acactggcga
E2 orf start -> E2 cds ->
2701 accgtttaga tgcgtgccag gagacgttgc tagaactgta TGAaaaggat agcaacaaac
<- E1 end
2761 ttgaggatca gattaagcat tgggcgcagg tccggctaga aaatgcatg ctgtttaagg
2821 cccgagaatg tggaatgaca cgagtcggct gtacagctgt gcctgcctc accgtgtcaa
2881 aagctaaggc atgtcaggcc atagaggtag agctggcatt acagacattg atgcagagtg
2941 cctatagcac ggaggcatgg accctacgag acacgtgtct ggagatgtgg gacgcacctc
3001 caaagaaatg ctggaaaaaa aaaggacaat cagtattagt gaaattgat ggcagcagtg
3061 acagagacat gatataatac agctggggat tcatttatgt gcaggacact atcactgatt
3121 cctggcataa ggtgccaggg caggtggacg aactgggatt atattatgtg cacgatggtg
3181 tacgtgttaa ctatgtggac tttggaacag agtcctTGAc ctATGgggtc accgggacgt
E4 cds ->
E4 orf start ->
3241 gggaggtgca cgtggctggg actgttattc accatacacc cgcactctgtg tctagcacc
3301 aggccagcgc ctcgagcagc gaaccactat cccctattag aactgctgta tccccagtc
3361 cagccccagt cgcagcctca gcagaatcaa caggagcagg aagagcagct ccgcccacc
3421 aagcgttgtg ctccgcccag gcgccaacga gtccgcccgc caagcgcagc cgtgtcatcg
3481 tccggacagca gcatccccgg cccgactcta cgcgaacggt cggagagggg gaagtggagt
3541 gttacaacaa gcggagcacc agtgactcta accgcacaga ccccaggtgg ggccacggtg
3601 acactgactc tgtgcctgTA Atccactga gaggtgatgc aaattgttta aagtgttca
<- E4 end
3661 gatacagggt gcaaaaaacat aaagacgtac tgtatgccag ggtgtcctcc acgtggcact
3721 gggcgggtgg gaacggtgat aagacagcct ttgtaacct gtggtacacc agcgttgaac
3781 agcgtacaga gttcctgaca agagtcagta tacctaaggg attgatagca ttgccagggt
3841 atatgtctgc atttgtaTAA tcctacatgc ttgtataaac atatggtcca atacatttca
<- E2 end
3901 aggcctgcct ccgcaacaca gccctggact actttctctg cgtggttgca ggggtggacac
3961 atctgcttgt gctactgctc ttctgtggc tetctcaact aacccccctt gtggcCTATC
->
4021 TGGtgttctt tttctgtgtC TATCTGGggc tgtggttgat atatgtgcag gccttttggg
repeat region start
4081 ttttACCATA GTCGTTatta tttcgccata cgttgctgct agcttgata catagtctat
-> E2-bind
4141 ataccattg tgtgagattt gcaatgtacc ctgttggtgta taagggatct gagggaaacat
4201 atcctgtggg actgtggggg catgaTGAtg ttcaATGtct gttgggtgatt cttatcctaa
L2 cds ->
L2 orf start ->
4261 tcgccttttt attggtgatg ttttatgtcc gttgttaaa ccacacctaa cacccccact
4321 tttttatatt gTTTTGATAC ATTTTcaTTT TGATACATTT gtgttttttt tgtatttggc
repeat region end <-

```

HPV2a

4381 gcgttttAAT AAACgtgcaa ccatgtctat acgtgccaag cgtcgaaagc ggcctcccc
signal ->
4441 cacagacctc tatcgtacct gcaagcaggc aggtacctgc ccccagaca ttatcccag
4501 agtggaacag aacactttag cagataaaat ccttaagtgg ggcagtttag gtgtgtttt
4561 tgggggtcta ggtataggca cgggcagcgg cacagggggg cgtactgggt acattcctgt
4621 aggttcgcga ccACCACTG TAGTTgacat tgggtccaacg occagggccg cgtttatcat
-> E2-bind
4681 tgaacctgtg ggggctctg aacctctat tgtcactttg gtggaggact ctagcatcat
4741 taacgcagga gcgtcacatc ccACCTTAC TGGTactggg ggcttcgaag tgacaACCTC
-> E2-bind E2-bind ->
4801 CACCGTTaca gaccccgccg tcttgatata cACCCCTCA GGTaccagtg tgcaggtcag
-> E2-bind
4861 cagcagtagc tttcttaacc cactatacac tgagccagct attgtggagg ctccccaaac
4921 aggggaagta tctggccatg tacttgtagg tacagccacc tcaggggtctc atggctatga
4981 ggaataacca atgcagacgt ttgccacgtc ggggggcagc ggtacagagc ctatcagtag
5041 cacaccctc cctggcgtgc ggagagtgc cggaccccg cgttacagta gagccaatca
5101 gcaagtgcaa gtcagggatc ctgctgttct tgcaaggcct gctgatctag taacatttga
5161 caatcctgtg tatgaccag aggaaactat aatatttcag catccagact tgcagagacc
5221 accggatcct gatTTTTTgg acatagtggc gttgcatcgt cccgcctca cgtccagaag
5281 gggtagctgc cgttttagta ggttgggagc cagggctaca ctccgcACCC GTAGTGGTaa
-> E2-bind
5341 acaaatggg gcacgggtgc acttctatca tgatattagc cctataggtg ctgaggagtt
5401 ggagatggag cactgttgc cccagcttc tactgataac acagatatgt tatatgatgt
5461 ttatgctgat tcggatgccc ttcagccatt gcttgatgag ttaccgcgcg ccctcgcgg
5521 ttcactctct ctggctgaca ctgctgtgct tgccacctcc gcatctacac tacgggggtc
5581 cactactgct cctttatcaa gtggatttga tgtgctgtg tacaccggtc ctgacattga
5641 accaccaat gttcctggca tgggacctct gattcctgtg gctccatcct taccatcgtc
5701 tgtgtacata tttgggggag attattattT GATgccaagt TATGtcttgt ggctaaaacg
L1 orf start -> L1 cds ->
5761 acgtaaactg gtccactatt tctttgcaga tggctttgtg gcgccTAAat gaaagcaagg
<- L2 end
5821 tatacctacc tccaacctt gtttcaaagg tgatcagtag ggatgtctat gtcacgcgga
5881 ctaatgtgta ttaccatggt ggcagttcta ggcttctcac tgtgggtcat ccatattact
5941 ctataaagaa gagtaataat aagggtggctg tgcccagggt atctgggtac caatatcgtg
6001 tatttcacgt gaagttgcca gatccaaata agtttggcct gcccgatgct gatttgtatg
6061 atccagatac ccagagactt ctgtgggctg gcgtgggagt agaggtgggc cgtgggcagc
6121 ctttgggtgt ggggtgtgct ggtcaccat attacaatag actggatgac actgaaaaatg
6181 cacacacacc tgatacagct gatgatggca gggaaaacat ttctatggat tataaacaga
6241 cacagctggt cattctgggc tgcaaacccc ctattggtga gcaactggtc aagggtacca
6301 cctgtaatgg gtcttctgct gctggtagt gccgcacct ccaatttact aacacaacta
6361 ttgaggacgg ggatattggt gaacaggggt tcggtgcctt ggattttgcc actctgcagt
6421 caaataagtc agatgttctt ttggatattt gtaccaatac ctgtaaatat cctgattatc
6481 tgaagatggc tgcaagacct tatggtgatt ctatgttctt ctgcctgctg agggaaacaa
6541 tgttcaactg tcatttttcc aatctgggtg gtaagatggg tgacaccatc ccggatgagt
6601 tatacattaa aagtacctca gttccaactc caggcagtcg tgtttatact tccactccta
6661 gtggctctat ggtgtcctct gaacaacagt tgtttaataa gccttactgg ctaccggagg
6721 cccaagggca caacaatggt atgtgctggg gcaatagggt ctttctgact gtgggtggaca
6781 ccacacgtag cactaatgta tctctgtgtg ccactgaggc gtctgatact aattataagg
6841 ctaccaattt taaggaatat ctcaggcata tggaggaata tgatttgcag ttcactctcc
6901 aactgtgcaa gataaccctt actcctgaaa ttatggccta tatacataat atggatcccc
6961 agttgttaga ggattggaac ttcgggtgac cccctccgcc gtctgccagt ttacaggata
7021 cctatagata tttgcagtcc caggctatta catgtcaaaa acctacacct cctaagacct
7081 ctaccgatcc ctatgcctcc ctgacctttt gggatgtgga tctcagtga agttttcca
7141 tggatctgga ccaatttccc ttgggtcgca agttttgct gcagcggggg gctatgccta
7201 ccgtgtctcg caagcgcgcc gctgtttcgg ggaccacgcc gccactagt aaacgaaaac
7261 gggtaaggcg tTAGctctca gtgtgcacac atttctctg ttctactttt tacatattat
<- L1 end
7321 tttgtgtct gtaatatggt tatgtttgtt ttgtgcttat atTACATGTA TACATGTAtg
-> repeat region start
7381 gtatgtatcc cctcccgtat gAATAAAcgt gtgtcatgtg ttgtgtgttc tgtaactgta
signal ->
7441 cgttctgggt cacagatttc tgcaccccat cgccttgtgt gtagcccca gtttcatgca
7501 ACCGTTTTTCG GTTgcgtgca gTTTCGGTcg gcgcggttgc caaccagct taatccttta
-> E2-bind <- repeat region end

```
7561 attgctctca tcctaaagtg ttatctgtgc cagcgacgat gagtttggat tttggttgtt
7621 taatgctttt tcttttcagt ttttcctttg tttgtgccag gccgcgagag ggcgtgcaca
7681 ttcttaggct gattatctta atgtgtTTGG CAcatccttg tactgcgtct gcagaaaaac
      NF-1 bind ->
7741 ctgcagcaac agcactttgg gcgcgtcgtt tttgcagcca actttcactt gccaaactgc
7801 cttgccgcgc attccaagaa acacACCTAT TCCGGTcgca atgtctacta tgtgtggttt
      -> E2-bind
```

HPV3

LOCUS HPV3 7820 bp ds-DNA VRL 04-OCT-1993
 DEFINITION Human papillomavirus type 3 (HPV-3), complete genome.
 ACCESSION X74462
 SOURCE Human papillomavirus type 3 DNA.
 REFERENCE 1 (bases 1 to 7820)
 AUTHORS Delius,H. and Hofmann,B.
 TITLE Primer-directed sequencing of human papillomavirus types
 JOURNAL Curr. Top. Microbiol. Immunol. 186, 13-31 (1994)
 REFERENCE 2 (bases 1 to 7820)
 AUTHORS Delius,H.
 TITLE Direct Submission
 JOURNAL Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG
 COMMENT HPV-3 was first isolated from the benign warts of a patient with EV by Kremsdorf in 1983 (J Virol 48: 340-51) and subsequently sequenced by Dr. H. Delius. HPV-3 is most closely related to HPV-10 and HPV-28. Each of these viruses display similar clinical properties. They are most often associated with plane wart-like lesions showing a distinctive histological pattern. Although they can almost always be detected in the flat warts of epidermodysplasia verruciformis patients, neither of these types is solely EV-specific, as they are also found in the general population. Frequently, these viruses have been detected in biopsies taken from cutaneous warts of immunosuppressed patients. They seem to be almost exclusively associated with the induction of benign, cutaneous lesions, although related viruses have been detected in genital carcinomas.
 BASE COUNT 2171 a 1637 c 1923 g 2089 t
 ORIGIN 101 bp upstream from beginning of E6 cds
 1 tctaactata attataaata acaatgcaca taataaaaag tagggagtaA CCGAAAACGG
 -> E2 bind
 61 TacgACCGAA TGGGGTAcataaaaaaggag gcacaTAAtg cATGgcagta gccatgtcta
 -> E2 bind E6 cds ->
 E6 orf start ->
 121 tggatgcaaa ctgccccaaa aacatatcttc tactgtgcag aaacaccgga ataggatttg
 181 acgaccttcg cctgcactgc atattctgta cgaaacagct gactacaact gaactacaag
 241 catttgcatt acgggaactg aatgtggtgt ggagaagggg agcgccttac ggtgcttctg
 301 cacggtgttt actttagtag ggcattgcac gacgcctaaa atattgggaa tattcatatt
 361 atgtatctgg cgtggaagaa gagacaaaac aatcaataga tacacagcaa attagATGct
 E7 orf start and cds ->
 421 acatgtgtca caaaccactg gtaaaggaag agaaggacag acaccgcaac gaaaagcgaa
 481 gactgcacaa aatatctggt cattggaggg ggagctgtca gtactgctgg tcacgatgca
 541 cggctccgat ccacgaTAA aagatataga attgagtctt gcaccagagg acgtccctgc
 <- E6 end
 601 actatgcaat gtgcaattag atgaagatga gtatataaat gctgtggaac cagcgcaaca
 661 agcgtattgt gtagtcacag tgtgtccgaa gtgtagtcca caacttcgac tggtggtaga
 721 gtgcagccac gcagatataa gggccttcga gcagcttctg ctgggcacac tgacggttgt
 781 gtgtccccgc tgcgtgTAAc aggacATGga tgatacttca ggtacagagg gggaaatgttc
 E1 cds ->
 E1 orf start ->
 <- E7 end
 841 cgagtggaa cgggctggag gatggtttat ggtagaggca atagtagaca ggcggacggg
 901 cgatacagtg tcaagcgatg aggatgagga ggaggacgga ggggaagatt tagtgattt
 961 catagatgat aggcctgtag gggacggaca ggaagtggca caggaactgt tgctgcagca
 1021 agcagctgcy gatgacgatg tagaagtgca gacagtaaaa cgaaagtgtg ctcccagtc
 1081 gtatttttagc cctgtgtgtg tacatcccag catagaaaat gagctaagtc cgaggctaga
 1141 tgcaataaag ctggggagac aaacatcaaa agccaaacgc cggctatttg agctaccgga
 1201 cagtgggtat ggccaaacac aggtggatac ggaatcgggA CCAAACAGG Tacaggacat
 -> E2 bind
 1261 ttgtaagaca agccaacaag atggctgcca ggggtcggat gaggggagag gtaggaatgt
 1321 ggggggaaat ggcagccagg aggaggagcg tgcaggaggg gatggggagg aatcgcagac
 1381 tgagagtgtg cagacagata cgacagcctg tggagtgttg gcaatattaa aagctagcaa
 1441 tcacaagca acgctactgg gtaagttaa agaacaattt gggtaggat ttaatgaact


```

1501 gattagacac tttaaaagta acaaaacagt atgtagcgat tgggtggtat gtgtgtttgg
1561 tgtatactgt acattggcag aaagctttaa gacgtaata caaccacagt gcgaatatgc
1621 acatatacag gtactatcct gtcaatgggg catgacagtg ttaacgttgg tacggttcaa
1681 acgggccaana aacagagaga cgggtggctaa aggtttcagc actttgctaa atgtgccaga
1741 aaaccacatg ttaatagagc caccaaaatt aagaagcgct ccagcagcgc tgtactggtt
1801 caaacaagc ctatcaaatt gtacgaggtg gtttggggaa acaccagagt gगतatgtag
1861 gcagacatg gtgggacatg cattagagga agcgcagttc agtctgtcag aaatggtgca
1921 gtacgcatac gaccacgaca taacagatga aagcacgttg gcatatgaat atgcactaca
1981 agcagataca gatgcaaatg cagcagcgtt cctagctagc aattgtcagg caaaatatgt
2041 aaaggacgca tgcacaatgt gcagacatta caaaagaggt gaacaggccc gaatgaacat
2101 gtcagaatgg ataaagttta gaggagataa aatacagggg gatggcgatt ggaaaccaat
2161 agtacagatg ttaaggtacc aggacgtaga atttatacca tttctatgcg ctctgaaatc
2221 attcctacaa ggaataccaaa aaaaaagttg tatagtgttt tatggaccag cagatactgg
2281 gaagtcatac ttttgcatga gcctgttgaa atttctgggc ggggtagtta tatcttatgc
2341 caattccagc agccattttt ggttgcaacc attagcagaa gccaagatag gtttgctgga
2401 cgatgcaact agtcagtgtt ggtgttatat agacacgtat ttaagaaatg ctttagatgg
2461 aaaccaggtg tgcatagata gaaagcatag ggccttgcta caactgaaat gtocctcgtt
2521 attgataaca actaatataa atcctttggg gगतgaaaga tggaagtac tgcgcagcag
2581 actgcaggtg tttacattta acaacaaatt tccattaact acacaaggag agccactgta
2641 tacattaat gatcaaaact ggaaatcctt ttttcaaagg ttatgggcac gttTAAacct
                                     E2 orf start ->
2701 taccgatcct gaagacgagg aggacaATGg aaacactagc gaaccgttta gatgtgtgcc
                                     E2 cds ->
2761 aggacaaaat actagaactg taTGAaaagg atagcgacaa acttgaggac caaataatgc
                                     <- E1 end
2821 attggcaatt gatgcggtta gagcaagcct tgtgtgtaaa agcaagggaa tgtggattaa
2881 cacacattgg ccaccaggtg gtgccacctc ttagtgtaac caaagcaag gcacgcagtg
2941 ccattgaagt gcatgtatct ttgcaacaat tacagcacag tgcatatgca caagaccctc
3001 ggacactgcg agacacgtca cgggaaatgt gggacacagt tcccaagaag tgctggaaaa
3061 aaagaggttt aactgtggaa gtcagatatg atggagacga aaacaaagca atgtgttatg
3121 tacaatggag ggaataaatt gtgcagaact atacagatga taactgggtg aaggtggcag
3181 gactggtgtc tcatgaggggt ctatattaca tgcaagagg acagaaaact ttttatgtaa
3241 aatttaaaga tgatgcgcgc gtgtatgggg acacaggaac atgggacgta catgtggggag
3301 gcaagTAA taccacgat tcatttgacc ctgtatctag cacacgagag ataccgcctc
E4 orf start ->
NH2 terminus unknown
3361 ctggacctct gtacgcctgt accaccaag cggccacca agcccaggtg ggcgctccg
3421 aaggaccgga gcaaaagcga cagcgactcg agacggctca cggggagcag cagcagcaac
3481 agcagcagca acagcaacag caacaacata cccaaacccc cggcccgcaa accactgaac
3541 gagcacgtca accattggac actgacagga cccgggaccg tgacactacg tgtccacacc
3601 ccacgggca tcgaagtgc cctgactgtg tgctgTAA taccctaaga ggtgatccta
                                     <- E4 end
3661 actgtttaaa atgttttaga tataggttaa acaaaggtaa aaataagtta tattcaagga
3721 cctctccac atggaggtgg tcctgtgaat cagaaaatca gtgtgctgac gtaaccattt
3781 ggtatacaag ttatggtcag cgggaagcat tttgtccac cgtaaaagtg ccaccaggtg
3841 ttcaagtgat actgggacac atgtcaatgt tcacaTAA tgtgtccccg atgttacagt
                                     <- E2 end
3901 ctggattact atttgtcag gctttctctg tgggtgtatt ttgtgctgct gctgtgtttg
3961 ttttggctgt gtgtgcttcc tgcgctaacg tgctatctgg caattgtgct ttgtgtgac
4021 ttggctcga tagcattgta tttacaaatt gtatcacgca ttgtacagaa taacacatag
4081 gttttactat gtatcctctg gtactcacag acaacaatgg cgaccatctt gtctgtttg
4141 ttgagcctgg agacgtgtac atattattgc tgtttatgtt agctgtcata cttacattgt
4201 ttattatgta tagacatctg ggactcctgt aaggtttag ttgcaggtca cctgtatgta
4261 ttcttccttg atgtatatgc ccTAGTgtgg tattgtacca cgtctttta tactgctatg
                                     L2 orf start ->
4321 tttttttta cagttcaata aagcaaccAT Ggtggcacat cgtgcaaggc gtcgcaaggc
                                     L2 cds ->
4381 tgcactctgc acacagcttt atagaacctg caaggccgca ggcacatgct cccctgatgt
4441 tattcccaaa gttgagggca ccactttggc cgatcgtatt ttgcaatggg gtagcttggg
4501 tgtttatttg gggggtctgg gcattgggtac tggatccgga actggggggc gcacaggtta
4561 tgcgccaatt agtacagggc ctgggtactgt tgttgatgtt agtgttctg caaaACCTCC
                                     -> E2 bind
4621 TGTGGTaaat gagcctggtg ggccatcgga ccctccatt gttaacctat tggaagactc
4681 cagtattatt aattccgggt ccaccatacc gACCTTACT GGTactgatg gattcgaagt

```

HPV3

```

                                -> E2 bind
4741 tattttcttca gccacaacta ccctgtctgt attagatatt acacctgcca gtgacaatgt
4801 ggtgggttagt agtaccgaatt ttagcaatcc agcttttaca gaACCTTCCC TGTGggaggt
                                -> E2 bind
4861 tcctcagaat ggtgaggttt cagggcacat acttattage accccacat ctggtacaca
4921 tggttatgaa gaaattccta tggaaacctt tgcttcgcca ggtacgggaa ctgaacctat
4981 tagtagcacc cctgtacctg gtgtaagtag aattgcaggt ccccgctat atagcaaagc
5041 tgtcacacag gttaaggtaa cagatcctgc tttcttgacc cgtcctcgct cgttaatgac
5101 atttgacaat cctgtgtttg agccagaaga tgagactata atatttgaac gtccgtactc
5161 tcctcacag gtgcctgact ctgacttctt tgacatttta cgtttgaca ggctgtcttt
5221 aacttctcgt aggggtactg tgcgttacag tagggtaggc caaaaattaa gcatgcgac
5281 tcgcagtgac aagggtcttg gtgctcgagt gcattattat caagatttaa gccccatagg
5341 tcctacggag gacattgaaa tggaaacctt gattgctcct gcactcgcct cagcctatga
5401 ctctctgtat gatgtgtatg cagatgtgga cgatgctgac ataggtttta catctggagg
5461 tcgtagtgac actctgtcta gaggccgtgc tacagtgtcc cccctgtcct ccaactctgc
5521 cacaaagtat ggcaatgtca ccattccctt tgtgtctcct gtggatgtgc ctttacaacc
5581 tgggacctgat attttactgc ctgcatcagc tcagtggcgg tttgttcctt tgtctcctgt
5641 tgacacaact cattatgtct acatagATGg cggggatttt tatctatggc ctgtcacctt
                                L1 orf start and cds ->
5701 ctttttgccc cgacgtcgtc gccgtaaacc tgtctcatat tttcttgca atggcactgt
5761 ggcgctcTAG tgacaacctg gtgtacctgc ctcctacccc tgtttccaag gttctcagca
                                <- L2 end
5821 cggacgacta tgtgacacgc accaacattt attattatgc aggcagttct cgcttgctga
5881 ccgtgggtca tccttatttt gctatcccca aatcttctaa ttccaagatg gatattccta
5941 aggtgtccgc ctttcaatat agagtgttta ggggtcgggtt gcccgacca aataagtttg
6001 gcctaccaga tgcacgcata tataaccagc acgccgaaag gctggtctgg gcttgcaactg
6061 gggttgaggt aggccgcggg ctgcctttgg gtgtaggcct cagtggacat cctctttata
6121 acaagctaga tgacactgaa aactctaaca tagcacatgg ggacataggt aaagattccc
6181 gggacaacat atctgttgac aataagcaaa cgcagctatg tattgtgggt tgtaccccac
6241 ctatggggga gcattggggc aaaggaacac catgtaagca gaatgcgtca ccgggtgatt
6301 gtctctctct agagcttatt actgcaccta tacaagatgg cgatatgggt gacacaggtt
6361 atggtgccat ggactttggt aacttgcagt ccaataagtc agacgtgcca ttagatattt
6421 gccagaccac ctgcaaatat cctgattatt tgggtatggc cgctgagccc tatggcgaca
6481 gcatgttttt ttatttgca aaggagcagt tgtttgcaag acattttctt aacagagctg
6541 gtagggctgg agacaccgtg cctgacgcgt tgtacattaa aggtgacagt cagagcggcg
6601 gtcgggataa aattggtagt gctgtgtact gtccctacccc tagtgggtcc atggtaacat
6661 ctgaaacgca gctattcaat aagccatatt ggctgcggcg tgctcagggg cacaataatg
6721 gtatatgttg ggccaaccaa ttgtttgtga ctgtggtgga taccacacgt agtactaata
6781 tgacattgtg tgtttctact gaaacctcgg ctacatata tgctactaaa tttaaagagt
6841 atttaagaca cggggaggaa tatgatttac agtttatatt ccagttgtgc aaagttacat
6901 taactcctga aattatggcc tatttacaca caatgaacag tactttggtg gaggattgga
6961 actttgggtt aaccttgcca ccgtccacta gcttggagga cacctataga tttttaactt
7021 cctctgcat tacctgccag aaagatgcac ctcccactga gaagcaagac ccctacgcca
7081 aactaaactt ttgggatgta gatcttaagg atcgtttttc cctggatctt tcgcagttcc
7141 cccttggcag gaaatctctc atgcagctcg gtgtaggtac ccgctctagt atactgttc
7201 gtaaacgctc ggcgacaacc acatctagaa cagctgctgc aaaaaggaag cgcacaaaaa
7261 aaTAGccaca tttgtgtttt gtatgtgtaa cctgtgtgta tgttttttat gtatgtactg
                                <- L1 end
7321 tgtgtgtaat gtgtactgtc tgtgctatgt gtttgtactg tattatggtg tgtgtatgtg
7381 tcaataaact gtgtcacata gttttatatt ttttaatttt tgtaattact gttcctgtga
7441 gtaagaaagg taattctggg tcatgcgACC GATTTCGGTt ctcaaatggc ccgcctttgc
                                -> E2 bind
7501 aggtgtgcac acaacaactt agtcatactg atctatatcc tgcgacctgc cttgtcacgc
7561 atagtttttg ctgtgatatt atcttttcta tagtttattt tattgctgca tcattctccc
7621 tggcagctct atctgtctcc attgcaaatt aacagcttct gggcactaac ttattatgac
7681 tactttcaca taattactgt cttggctgcy ttttctagtc tgccttgcca atagtgtctt
7741 ccaaatctcc accaagacac ACCTAATCCG GTCgctgctt gctttctagc cataatztat
                                -> E2 bind
7801 gcagttgcta cacgttcctt

```

LOCUS HPV7 8027 bp ds-DNA VRL 04-OCT-1993
 DEFINITION Human papillomavirus type 7 (HPV7), complete genome.
 ACCESSION X74463
 SOURCE Human papillomavirus type 7 DNA.
 REFERENCE 1 (bases 1 to 8027)
 AUTHORS Delius,H. and Hofmann,B.
 TITLE Primer-directed sequencing of human papillomavirus types
 JOURNAL Curr. Top. Microbiol. Immunol. 186, 13-31 (1994)
 REFERENCE 2 (bases 1 to 8027)
 AUTHORS Delius,H.
 TITLE Direct Submission
 JOURNAL Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG
 COMMENT HPV-7 was first cloned from a hand wart of a butcher in 1986 by Oltersdorf et al. (Virology 149: 247-50) and subsequently sequenced by Dr. H. Delius. HPV-7 is prevalent in the warts commonly developing on the hands of butchers and other persons occupationally involved with handling meat or fish. Further, reports indicate that the warts tend to disappear if the person afflicted ceases to practice this type of occupation. These lesions display typical histological features but clinically are more hypertrophic and hyperkeratotic than common warts. HPV-7 has also frequently been detected in seropositive HIV-infected patients, in both lesions of the oral mucosa and of the facial skin. Unlike the skin lesions induced by HPV-7 these oral mucosal lesions vary widely in their clinical features and display no characteristic histology. Unexpectedly, HPV-7 has been detected in a significant proportion of oral papillomas (9%) and in one case of a tonsillar carcinoma (de Villiers et al. Laryngol Rhinol Otol (Stuttg) 65: 177-9 and Snijders et al. in \it Human Papillomaviruses, edited by Harald zur Hausen, Springer-Verlag, Heidelberg, 1994, pp 177--197). HPV-7 has only rarely been detected in the skin lesions of otherwise healthy individuals not belonging to the occupational group described above.
 BASE COUNT 2458 a 1450 c 1722 g 2397 t
 ORIGIN 101 bp upstream from beginning of E6 cds
 1 tgtttaataa ttgtacagtt actaaagggc gtaACCGAAA ACGGTccgAC CGAAAACGGT
 -> E2 bind -> E2 bind
 61 acatataaaa accaacccaa aaaaccTGAT ctggggccac tATGtctgca cgttgcygct
 E6 orf start -> E6 cds ->
 121 ccacagctag aactttatgt gaattatgtg accagtgcaa tataacattg cctacgttgc
 181 aaattaattg catattttgt aacagcattt tacaacacgc tgagggtgctg gcctttgcat
 241 ttgagagatt atatgtatgt tggcgcaacg actttccctt tgcagcgtgt gtaaagtgtt
 301 tagaatttta tggaaaagtg aatcagtata ggaactttag atacgctgca tatgcaccaa
 361 cagtgggaaga agaaacagga ttaacaattt tagaagttag aataagatgc tgcaaatgcc
 421 acaaacatt gtctcctgtg gaaaaacca accacattgt taagaagaca cagttttTA
 481 Aactgcaaga ttcgtggaca gggactgtc tgcactgttg gaagaaATGc atggagaaag
 E7 orf -> E7 cds ->
 start
 541 gccaacgctc ggagacatcg tgtTAGactt gcaaccgaa ccagtaagtt taagttgcaa
 <- E6 end
 601 cgagcaatta gacagctcag actcagaaga tgaccatgaa caagaccaac tagacagctc
 661 acacaataga cagcgtgagc aaccacgca acaggacttg caagtaaatt tgcaatcatt
 721 taaaatagta acacattgtg tattttgtca ctgtttagtt cgctagtag tccattgtac
 781 tgctacTGAT ataagcagg ttcacagatt gctgatggc aactaaata tagtgtgccc
 E1 orf start ->
 841 caactgtgca gctacagcgT GAcacaATG gcagacgatt caggtacaga ggatgtgggg
 E1 cds ->
 <- E7 end
 901 tctgggtgtt caggatggtt ttagttgag gctgtggtag ataaacaac aggtgatgta
 961 gtttcagaag atgaagatga ggatgctata gaggacagtg gatatgatat ggtagatttt
 1021 attaatgata ctgtagtaag tgaacatgaa gaactaagta atgcacaggg tttgttacat

HPV7

```

1081 gcacaacaga catgtgcaga tgctgtagag ttgtgtgagc taaaacgaaa gtacattagt
1141 ccatatgtaa gtcctattca gtgctcagaa cggcccgtagg acgggggattt aagtccaagg
1201 ctgcatgcca taaagcttgg cggcggtaaa aaggctaaaa ggccggttgtt tgagcgattg
1261 gagcagcgag acagtggcta tggctattca caagtggaaa caacagagac acaggtagag
1321 gaagaacatg gcgaaccgga aggtatagag gggggcagtg ggagggctgc gacagttgaa
1381 acggaagcgg ttgaagtgtc agaagaagc agtgatgta tacagcaact tagtcccgct
1441 acacaggtgg tagagctggt taaatgcaag gatttaaatg ctaaactgtg tggtaagttt
1501 aaggaacttt ttggagtggg ctttcacgat ttggtagac agtttaaaag tgataaatca
1561 acgtgtacag attgggtgta cgcagtgttt ggggttaatc ccactatagc agaaggcttt
1621 catacattat taaaaggaca ggcattatac ttacatacac agtggacaac gtgtagatgg
1681 ggtatggat tgcttgcat gtgtagatat aaggtagcaa aaaaatagaga aacagtagtg
1741 cggcagcttg ccaaaatggt aaatgtacca gataatcaac taatggtaca accacctaaa
1801 ttacaaaagt ctgcagcggc ttattttgg tttagatcag gaatgggtaa tggaaagtga
1861 gtgtctggca caacaccgga atggatagct aaacaaacaa tgttggaaaca tagttttgct
1921 gaagcacagt ttagttaaac tcagatggtg cagtgggcat atgataatgg gcatacagat
1981 gaatgtgaaa tagcatatta ttatgcacaa atagcagata tagatgcaaa tgcagcagcg
2041 tttttaaaaa gtaacaatca agctaaatat gttagagatt gtgcagctat gtgtaagcat
2101 tataggttgg cagaaatgag cgggatgtca atggcagact ggataaagca tagaggtgaa
2161 aaatgtgatg aaggagattg gaagcctata gttaaactac taagatatca acatatagac
2221 ataattgtgt ttttagctgc gttgaaaaaa tggctacatg gaataccaaa aaaaaattgt
2281 atttgtattg ttggacctcc agatactgga aagtcatggt ttggtagatg tttgatgcat
2341 tttttgcaag gtactataat ttcatttgtt aattcgtgta gccatttttg gttgcaatca
2401 ttagtggcag caaaagtagc tatgttagat gatgtgactt ctgcatgctg ggctatatg
2461 gatcacaca tgagaaattt attagatgga aaccacaaca gtatagatag aaaacataaa
2521 tcattggctg tgattaaatg tctccatta ttgttaacat caaatataaa catataaacat
2581 gattgcaaat atcaatattt acagagtaga gtgacagtgt ttgaatttcc aaatccattt
2641 ccatttgaca gcaacggaaa tgctgtgtat gaattaagtg atgcaaattg gaactccttt
2701 tttaaaaggt tggcgtccag ttTAGagctg cagacaacag aggacgaggA TGgagaaact
          E2 orf start ->          E2 cds ->
2761 agccaggcgc ctagatttgt gccaggaaca gttgttagaa ctttaTGAac aagacagcaa
          <- E1 end

2821 acagctacag caccatata tgcactggaa atataacgt tatgaaagtg taatatatta
2881 tacagcaaga caaatgggca ttaaactgtc gggccaccag gtggtgccaa gtttagatgt
2941 gtcaaaagcc aaagcccatg cagcaattga aatgcaaatg tgtctagaat ctttgcaaac
3001 tactgaatat aacttagagc catggacggt acaggacaca agtcaagaac tatggcttgc
3061 agaaccaaag aaatgtttta aaaaaggagg aaagacagta gaagttagat ttgactgtaa
3121 tgaacataat gcaatgcatt atactctatg gactgcagta tatgtacagg tggaggatac
3181 atggacaaa gttgaaggcc aggtggacca cagaggccta tttatacag tgcattgggtg
3241 cacaacatat tatgTAGact ttggaaaagga agcacataca tATGggaaaa caaatgactg
          E4 orf start ->          E4 cds ->
3301 gactgttatt gtgggtcac gcgttatatg ttctcctagt actgtogaag ggctaccat
3361 tgttgcgctt gttgacatca gacatcccgc ggccaccgac gccaccgacg ccaccaaggt
3421 gcacgacgcc ccctacgccc tgcccgcgtc gaccacaaa gtatacaacg acagccacgc
3481 accgcccga aagcgaagga gagacggaga cttgtccatc agtgcagtgg acggatgtag
3541 tggaaagaaa tacgtggaca ctggaaacag agcacgctcg cctgatattg aaagcaacaa
3601 caaaatcagg aacagtggtg gaggtcattc tacacctaTA Atacaactgg aaggtgatgc
          <- E4 end

3661 caattgttta aagtgtttta gatataggct aacaaagggt agccatttat atacaaattc
3721 ttcaactaca tggaggtgga ctacagaatc tagaacaat aaaaatgcca ttataacatt
3781 aacatatagt agtgtacacc aacggtcaca atttctagca cttgtaaaaa tacttaaac
3841 tattaaacat agtttaggca tgttaactat aatgTAAata tgtttgtata tttgtaaaaa
          <- E2 end

3901 acatatgtat ggaacgcaac tgtgaagggt aatgcatatg catttaccac acaaccagcc
3961 aaactgctgt tattggtggt aatagagtca acgcttggca tactaatagt atataactc
4021 tgtcttataa ctgcaatttt aatagatcat catataacct atttctgtgt ctggagctct
4081 tttagggcca ctatttctat tctttgcttt ataacgtggg gtgcaactac atctattatt
4141 aacgtatttt ttttagtggt gttagtgtgg tatttgctcg cactgttctc tcatatgtca
4201 attgtgtatg ctatacaaca agaacaacaa taacaataat tgcaaatatg ttaacctgta
4261 aattagatga tggatgatac tggatggcat tatggatggt acttacattt ataactgtat
4321 tgttactggt attggcgttt cattgtagga cattatatgt atacaaatat agtaagtaaa
4381 atactgtgta ttaaaTAAat atttttatat ttgcagcggc ttggtATGgt gtccagcagg
          L2 orf start ->          L2 cds ->
4441 ccccgtaggc gtaagcgggc gtccgcaaca cagttatatc aaacatgcaa ggcagcaggc
4501 acctgtccac cagatgttgt taataagggt gagcaacaaa ccgttgcaga tcaaatttta

```

```

4561 aaatggggca gcatgggagtg gttttttggc ggtcttggca taggttcagg ctctggctca
4621 ggcggccgtg ctggctatgt gcctttgtct acaggttccc gtgcaatacc tcctaaatca
4681 ttagctccag atgttattgc taggcccct gttgtgggtg atactgtcgc cccactgat
4741 ccatccattg tatctttaat tgaggaaagt agtattatc agtctggggc tccttcgcca
4801 gtaattccca cagaggggtg gttttcaata acatcatcag gtacagatgt ccctgcaatt
4861 ttatgatat cttctactaa tacagttacat gttacatcta ccacacacca taaccacata
4921 tttactgata cttcggttgt gcagcctatt ccacctgtag aggctagtgg tcgfatcatt
4981 gtgtcgcatt cctctattac tactgggtgca gctgaagaaa tacctatgga cacatttgtt
5041 gttcatagtg atccactgtc cagtacacct gtgctcgttg tgcagcgcg gcctaaagt
5101 gggctatata gcaaagcttt gcagcaagta gaaatagtag atccaacatt tatgtccacc
5161 cctcaacgtt taattactta tgacaatcct gtatttgaca acattgaaga tacactacat
5221 tttgaacagc cttctattca taacgcacca gatcctgcct ttatggatat cattacttta
5281 cataggcctg ctttgacctc taggcgtggg gtggtagctt ttagtagggg ggggtcaacgt
5341 ggaacatgt atacacgtcg tgggACCCGT ATTGGTggtc gtgtacactt ttttaaagt
      -> E2 bind
5401 attagtccca tagcttcacg tgaagaaatt gaattgcacc ctctagtggc ctcacctaatt
5461 aacagtgacc tttttgatgt ttatgcagat atagatgata ttgatgaaa tatattatat
5521 tctactatag acaataatac accaacttct acctattcct tgtatccagg taattctaca
5581 cgcatagcaa atacatctat acctcttgcc acaattcctg atacattcct aacatctggg
5641 cctgacatag tgtttccttc tgttctcgtc ggtacacatc atttgctgtg gtcaccttct
5701 atacctgcca tatctgtact gattcgcggg actgattatt atttgaatcc tgcatactat
5761 tcagaaaac gccgaaagcg catatTAGca tatTAGgATG tggcaactta atgaaaacca
      L1 orf start ->      L1 cds ->
      <- L2 end
5821 agtgtattta ccaccgccc cgctgtgtgc tacaattggt agcacagatg agtacgtgca
5881 acgcaccagt ttatattatc atgcaggtag taccagggta ttaaccatag gacatccata
5941 ttttgaattg aaaaagccta atggcgatgt atcgggtgct aaagtgtctg gacatcaata
6001 cagagtgttt agagtagcgt tgcccgacct taataaattt ggattatcag acacgtcttt
6061 atttaattct gaaacccaac gccttgtatg ggctcgtggt ggtgttgagg tcggtcgagg
6121 tcagccatta ggtgtaggca ttagtgggtc tccatacttt aataaagatg aagatgtgga
6181 aaactcgtct gtatatggaa cagtaacctg tcaggacagc agagaaaatg ttgctatgga
6241 ttataaacia actcagttat gtattgtagg ctgtactcct cctattggag aatattgggg
6301 tatgggtaca ccgtgcaatg cttctaaagt gtctcctggt gactgtcctg tactagaatt
6361 aaaaagtga gttattgagg atggcgacat ggttgatgca ggctttggg ccatggattt
6421 tgcattcatt caggccaata aaagcgatgt gcctttagat ttatgtacat ctattagtaa
6481 ataccagat tattttaggaa tggctgcaga accgtatggt aatagttat tttttttct
6541 tagaagagaa caaatgtttg ttaggcactt ttttaatagg gcaggaacta ctggagacag
6601 tgttccaaat gatttatata taacagggtc atctaactgc gcttctattg caggcagtat
6661 ttattattcc acaccaagtg gctctctagt tacctctgat tctcagattt ttAATAAAcc
      signal ->
6721 tttgtggata caaaaggccc agggtcataa caatggcatt tgttttggca atcagttatt
6781 tgttacagtt gtagatacta ctctgtagcac aaatttaaca ttatgtgctg ctacacaatc
6841 gccacacca acaccatag acaatagtaa gtttaaagaa tatttacgct atggggaaga
6901 gtttgattta cagtttattt ttcagttatg tgttattaca ttaaatgcag aggttatgac
6961 atatatacat gctatggatt cttccttatt agatgattgg aattttaaaa ttggctctcc
7021 agcgtctgca accttgaag atacttatag gtttcttacc aataaagcca tagcatgtca
7081 gcgtgatgca ccccaaaaag aaaaggagga tccatataaa aaatataaat tttgggaagt
7141 aaatttaaca gaaaaatttt catctcagtt agatcaattt ccattaggac gtaagtttct
7201 tatgcaggca ggccctacgca cagggcctaa gtttaaatcc aggaagcgc ctgccctac
7261 ctcttcttcc tcttctgggt cagtcacccc caaacgtaag aaaacaaaac gaTGAtgtg
      <- L1 end
7321 tatgtgtgtg tgtgtgatac tcccttgca tcccttatgt gttgttactc tggaatgtat
7381 gatgtgtgtg tttatgtatg tgtgtttgta tatgtgtgtt atgtgtgaga atgtatttgt
7441 gtgtatgaat gtaatgtatt gttgtgtgtg ttgttaataa taaatatatt gtgtgtgtg
7501 tggtaaagta ctttctttgt tttaaaagt acttattaaa ctcaactatt ttaaaatgtc
7561 tgctctgcac actgcaACCG TTTTCGGTcg cggttggcaa ctattacat ttgtcagcat
      -> E2 bind
7621 gtttttataa catgtttaaa attgtagct ttatataact atataaatcc ttcaatttcc
7681 acccataacc gtttccagtc tcggttggca agtcaccatg tttgtcagca tatttgcatt
7741 gcatgtttca aattgctagg tcaaagttcc ctgcacaaaa tgcgcaccaa tatgtactat
7801 taggggtgagg ttgccacacc tttaattaca cttttattgc actgttactc atcttatttt
7861 atacgcttcc aaacttgctt ttaggcacat agttttactg cataaacatt tagctaagag
7921 cagtttggca ccacataaca ctatggttaa caaataacac attactcatg gtacacacct
7981 gcaaACCGCT TTCGGTgtct acacgtttta ttactttcta gttatta

```

HPV7

-> E2 bind

LOCUS HPV10 7919 bp ds-DNA VRL 04-OCT-1993
 DEFINITION Human papillomavirus type 10 (HPV-10), complete genome.
 ACCESSION X74465
 SOURCE Human papillomavirus type 10 DNA.
 REFERENCE 1 (bases 1 to 7919)
 AUTHORS Delius,H. and Hofmann,B.
 TITLE Primer-directed sequencing of human papillomavirus types
 JOURNAL Curr. Top. Microbiol. Immunol. 186, 13-31 (1994)
 REFERENCE 2 (bases 1 to 7919)
 AUTHORS Delius,H.
 TITLE Direct Submission
 JOURNAL Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG
 COMMENT HPV-10 was first identified by Orth et al. (Proc. Natl. Acad. Sci. 75:1537-1541) as HPV-3 and was subsequently renamed by Kremsdorf et al. (J. Virol. 48: 340-351) as HPV-10. It was isolated from the benign warts of a patient with EV and subsequently sequenced by Dr. H. Delius. The patient was a 23 year old Polish patient (J.G.) with nonfamilial EV. It has been associated with benign EV lesions and with flat warts in the general population.
 The flat warts commonly found associated with HPV-10 exhibit minimal or no papillomatosis. They are almost always multiple and common locations are the arms, hands and around the knees. A majority of these warts spontaneously regress but some may persist for much longer periods.

BASE COUNT 2169 a 1651 c 1981 g 2118 t
 ORIGIN 101 bp upstream from beginning of E6 cds
 1 ttataaacta taatctagac aataataaaT AGggagggAC CGAATACGGT gcgACCGAAT
 E6 orf start -> -> E2 bind -> E2 bind
 61 GGGGTacata taaaacaagg cccgtagcat ctgcagaagc tATGtccatg ggtgcacagg
 E6 cds ->
 121 aaccagaaa catattgctt ttgtgtagaa attgtggaat acctttggag gaccttcgcc
 181 tgtgctgtat attttcaca aaacagctga cgcagcggga attggcagca tttgcaacta
 241 gagaattata tttggtgtgg agagcgggag tgccatacgg tgctgtgca cgggtgttac
 301 tcttacaggg cattgtacga cgcctaaaat attgggacta ttcattatg gtagaagggtg
 361 tggaagagga gaccaaaaca tctatatata cacagctgat cagatgctac atgtgtcaca
 421 aaccgctggt aagggaaagaa aaagacagac atcgtaacga acggcgacga ctgcacaaaa
 481 tatcagggta ctggagaggt agttgTGAgT attgctggtc acgATGcagc gtccgcatcc
 E7 orf start -> E7 cds ->
 541 cacagTAAaa gatatcgaat tgagtcttgc accagaggat atccctgtat gcaatgtgca
 <- E6 end
 601 attagatgaa gaagattata cagatgagggt ggaaccagca caacaagcgt atagggtggt
 661 aacagaatgt acaaagtgta gtttaccact gcgactggtg gttagagtgca gccacgcaga
 721 tataagggca ctggaacagc tgctactagg cacattgaag ctcggtgtgc ctcgctgctg
 781 gTAAcaggac ATGgacgata atacaggtac agaggggggc gcatgttccg aatcggaacg
 E1 orf start -> E1 cds ->
 <- E7 end
 841 ggcgggtgga tggtttatag tggaagccat tgtagatagg cggacaggcg atccaatc
 901 tagtgatgat gatgaggagg aggacgaagc aggggaagac tttgtagact ttatagatga
 961 tactaggtcg ctaggggatg gacaggaagt ggcacaggaa ctgttccagc agcagacagc
 1021 tgcagatgac gatgtagctg tgcagactgt aaaacgaaag tttgcacca gtccttattt
 1081 cagcccgtg tgtgagcaag ccagcataga acatgaacta agccaaggc tagacgccat
 1141 aaagctggg agacagtcag caaaggccaa acgtcggctg tttgagctac cggacagtg
 1201 ctatggccaa acacaggtgg atacggaatc gggACCAAAA CAGGTacagg gcagtagtga
 -> E2 bind
 1261 aacgcaagat ggccgacagg atgatgatga ggggagtgtg gtacagagca cacttgacac
 1321 aggcaaccaa aatggccgcc agaacaatga tgaggggagt gggaggaatg tgggggaaca
 1381 tggcagccaa gaggaggagc gtgcaggagg gtagggggag gaatctgact taaaaagtac
 1441 aagcactggc aagggagcag gtggcgtggt agaaatatta agagccagca ataagaagc
 1501 aactctactg ggtaagttaa aagaacagtt tgggttggg tataacgaac taattaggca
 1561 cttttaaagt gatagaacat ctgtgcccga ctgggtggtg tgtgtgttcg ggggttttg
 1621 cacagtggca gagggcataa agaccctaata acaaccattg tgtgattatg cacacataca

HPV10

```

1681 agtgtacca tgccaatggg gaatgacagt gcttatgctg gtacggtaca aacgtgccaa
1741 aaacagggaa acagtggcta aaggcttaag cacattatta aatgtaccgg aaagccagat
1801 gtttaattgaa ccacaaaaat tacgaagtgg tccagcagcg ttgtattggt acaagactag
1861 tatgtccagc tgtagcgagc tgtatgggga aacaccagag tggatagtca ggcagacaat
1921 ggtgggacat gcaatggagg atgcgagctt tagcctttca gagatggtgc agtgggcata
1981 tgaccatgat attacagagc agagcacgct ggcatacagag tatgcactga ttgctgacac
2041 agattc AAAAT gcagctgctg ttctaagtag caattgtcag gctaaatatc taaaggtgac
2101 atgcacaatg tgcagacatt ataaaagagg agaacaggcg cgcatgagca tgtcagaatg
2161 gatatggttt agaggcgaca aagtacaggg agatggagat tggaaaccaa tagttcaatt
2221 ttttaagatat caggatgttg aatttatccc attcctatgt gcctttaaaa cattcctaca
2281 aggagtacca aaaaaaagt gtttagtgtt ttatggacca gcagacactg gcaaatcata
2341 cttttgtatg agcttactta gatttttggg gggggctggt atatoctatg ctaactcaag
2401 cagccatttt tggttgcagc cattatctga agccaaaata ggactgtag atgatgcaac
2461 tagtcagtgt tggaaactata ttgacactta tttaaagaaat gccttagatg gcaaccaaat
2521 atgtgtagac agaaagcata gagcattgtt acagctaaaa tgcctccac tattaataac
2581 aacaaatata aatccattga cggatgaaag atggaagttt ttgagcagca gattgcagct
2641 cttcacattt aaaaaccctt ttccagtgc aacacaagga gaaccaatgt atacattaaa
2701 tgatcaaaaat tggaaatgct tttttcgaag gttatgggca cgttTAAGcc ttaccgatcc
                                     E2 orf start ->
2761 tgaagacgag gaggagCATG gaaaccctag cgaaccgttt agatgctgac caggacaaaa
                                     E2 cds ->
2821 tgctagaact ataTGAAAag gatagcgaca aacttgagga ccagatcacg cattggcact
                                     <- E1 end
2881 tattgctgtg agaaaatgct ttgctgtaca aagcaagaga atgtggactg acacatattg
2941 gccatcaggt ggtgccacct cttagtgtaa ctaaagccaa ggcacgcaat gccattgaag
3001 tgcattgtagc tttacagcaa ttgcaagaaa gtgcctatgc acacgaacct tggacattgc
3061 gggacacatc acgtgaaatg tgggacactg ctccctaaagg gtgctggaaa aaaaggggga
3121 taactgttga agtcagatat gatggagacg aatctaaagc catgtgctat gtacaatgga
3181 gggaaacttta tgtgcagaac tatagtgacg atagatgggt gaaggtgcca gaaaaagtct
3241 catacggagg tctatattat acacatgaaa atatgaacat atattatgtg aatttcaagg
3301 atgacgcttg tgtatatggg gaaacaggca aatgggaggt acatgtggga gcaaaagTAA
                                     E4 orf start ->
3361 ttcaccATGA tgcatttgac cctgtatcta gcacacgaga aatatccact cctggacctg
                                     E4 cds ->
3421 tgtgcaccag taacaccacc ccagcgtcca cccaagccca ggtgggagcg tccgagggac
3481 cggaaacaaa gcgacagcga ctgagggcgg tgcacggaca gcaccagcag cagcgacaag
3541 ggtccaaaga ttccaccag aaggccgagg aacgagcggg tggacaagtg gacagtgaca
3601 ggaccggtt gtgtgacact agaagtgcac accccgtccg gcaccaagt gaccctgact
3661 gtgcacctgT AAtacaccta cgagggtgat ctaacagttt aaaaatgctt agatatagat
                                     <- E4 end
3721 tacaccacgg aaaaaggaaa ctatactcac ggtcatcctc cacatggagg tggctttgtg
3781 agtcagaaaa ccaggcagcg tttgtaacgc tttggtatag cagcgataca cagcgtactg
3841 aatttcttaa tgttgtaaag gttcccctg gcatacaagt gattttggg tatatgtcaa
3901 tattcTAATA tgtcagatat ttctatacat atatagatct ataggtgat tgtacattct
                                     <- E2 end
3961 ggatttttac ttgtgtcgac tacattgctg ggctttttt gtgctgctgc tgtgtttgtt
4021 ttgctgtgt gtgcttcccg cgctgacgtg ctatctggca tttgtgctt gtgtgtactt
4081 gggccttata gcattatatt tacaaattgt gacacgtatt gtacaggca acacataggt
4141 ctcaaatgt atcctcttat attaaagat cacactgggt accatcctgt attgtgtttt
4201 gagccgggg acgtgtatct actactgtgt gtaatcattt ttgtataaT AGcactgctt
                                     L2 orf start ->
4261 gtgttatata gacatcttgg tgtattgtaa cataatatgt ggtagcctgt acggatccca
4321 cataggggg tgtatatcat atgttgtgt tgtgtgtacc ctagtgtgct attgtggcac
4381 cattcttttt ctatttttgt tttttttttt acagttaaat aaagcaaccA TGgtggcaca
                                     L2 cds ->
4441 acgtgcaagg cgtcgcaagc gtgcacccgc cacacagctt tataggacct gcaaggcctc
4501 aggcacatgc ccccagatg ttattcccaa agtagagggc accacccttg cagatcgcat
4561 tttgcagtgg ggtagccttg gtgtatattt ggggggocct ggcattggaa ctgggtctgg
4621 tactgggggt cgcacgggg atgttcctat cagtacacga cctggcactg ttgtggatgt
4681 tagtgttctt gccaggcctc ctgttgttat tgagcctgta ggcccgtcag atccttccat
4741 tgtaaatatt tggaggact ccagcattat taattcaggg tccaccatac ctacatttct
4801 ggtactagt ggcttgaag ttacatcctc tgccacaacc accccagctg tgttggacat
4861 cacactgcc agtgagaatg tggttattag tagtacaac tttacaaatc ctgcattcac
4921 agagccgtcc cttgtagaag ttccacagag tgggtagggt tcgggacaca tacttataag

```



```

4981 cacacctaca gctggcacc atggatatga ggaataacc atggacacgt ttgcttcttc
5041 gggactggc acggaacca taagcagtac ccctgtgcct ggtgtcagta gaatagcagg
5101 cccacgctta tatagcaggg ctaacacaca ggtcaaggtg tccgatcctg catttctgtc
5161 acgtccttcc tccttgtaa catttgacaa tcctgtgttt gaaccggagg atgaaacaat
5221 catattcgaa cgtccgtact ctcctcacg ggtgccggac cctgatttcc ttgacattgt
5281 gcgtctgcat cgccctgcat taacatctcg caggggtact gtacgtttca gccgcctggg
5341 tcaaaagtcc agcatgcgca cacgcagtgg caagggatt ggagcccggg tacattacta
5401 tcaggacctt agccctatag cccaatcga ggacattgaa atggaacctt tacttgctcc
5461 agctgcctct gacactatth atgacattht tgctgatgth gatgatggth atgthgctth
5521 tacagagggg tatcgtagta ccacacagtc caggggatat aataccactt ccccttgtc
5581 ttctacactt tccactaagt atggaaatgt aacaattccc tttgtgtctc ctgthgatgt
5641 aactttacat actgggctcg atattgtact acctacttca gcacagtggc catatgtgcc
5701 cctgtcacct gctgacacca cccattatgt gtacatagAT Ggcggggatt tctatctttg
      L1 orf start and ods ->
5761 gcctgttacc tttcacttct cccgacatcg tgcgcgtaaa cgtgtctcat atttttttgc
5821 agatggcact ctggcgctcT AGtgacaact tgggtgacct gcctcccact cccgtgtcta
      <- L2 end
5881 aagtctctcag cacggacgac tatgtgacac gcaccaacat ttactattat gcaggcactt
5941 cacggttgct tactgtaggc catccatatt tcctataacc taagtcaagt aacaataagg
6001 tagatgtccc caaggtatct gcatttcaat ataggggtgt tcgggtgctg ttgcctgacc
6061 ctaataagtt tggactgcct gacgcccgca tataataccc tgacgcccag cgactggtct
6121 gggcttctac tggggttgag gtaggtcgtg gacagccact ggggtgtggg ctcagtgagg
6181 accctctata taataagcta gaggacacag aaaactctaa tatagcacat gggccaattg
6241 gtcaggatcc acgggacaat atttctgthg ataataagca aacacagcta tgtattattg
6301 gttgtacacc tccaatggga gagcattggg gcaagggaac cccgtgcagg aaccacctg
6361 cacagggcga ttgccctccc ctggagctta taacttcccc tattcaggat ggtgatatgg
6421 tggacactgg ctatggggcc atggacttht ctgctthtaca attaataag tctgacgtgc
6481 ctatagatat ttgccaatct acttgtaaat atccagatta tttgggaatg gccgcggagc
6541 cttatggcga cagcatgttc ttttacttgc gcagggaaaca actgthtgc agacatttct
6601 ttaatcgggc tagtctgtht ggagacgcca tcccagacac tthcatatta aagagcaacg
6661 gtggggggcg agacgttggg agtctgtgtg atagccccc acccagtggg tccatggtaa
6721 cgtctgaggg tcaattgtth aataagccat attgctgctg gcgggcccag gggcacaca
6781 atggatatat ctgggctaac caattgttht ttactgtgth agacacgact cgcagtacca
6841 atatgtgctt gtgtgttctt tctgaggcct ccctgcccac tacgtatgac gccaccaaat
6901 ttaaagaata tttgaggcac ggagaggaat atgatttgc gttcatttht cagthtgcga
6961 aggtaacatt gaccccggat attatggcct atthgcacac catgaatagt agthtattgg
7021 aggattggaa cthtgggtht actthgccac cgtccactag cthtggaggac acatatagat
7081 thttgtctc thcagccatt actthtcaga aagatacacc cccaccgag aagcaggatc
7141 cctatgcaaa actthactth tgggacgtag atctthaaag taggthtthc ctggacctgt
7201 ctcagthtcc cctgggacga aaaththtgc tgcagctggg tgtacgtthc cgtcccgcg
7261 tctccgtccg caaacgcccg gcgactthc cgcaggatc cagggctgca aaaagaaaa
7321 gaactaagaa aTAActgcat gthtgtgtht aatgtatgtg catathtgtg tatgtggthg
      <- L1 end
7381 gctgthtaata ggtgthtgc atgtctatgt atgtgtgtat gtatgtaatg tgtthtgcga
7441 ttatgagcat gtatgtgtgt atgtgthaat aaagthtgtc acatagtht ataththttht
7501 atthtgtgtht thgctgthtcc tgtgagtht taagthtgtt ctaggthcagg agACCGATTT
      -> E2 bind
7561 CGGTacaaga tggccgctt tccaggtgtg cacaccacca attagtcatg ctgatctata
7621 tcctgcgacc tgccctgtca cgccaththt thggthtaaga thgtatagth tctatagtht
7681 agthtattgc tgtatcatgc thtctggcac ggcaactgt cthcattgca aaththaacag
7741 cthctgggca ccaacttht atgactactt tcacataat actgthtctg ctgctgthtthc
7801 tagthtgcct tgccaataaa atgtgctthc aaatctccac caagacacAC CTAATCCGGT
      -> E2 bind
7861 cgtgctthc thtctagctt taathaatgc agthtgcata cgtctctthc taactataa

```

HPV27

LOCUS HPV27 7823 bp ds-DNA VRL 04-OCT-1993
 DEFINITION Human papillomavirus type 27 (HPV-27), complete genome.
 ACCESSION X74473
 SOURCE Human papillomavirus type 27 DNA.
 REFERENCE 1 (bases 1 to 7823)
 AUTHORS Delius,H. and Hofmann,B.
 TITLE Primer-directed sequencing of human papillomavirus types
 JOURNAL Curr. Top. Microbiol. Immunol. 186, 13-31 (1994)
 REFERENCE 2 (bases 1 to 7823)
 AUTHORS Delius,H.
 TITLE Direct Submission
 JOURNAL Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG

COMMENT HPV-27 was first identified in a common hand wart from a renal transplant patient (Ostrow et al. J. Virol. 63, 4904) and subsequently sequenced by Dr. H. Delius. Common warts are dome-shaped with multiple conical projections. A majority of the warts regress spontaneously within 2 years, but some may persist for longer periods. HPV-27 was identified as a novel type when it was found to hybridize less than 50% to all other known types. HPV-27 has subsequently been detected in cases of anogenital lesions.

After sequencing 1535 bp of the HPV2c genome which includes all or part of the L1, E6, LCR, and E7 regions, it was determined by Chan et al. (Virology 201, 397-398) that HPV2c and HPV2a differed by 12%. In contrast, there was complete agreement between HPV2c and HPV27 over the sequenced region of HPV2c. Based on these results, Chan et al. believe that HPV2c should not be considered a subtype of HPV2a but either identical to HPV27 or as a variant of HPV27. HPV2c was first described as being very prevalent in common warts. Thus, HPV27 should be considered to be a significant etiologic agent of common warts and may be the most common type associated with this pathology.

BASE COUNT 2003 a 1782 c 2023 g 2015 t
 ORIGIN 98 bp upstream from beginning of E6 cds
 1 tatgtggttt aTAAatata actataatcc tttatttaaa aataggggtgt aACCGAAAAC
 E6 orf start -> -> E2 bind
 61 GGTccgACCG AAATCGGTcg tataaaaaca ggagcaggAT Gcgacaagg gcagggatgt
 -> E2 bind E6 cds ->
 121 cagaagagaa tccatgcctt aggaacatct ttttgctttg caaacagtat ggtctggagc
 181 tagaggattt gagattgctc tgcgtgtatt gcagacgagc gctttcagac gctgatgtat
 241 tggcatttgc aataaaaagaa ctgtctgtag tgtggagaaa gggcttccct tttggagcat
 301 gtggaaaatg cctgattgca gcaggaaaac ttagacaata cagacattgg cattactcat
 361 gctacggaga cacagtggag accgagacag gaatacccat ACCTCAGCTG Tttatgagat
 -> E2 bind
 421 gctatatctg ccaTAAgccc ctgagctggg aggagaagga ggcattactg gttggaaaca
 E7 orf start ->
 481 agcgattcca caacatatcc gcccggtgga cgggacactg catgcagtgc gggtaaacAT
 E7 cds ->
 541 Gcacggcacc cgaccagcc tcgcgacat tacatTAAta ttggaagaaa taccgaaat
 <- E6 end
 601 tattgacctt cattgagcag agcaatttga cagctcagaa gaagagaata accatcaact
 661 gacagaacca gctgtgcagg cctacggggg gtaacaacc tgctgcaagt gcggcagagc
 721 cgtccggctg gtggttgagt gcggaccaga agacataaga gatctggaac agctgtttct
 781 gaagacgctg aatcTAGtgt gccccactg cgcgTAGcgt tATGgaggat tccgaaggta
 E1 orf start -> E1 cds ->
 <- E7 end
 841 ccgacgggac agaggaggac gggtgccggg caggagggtg gttccatgtg gaggccatta
 901 taacacacgg ccagaggcag gtatccagtg acgaggatga ggactgcaca gaaacagggg
 961 aggatgtaga cttcatagac aatagggttc ccggagatgg gcaggaaatt cccttcagc
 1021 tatatacaca gcaaatcgca caggatgacg aagcaacagt gcaggcccta aaacgaaagt

```

1081 tcgtggctag tcctttgtct gcatgctcat gcatagagaa tgatttaagt cccagattag
1141 atgcaatctc cctaaacaga aagtcagaaa aggcaaagcg gcgcctattc gagacagAAC
1201 caccagacag tgggtatggc aatacgcaga tgggtgttgg aacACCAGAG GAGGTAacgg
                                -> E2 bind
1261 gggatgagaa cagcgaaggg gggcgcccg tggaggataa ggaggaggag cgtcaagggg
1321 gagacggaga ggcagaccta actgtacaaa ctccacagtc aggaacagat cggcggggta
1381 gcgtgctgac ctactgaga agtagcaatc tgaaggcgac gttactgagt aagtttaagg
1441 aactgttcgg ggtgggatat tatgaactgg taagacagtt caagagcagc aagacagcat
1501 gtgcagactg ggttgtctgt gcctttggcg tgtactatgc tntagcagag ggtatcaaac
1561 aattaataca gccacatacg caatatgcac acatacaggt actgacctgc tcgtggggca
1621 tgggtgctct tatgctgcta cgatacaact gtgcaaagaa cagggacagc gtgtccaaga
1681 acatgagcat gctgttaaac atccctgaaa agcatatgct catagaacca ccaaaactaa
1741 gaagtacccc tgctgccttg tattggtaaca agacggctat gggtaacgga agtgagagt
1801 atggggaaac accagaatgg attgtaagac agacgttggg aggacatagt atggaagatg
1861 agcagtttag actatctgtc atggtaacaat ttgcatacga ccatgacatt gtagaggaaa
1921 gtgtgctggc attcagatag gcacagctcg cagatgtaga tgccaatgca gcagcatttc
1981 taacacgtaa ctgccaggcc aagtacgtga aggacgcagt gacaatgtgc agacactaca
2041 aacgtgcaga gagggcacag atgagtatgt cccagtggat aacattcagg ggaataaag
2101 tattagagga aggcgactgg aagcccatag taaagttttt aaggcaccaa ggggtagagt
2161 ttgtgtcgtt ccttgccgccc tttaaattat ttttaaaagg cgtgccaaag aaaaattgta
2221 tagtgtttta tggacctgca gacacaggca agtcatattt ctgcatgagc ttgttgcagt
2281 tcttaggtgg cgctgttatc tcatatgcta attccagcag ccatttttgg cttcagcctt
2341 tatctgatag taagataggg ttactggacg atgcaacacc ccagtgttgg agctatatag
2401 acacatattt aaggaatttg ctggatggaa acccagtaag catagacaga aaacataaaa
2461 cctgctgca gcttaaatgt ccacccttga tgattacaac caacattaat ccccttgagg
2521 aggacagatg gaaatatttg cgcagcaggc taacactggt tacatttaac aatccatttc
2581 cttttgcaag cccgggggaa cccctgtatc caataaataa tgcaaactgg aatgctttt
2641 tccaaaggtc gtggtcccgc ttagaccTAA acagtccaga ggagcaggac gacaATGgaa
                                E2 orf start ->                                E2 cds ->
2701 aactagcga accgtttaga tgcgtgccag gagacgttgc tagaacttta TGAaaaagat
                                <- E1 end
2761 agcaacaaac ttgaggatca gattaagcat tgggcgcagg ttcggctaga aaatgtcatg
2821 ctgtttaagg cccgggaatg tggaaatgaca cgagtcggct gtacaacagt gcccgccctc
2881 accgtgtcaa aggctaaggc atgtcaggcc atagaggtag agctggcatt acagacattg
2941 atgcagagtg cctatagcac ggaggcatgg accttacgag acacgtgtct ggagatgtgg
3001 gatgcacctc caaagaaatg ctggaaaaag aaaggactgt cagtattagt gaaatttgat
3061 ggcagcagtg acagagacat gatttataca ggctggcacc acatatatgt gcaggatgct
3121 aacgatgaca cctggcaciaa ggttccaggg aaggtggacg aactgggatt atattatgag
3181 cacgatggtg tgcgtgttaa ttatgtggac tttggaactg agtccTGAc ctATGgggtc
                                E4 cds ->
                                E4 orf start ->
3241 actgggacgt gggaggtgca cgtggctgga cgtgttattc accatacatc tgcactctgtg
3301 tctagcacc aggccaccgc ctccgacgac gaaccactat cccctattag acttgcctgta
3361 tccccagtc cagccccagc atcagcagca tcagcaagaa caggaacagc tccgcccaca
3421 aacttgctgt gcaccgcccc ggcgccaccg agtccgcccg ccaagcgcca gcgggtcatc
3481 gtcggacagc agcatctccg gcccgactct acgcgaacgg tcggagaggg gcaagtggag
3541 tgttacgaca aaaggagcat cagtaaacact aacagcacag ctccccggtg ggaccacggg
3601 gacactgaca ctgtgcctgT AAtccatttg cgaggtagtg caaattgttt aaagtgcctc
                                <- E4 end
3661 cgatacaggg tacaaaaaca taaagacaag ctgtatgaca ggggtgcctc cacgtggcac
3721 tgggccggtg gaaagtgtga taagacagcc tttgtgacag tttgggtacac cagtgtggag
3781 cagcgtaaag agttcctgac aagggtcaac atACCTAAGG GGGTgatagc attgccaggg
                                -> E2 bind
3841 tatatgtctg cctttgtaTA Gacctacatg cctgtataga acgtatggtc cagaacattt
                                <- E2 end
3901 caaggcctgc ttagtcaaca cagccctgga ctacttcctt tgcgtggtgg caggggtggac
3961 acatctgctt gtgctactgc tattcctttg gctctctcaa ctaactgctc ttgtggcctt
4021 tttgggtgtc tttttctgtg tctatctagg gttgtggttg atatacgtgc aggccttttg
4081 gtttttacia taattgtgac atatcgccac catttgctgc taacctgtat atatagtttt
4141 ACCACTTGTG TTatttgcaa tgtaccctgt tgtgtataaa gggtcggagg ggacatatcc
                                -> E2 bind
4201 tgtgatattg tggggctggt atgatattga ctgtctgttg gtgattctca cctcatttgg
4261 cttattattg ttggtgtttt atgtccggtt gatccaacac acctaacacc cctcaccctt
4321 ttTGAtacat tatctctttt catacattgt tttttttgta cttgctgctg tgtaataaac

```

HPV27

```

L2 orf start ->
4381 ttgcaccAT Gcctcgtgcc aagcgtcggg aacgcgcctc cccaccgac ctctatcgta
      L2 cds ->
4441 catgcaagca ggcaggtacc tgccccccag acattattcc aaggctagaa caaaacacct
4501 tggcagataa aatattaag tggggcagcc taggggtcct ctttggcggc cttggtatag
4561 gcaactggcag tggcacgggg ggacgtaccg ggtatattcc tgtaggtacc aggccgACCA
      ->
4621 CTGTGGTTGa tattggtgtg gacccaagc cacctgttgt tattgaacct gtgggggcct
      E2 bind
4681 cagagccctc tatagtcact ttggtagagg actccagcat cataaatgca ggcgcttccc
4741 atcccACCTT CACTGGTaca ggtgggttcg aggttacaac ttccaccgtt acagaccocg
      -> E2 bind
4801 ctgtcttggg taccACCCCC TCAGGTacca gtgtgcaggt cagcagcagt agctttttga
      -> E2 bind
4861 acccactata cactgaacc gcaattgtgg aggctcctca aacaggggag gtatctggcc
4921 atgtacttgt tagtacagcc acttcagggt ctcacggcta tgaggagata ccaatgcaga
4981 cctttgctac gtcaggtgga agtgggtcaag agcccataag tagcacacc cttcctggcg
5041 tgcgtagggt tgcagggccc cgcttgtaaa gtagggcaaa tcagcaggtg caagtcaagg
5101 atcctgcgctt tctggaaaagg cctgctgatt tggtaacatt tgacaaccct gtgtatgacc
5161 cagaggagac cataatattt cagcatccag actttcatga gccaccggat cctgatttct
5221 tggacattgt ggctttgcat cgtcctgccc ttacgtccag acaaggcaca gtcctgttca
5281 gtagattggg acgcagggcc acacttcgCA CCCGTAGTGG Taaacagatt ggggctcggg
      -> E2 bind
5341 tgcacttcta tcatgatatc agccctgtcg tccctgatga attggagatg gagccattgt
5401 taccocccagc ttctactgTA GgatcagATG ttttatatga tgtttatgct gatcctgatg
      L1 cds ->
      L1 orf start ->
5461 tctcgcagcc attggacgat tactaccag cccctcgcgg ctccctggct aataccacgg
5521 tatctgcctc ctctgcatct aactgcgag ggtccactac agccctctc tggggtggtg
5581 ttgatgtgct tgtgtatact gggcctgata ttgaaccacc cgttgttctt ggtttggggc
5641 ccctcattcc tgtggctcca tcgttgcctc cctctgttta catatttggg ggggattatt
5701 atttactgcc aagttatata ctgtggccta aacgcagtaa acgtgtcaac tatttctttg
5761 cagatggcct tgtggcggcc TAAtgaaagc aaggtatacc tacctccaac acctgtttca
      <- L2 end
5821 aaggatgatca gtacggatgt ctatgtcacg cggacgaatg tctattacca tgggtggcagt
5881 tctaggctcc tcaactgtcg ccaccatata tattctataa agaagggtag caataatagg
5941 ttggcagtg ctaagggtgc cggctaccaa taccgtgtat ttcacgttaa gctgccagat
6001 cccaataaat ttggcctgcc tgatgctgac ctatatgatc cagacactca aagactactg
6061 tgggctgctg tgggagtaga ggtgggcccga gggcagcctt taggtgtggg tgtgtctgga
6121 caccatattt ataataggca ggatgacact gaaaatgcac acacacttga ttcagctgag
6181 gatggaaggg aaaatatttc catggattat aaacagacc agctgtttat tcttggctgt
6241 aaaccctcta ttggtgagca ctgggtccaag ggaaccacct gtaattgggtc ctctgctgct
6301 ggtgactgcc cgcccctcca gtttactaat tcaactattg aggatgggga tatggttgaa
6361 acagggtttg gtgcattgga tttcgccact ctgcagtcca ataggctcga tgttctttg
6421 gatatttcta caaacgtctg taaatatcca gattacctga aaatggctgc agagccttat
6481 ggtgattcca tgttcttttc gctgcgtaga gaacagatgt tcaccogtca tttttcaat
6541 agggctggtg agatgggtga cacaatccca gatgagttgt acattaaaag taccactatc
6601 tgggaccocg gcagtcattg gtatacctcc actcctagtg gctctatggt gtccctctgaa
6661 cagcaattgt ttaataagcc ctactggcta cggagggccc agggacataa taatgggatg
6721 tgctggggca atcggatctt tctgactgtg gtggacacca cacggagtag caatgtctct
6781 ctgtgtgagc ctgaggtgtc tgataatact aattataaag ctacgaattt taaggaatac
6841 ctcaggcata tggaggagta tgatttgcag ttcattttcc aactgtgcaa aataaccctc
6901 actcctgaga taatggccta cacaataat atggatcccc agttgttggg ggactggaac
6961 ttgggtgtac ccccccgcc gctcgcaggt ttgcaggaca cttatagata tttgcagtc
7021 caggctatta cgtgtcagaa acctacgccc cctaagacc ctacagatcc ctatgccaac
7081 atgaccttct gggatgtgga cctacgggaa agtttttcta tggatctgga ccaatttctc
7141 ttgggtcgca agttcttatt gcagcggggg acgacgcccc ccgtgtctcg aaaacgacc
7201 gctgttgggc gcggccacTA Gtaaacgcaa acgggtgagg cgttacgtgt gagtgtctcg
      <- L1 end
7261 ataatttctc ctgtctacct tttacataat atttggtgtt gtttgtgctt atgtttgtgt
7321 tgttgtttgt acgtatgtta catgtataca tgtatggtat gtatcccctc ccgtatgaat
7381 aaacgtgtgt catgtgttgt gtgttctgta actgttcttt tgggtgcacg ggtttctgca
7441 ccctattgtc ctttgtgtag cccccagttt catgcgACCG TTTTCGGTtg cgtgcagttt
      -> E2 bind

```

```
7501 cggtcggcgc cgttgccagc acagcttaat cctttaattg ctcacatcct aaagtgttag
7561 ctgtgccagc aacaatgagt ttggatTTTT ggttgTTTaa tgctTTTTct ttttagTTTT
7621 tcctttctct gtgccaggcg cgagaggggtg tgtgcattcc taggctgatt atggttctgt
7681 gttggcacag atttctactg cgtctgcagg aaaacctgca gcaacagcac tttgggcgcg
7741 tcgtttctgc agccaacttt cacttgcaa cttgtcttgc cgcgcattcc aagaaacacA
                                     E2 bind ->
7801 CCTATTCCGG Tcgcaatgtc tac
```

HPV28MY911

LOCUS HPV28MY911 449 bp ds-DNA VRL 16-OCT-1994
DEFINITION Human papillomavirus type 28 (HPV-28), partial L1 cds, My09/My11 region.
ACCESSION U12502
SOURCE Human papillomavirus type 28 DNA.
REFERENCE 1 (bases 1 to 449)
AUTHORS Bernard,H.-U., Chan,S.-Y., Manos,M.M., Ong,C.-K., Villa,L.L., Delius,H., Peyton,C.L., Bauer,H.M., and Wheeler,C.M.
TITLE Identification and assessment of known and novel human papillomaviruses by PCR amplification, restriction fragment length polymorphisms, nucleotide sequence, and phylogenetic algorithms
JOURNAL J. Infect. Dis. (1994) In press
COMMENT HPV-28 was originally isolated from butchers' warts (Favre et al. J. Virol. 63, 4905). Subsequent tests revealed it to be present in 7 of 130 butchers suffering from warts, in 3 of 66 wart patients and in 1 of 55 immunosuppressed patients. It has been detected in flat warts and intermediate warts. It is most closely related to HPV-3 and HPV-10.

Cloned HPV-28 DNA was obtained from the Papillomavirus Reference Center, Heidelberg and subsequently sequenced by Dr. H. Delius over the L1 MY09/MY11 segment. HPV-28 and the several other HPV types recently sequenced over the MY09/MY11 primer region by Dr. Delius were used as type-specific probes to screen DNA for novel genital HPV types. The screened DNA was obtained from four recent epidemiological studies [1]. Primer regions are annotated in the sequence; information in this region is not accurate due to primer degeneracy.

BASE COUNT 123 a 95 c 95 g 136 t
ORIGIN
1 gctcagggac acaataatgg tatctgttgg gccaaccaat tgtttgtaac tgtagtggat
L1 cds ->
-> MY11 PCR primer <-
61 actacacgca gtacaaacat gacgttgtgt gtttctactg actcttcagc tacgtacgat
121 gctagtaaat ttaaggaata ctaaggcac ggggaggagt acgatttgca gtttatatc
181 cagttgtgta aagtaacctt gaccctgat attatggcat attacatac catgaacaat
241 agtttattgg aggactggaa ctttgggttg actttaccac catccactag cttggaggac
301 acgtataggT tcatatcttc ctctgccatt acctgtcaaa aggatgcttc cccactacc
361 aaggaagacc cttacgctaa actaaacttt tgggaagtgg atcttaagga tcgcttttct
421 cttgatctat cgcaattccc tctgggaag
L1 cds ->
-> MY09 PCR primer <-

LOCUS HPV29MY911 455 bp ds-DNA VRL 16-OCT-1994
DEFINITION Human papillomavirus type 29 (HPV-29), partial L1 cds, My09/My11 region.
ACCESSION U12503
SOURCE Human papillomavirus type 29 DNA.
REFERENCE 1 (bases 1 to 455)
AUTHORS Bernard,H.-U., Chan,S.-Y., Manos,M.M., Ong,C.-K., Villa,L.L., Delius,H., Peyton,C.L., Bauer,H.M., and Wheeler,C.M.
TITLE Identification and assessment of known and novel human papillomaviruses by PCR amplification, restriction fragment length polymorphisms, nucleotide sequence, and phylogenetic algorithms
JOURNAL J. Infect. Dis. (1994) In press
COMMENT HPV-29 was originally isolated from a skin wart. Subsequent tests failed to detect it in cutaneous wart specimens from 119 patients, including both butchers and immunosuppressed patients. These latter were included in the study on the basis of HPV-29's similarity to members of the group including HPV-10 and HPV-3, which are frequently detected in such patients. It thus seems to be a relatively uncommon member of this group (Favre et al. J. Virol. 63, 4906).
Cloned HPV-29 DNA was obtained from the Papillomavirus Reference Center, Heidelberg and subsequently sequenced by Dr. H. Delius over the L1 MY09/MY11 segment. HPV-29 and the several other HPV types recently sequenced over this region by Dr. Delius were used as type-specific probes to screen DNA for novel genital HPV types. The screened DNA was obtained from four recent epidemiological studies [1]. Primer regions are annotated in the sequence; information in this region is not accurate due to primer degeneracy.
BASE COUNT 132 a 92 c 99 g 132 t
ORIGIN
1 gcgcagggac acaacaatgg tatatgctgg gccaatcagg tatttttaac tgtgggtggac
L1 cds ->
-> MY11 PCR primer <-
61 accacacgca gcaccaatat gtcgttgtgt gctaccacag agtctcaacc gttgaccact
121 tatgatgcta ccaagattaa agaataattg agacatgggg aggaatatga tttgcagttt
181 attttccagt tgtgtaaagt tacattgaca cctgaaatta tggottacct tcatactatg
241 aacagtgccct tacttgaaga ctggaatttt ggattgacat tgccaccttc cactagcttg
301 gaagacacgt ataggtttgt aacatcctct gccataactt gtcaaaaaga tttggccctt
361 acagaaaagc aggatccgta tgcaaagcta aatttctggg atgtagattt aaaggataga
421 ttaccctgg atttgcaca gtttcccctg ggacg
L1 cds ->
-> MY09 PCR primer <-

HPV32

```

LOCUS      HPV32          7961 bp ds-DNA          VRL          04-OCT-1993
DEFINITION Human papillomavirus type 32 (HPV-32), complete genome.
ACCESSION  X74475
SOURCE     Human papillomavirus type 32 DNA.
REFERENCE  1 (bases 1 to 7961)
AUTHORS    Delius,H. and Hofmann,B.
TITLE      Primer-directed sequencing of human papillomavirus types
JOURNAL    Curr. Top. Microbiol. Immunol. 186, 13-31 (1994)
REFERENCE  2 (bases 1 to 7961)
AUTHORS    Delius,H.
TITLE      Direct Submission
JOURNAL    Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H.
           Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im
           Neuenheimer Feld 506, W 6900 Heidelberg, FRG
COMMENT    HPV-32 was originally isolated from a lesion of oral focal
           epithelial hyperplasia by Beaudenon et al. in 1987 (J Invest
           Dermatol 88: 130-5) and was subsequently sequenced by Dr. H. Delius.
           It has been detected in a significant number of oral papillomas.
           When 113 benign and malignant tumors of the ororespiratory system were
           analyzed for the presence of HPV-32 DNA, 9% of the oral papillomas
           were positive for HPV-32 (de Villiers et al. Laryngol Rhinol Otol
           (Stuttg) 65: 177-9).

BASE COUNT 2390 a 1533 c 1729 g 2309 t
ORIGIN     101 bp upstream from beginning of E6 cds
           1 taatctttga attataaaaa agtagggagg aACCGATATC GGTttaACCG AAAACGGTgc
                                     -> E2 bind -> E2 bind
           61 atatataaac caccctgggc agtggctcctt gtTAAGgcag aATGgcaagt acttctgcct
                                     E6 orf start ->
                                     E6 cds ->
           121 catcacagcc aagtacatta taccaattgt gcaaagattt tgggctgacc ctgcggaatt
           181 tacaatctg ctgtatttgg tgtaaaaacc acttaaccag tgctgaagcg tatgcataac
           241 attttaaaga ttgcacgta gtgtggaaga aaggctttcc atatgccgcc tgtgccttct
           301 gcttagaatt ttattctaaa gtgtgtgcac tggcagacta cgacagatca gcattttggc
           361 atacagtaga acaagaaaca ggactactgt tggaagaaca aataattcgc tgtgctatat
           421 gtcaaaagcc tttatcgcca agtgagaaag atcatcatat ttataacgga cggcatttca
           481 gattcatttt aaaTAGgtgg acgggtcgct gtaccagctg cagagaaTAA TGcgtggaaa
                                     E7 orf start ->
                                     E7 cds ->
                                     <- E6 end
           541 cgcaccaacy ctaaaggaca ttattttcta tgacctgcca acgtgtgacc cgacaacgtg
           601 cgacacaccg ccggttgACC TGTATTGTTa tgaacaattt gacacctcag atgaagatga
                                     -> E2 bind
           661 tgaagacgat gaccaaccta taaaacagga catacaacgt tacagaatag tgtgtggttg
           721 tacacagtgt ggacggtcag ttaaacttgt tgtcagTAGt acaggcgcg acatacaaca
                                     E1 orf start ->
           781 gctgcatcag atgcttctgg acacactggg cattgtgtgt ccattgtgtg cctgcgtgga
           841 gTGAActgcaA TGcggatga tacaggtaca gaggaggggc tagggtgttc tggttggttt
                                     E1 cds ->
                                     <- E7 end
           901 tctgtagaag caatagtaga aaggactaca gaaaatacta tatcagacga tgaggatgaa
           961 aatgtagagg acagcgggtt ggacctcgta gattttgtag acgacagcag aataatacct
1021 acaaatcaAT TAAAggcgca ggcattATTA AAtaggcaac aagcacatgc agataaggag
           signal -> signal ->
           1081 gcagtacagg cactaaaacg aaagttatta ggcagtccat atgaaagtcc cgccagtgat
           1141 ttacaggaga gcataaaca agagctaagc cctaggcttg gtggattaca gttatgtcgg
           1201 gggcccccaag gggccaaaacg acgactatth caatcattgg aaaaatcgca cagtggatat
           1261 ggctattctg aagtggaat acggcaggaa caggtagaaa atggacatgg cgccccagac
           1321 gggagtatgg gtaacggggg gggcatgggc agtgtacatg gggcgacgaa aaatcaggaa
           1381 ataggcacia atacgcctac aacaaggggtg gtggaattgc ttaagtgtaa gaacttgcaa
           1441 gcaacattgt taggtaagtt taaagagctg tttggattgt catttgggtga ttagtaaga
           1501 caatttaaaa gtgacaaaag tagttgtaca gattgggtag ttgcagcatt tgggtgacat
           1561 catagtattg cagaaggctt taatacacta attaaagcag aggcgttata tacacacata
           1621 caatggttaa cctgtacctg gggaatggta ttgttaatgc taattagatt taaatgtggc
           1681 aagaatcgca ccacagtgtc taaaggaatg tgcaactat taaatatacc tgctaatacaa
           1741 ctgttaatag agccgccagc attacaaagt gtggcagcag caatatattg gtttcgagca

```



```

1801 ggaatatcta atgccagtgt agtaaccggg gaaacaccgg aatggatata aagacaaaaca
1861 attgtagaac attgctttgc agatacacag tttaatTTaa cagaaatggt gcaatgggca
1921 tatgataatg atttAACCGa agatagtGac attgcatatg aatatgcccc acgtgctgac
1981 acagatagta atgcagctgc atTTTTAAAA agcaactgtc aggcaaaata tgtaaaagat
2041 tgtggaatta tgtgtagaca ctataaaaaa gcacaaatga aacgtatgtc aatgccacag
2101 tggattaaac atagaagtga aagaactggc gataatggtg attggagacc tatagtaaaa
2161 tttattagat atcaaggaat agattTTTTa acattttatg ctgcattcaa aaaatttttg
2221 cataatatac caaagaaaag ctgtttagta ttaattggtc cgccaaatac aggaaaatca
2281 cagtttgtaa tgagttagt aaagtTTTTa gcaggaacag taatttcctt tgtaaattca
2341 cacagccatt tttggttgca gccattagac agtgcaaaaa tagcaatgct agatgatgca
2401 acacctccat gttggacata cttagatata tatttaagaa atttactaga tggcaatcca
2461 tgcagtattg atagaagca taaagcatta acagttgta aatgtccacc attaataata
2521 acatcaatac cagatattag aacagaagac agatggaaat acttatatag tagaattagt
2581 ttgtttgaaT ttccaaacc atttccatta gataaaaaat gaaatcctgt atatgtgta
2641 aatgatgaaa attggaatc atTTTTTcaa aggttgtggt ccagctTAGa atttcaagaa
                                     E2 orf start ->
2701 tcagaggacg aggaagaaaA TGgagacact ggccaaacgt ttagatgctg gccaggaaca
                                     E2 cds ->
2761 gttgtagaaa ctgtaTGAgg aagatagtaa acatttagaa aaacatgtgc agcactggaa
                                     <- E1 end
2821 gtgtttacgc atagaagcag ccttattatt taaggctcgt gaaatgggtt atgcacaagt
2881 aggacatcaa atagtgccag cactggaaat atccagggcc aaggccacg ttgcaattga
2941 aattcaattg cgttagagca cattattgca gtccacattt ggtacagaac catggacatt
3001 gcaagagaca agttatgaaa tgtggcatgc ggagcccaa aagtgttaa aaaaacaggg
3061 acgcactgtg gaggttgtat tctgagaaT tccatggaaa tccatgagaa gcaatgcatt atacagcatg
3121 gacatttata tacgtgcaaa cactagatgg cacatggtgt aaagtatacg gacacgtatg
3181 ctatgcagga ctatactaca ttgtggacaa catgaaacag ttttattgta actttaaaaa
3241 tgaggcaaaa aaatatgggg taaccggaca atggggaggt catgatggca ctcaggTGAT
                                     E4 orf start ->
                                     NH2 terminus unknown
3301 tgtttctcct gcatccatat cttagcaccac aaccaccgaa gcagaggtat cctcttctgg
3361 acttactgaa ttggtacaaa ccaccgacct atacaacacc acacctacac ccacaacct
3421 cacaaggagt aactgcgacc cagacggcac agacggaata ttatacaaa agcccccccc
3481 gaccaccccg ccgcgaaaac gataccgaca gtctttgcag ccaccaaca agcacctgca
3541 gcaactcggc gtcacaaacg taccgtgga ccccgatcg caaagggtca catctgacaa
3601 taacaataac cagcgacgga acccgtgtgg aaatcagact acgcccagTAA tacacttaca
                                     <- E4 orf end
3661 agtgatcctt aattgcctaa agtgtttaa agtgagggtta aagaaaaatt gttctcactt
3721 atttactcaa gtgtcatcca catggcacct gacagaaaaa gactatacac gtgactcaaa
3781 ggatggtata ataacaattc attattataa tgaggaacaa cgagataagt ttttaagtac
3841 tgtaaaactt cctcctggta ttaaactctg catgggttac atgtcaatgt tacaatttat
3901 gTAGTgtgta catatgtaca aacatataa ggggaacatt actgtgagtc cactattgtg
                                     <- E2 end
3961 tgggacaacc ggccagacac tgctctatt attgtttata gttgttctg cgtgtgtgtt
4021 gtgtgtgtgg attagttaa tagaccccc atatccctg tgggcctcct gtcttctgtag
4081 ctatctaata ttagtactat tgcctctggt gcagctacta acatctgtgc aattttttt
4141 tttagctttg cttgttgttg tgttctctgc ctttttacta accctatata tacactttgc
4201 actacattaa tgtagaaggt gtacatacaa tagtatagct gtgtgtgtgt ggtgtgcatg
4261 gtatacaca taattgtaca tacgtgataa tatggtacat tgtactttag aaaatggcga
4321 tgtgtgctgc ctgctgtggt ttttagccac agtgtttatt tgtatattag tactgtgtT
4381 AAactactgc agacatacgt gtgtacattg ctgtaataa agttgggttt ttgtattgtt
L2 orf ->
start
4441 tgtaccaaag tgtattatca ATGccaccac accgaagtgc cagacgcaaa cgggcatctg
                                     L2 cds ->
4501 ctacacagtt atatacaaa tgtaaggcct ctgggacatg cccaccgat gtaattccta
4561 agattgaagg tgcacgtggt gcagatcaaa tattaaaaat gggcagcact ggtgtgtttt
4621 ttgggggctt tggcataggt accggtgctg ggtcaggtgg gcgcacaggg tatgtgctta
4681 taggaacacg accgctgttt gttgcagac cgggcctctg aatacgtcct ccagtgtgtg
4741 ttgacactat tgggcctact gaccatctg taatttcctt attggaagag tctgcagtta
4801 ttgattctag cataccagct cctaccgaca cgtctcatgg tgggtttaa attacgtcct
4861 ctgcaagtgg ccgctcctct acacctgctg tattggacat ttctcctccc accaatacta
4921 taagggttgc atctactacg tccacaatc cagtatacag tgatcctttt actttacggc
4981 cttctttgcc agtagaggtt aatggtagac tcttaacatc ccaccaaca atagcccctc

```

HPV32

```
5041 attcgtatga ggaattcct atggacactt ttgtggtatc cacagataca agtaaactg
5101 ttaccagtac ccctattcct ggacctcgcc ctactatgcg ccttggttta tataaccaggg
5161 tcactcaaca acgtccagtt gctactacaa catttttaac atctcccgag cgtttggtaa
5221 cctatgacaa ccctgcatat gagggctctg ctgaggggtac gttggaattt gaacatccca
5281 ccattcatga ggctcctgat tctgatttta tggatattat tgcattacac cgtcctgtgc
5341 tatctgctag gcagggcact gtccgtgtca gtcgtattgg gcaacgggct tccttgcaaa
5401 cacgtagtgg ggctcgtatt gggcttaggg tacatttttt ccatgatatt agcccaatca
5461 ctaggccatc agaggctata gaattgcaac ctttaggctc ctctccaca gctgtatcta
5521 ctactgcttc atccgcaatt aatgatggcc tgtttgatgt ttatggtgac cctgacatac
5581 ctccctcaca cgcgttaccg cccctacggc ccccccacaca cgtgtcaact gtttctttaa
5641 ctagtcttgg tagtgttcct gcacaaactg caaatacaac tgtccctctt tccttaccta
5701 caaatattaa tgtaggcctt gacctttcac ctctgagtc tccgcccttt attagtacac
5761 gtccctgatac accttctttt gactctgtta tgggtATAGg atgggatttt atattgcatc
                    L1 orf start ->
5821 ccagttatat gtggcgtaag gcgcgtaaac ctgtaccata tttttttgca gATGtccctg
                    L1 cds ->
5881 tggcggcccTA Gtgacaacaa ggtttatctg cctcctctc ctgtttccaa ggtggctcagc
                    <- L2 end
5941 acagatgaat atgtgcaacg taccaactac ttttatcatg ccagcagttc taggcttttg
6001 gctgttgggc atccatatta tactattaag aagacacca atagaacatc tattccaaag
6061 gtgtctggat tgcagtatag agtatttagg gtttaggcttc cagaccctaa taaatttaca
6121 ttacctgaaa caaacttata taatcctgaa acacaacgta tgggtgtgggc ctgtgtgggt
6181 ttggaggttg gccgtggaca gccttttagg gttggtotta gtgggcatcc ttatataaat
6241 agattggatg atactgaaaa tgggcctaga tatgctgcag ggctggaac tgataataga
6301 gaaaaatgat ctatggattg taaacaaacc caattgtggt tgggtgggttg taaacctgcc
6361 attggcgagc attggggtaa gggtgctgct tgctctgcac aatcaaatgg cgactgcca
6421 cctttggaat tacaaaacag tgttattcag gatgtgata tggcagatgt agggtttggga
6481 gcaatggact ttagtgctt gcaaaactca aaagctgagg tgccattaga tattatgaac
6541 tccattagta aatctcctga ctatttaaaa atgctgcag aggcctatgg cgacaatatg
6601 tttttctttt tgagacggga acaaatgttt gttcgtcact tgtttaatag ggcaggaacc
6661 cttggatgaac ctgttctgga ggacatgtat ataaaagctt ctaatgggtgc ttctggcaga
6721 aataatthag ctagtagtat ttattatcca actcccagtg gttctatggt cacctctgat
```

```

6781 gcacaaatat ttAATAAACCC ATATTGGTTa cagcaggcac aaggccacaa taatggtata
      signal -> <-
              -> E2 bind
6841 tgttggggta atcaagtgtt tctaactggt gtggatacta cccgtagtac taacatgact
6901 gtgtgtgcta ctgtaacaac tgaagacaca tacaagtcta ctaactttaa ggaatatcta
6961 cgccatgcag aggaatatga tatacagttt atatttcaat tgtgcaaaat tacattatct
7021 gtagaggtta tgtcatatat ccacaccatg aatcctgaca tactagacga ttggaatggt
7081 ggtgtagctc caccgccctc tggtaacttta gaagatagtt atagatttgt gcagtctcag
7141 gccatacgat gtcaagctaa ggtaacagca cctgaaaaaa aggatccttt ttctgactat
7201 tcattttggg aagtaaatat atctgaaaag ttttctagtg atttagatca gtttccattg
7261 ggtaggaagt ttttactgca agctgggtta cgtgcaagac ctaaacttac agcagtaaaa
7321 cgaacagcat cttccagtca aaagtcttct tctcctgcaa aacgcagaaa aacacgtaaa
7381 TAAatatttt tctctgtgtg atgtgttgtg caacttgtgt gtgggtgtgca tgtattgtgt
      <- L1 end
7441 atatgtgtgt ttcggtgtgc atattaataa agtatgattg tgttcattgt atgttgttgt
7501 accaccattt tatttttcaa tcctccattt tagtacatgc aACCGAAATC GGTtgcctgt
              -> E2 bind
7561 ggataatgtc cattaatatt agcaagcaca tgcagttttt actgccaat atatgtactg
7621 ccaaatggta ttgctaagta gcaaaatggt tttacataca taacatacac gcccttttgc
7681 acaagcatgt tttagaaagg ttggcatagc tttgcattta cttatttctt tcttcttttg
7741 ttcattgtat gactactgta tttgttatgt ataattaaaa tgcttttagg cacatatttg
7801 tgggttttgg cacagtactt tacaagttac tctggcctag acaagactag tgctttgtca
7861 tgcccttatat aactaagtat tagtcactga ctgcatctaa ttaaactgtc aACCGAAATC
              -> E2 bind
7921 GGTgcataaa tcctaacttg ttctgttgtt attataaagt a

```

HPV40

```

LOCUS      HPV40          7909 bp ds-DNA          VRL          04-OCT-1993
DEFINITION Human papillomavirus type 40 (HPV-40), complete genome.
ACCESSION X74478
SOURCE    Human papillomavirus type 40 DNA.
REFERENCE 1 (bases 1 to 7909)
AUTHORS   Delius,H. and Hofmann,B.
TITLE     Primer-directed sequencing of human papillomavirus types
JOURNAL   Curr. Top. Microbiol. Immunol. 186, 13-31 (1994)
REFERENCE 2 (bases 1 to 7909)
AUTHORS   Delius,H.
TITLE     Direct Submission
JOURNAL   Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H.
          Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im
          Neuenheimer Feld 506, W 6900 Heidelberg, FRG
COMMENT   HPV-40 was originally isolated from a biopsy taken from a
          penile intraepithelial neoplasia in a 62-year-old male patient
          and has subsequently been sequenced by Dr. H. Delius. The
          lesion was characterized by some nuclear atypia, acanthosis
          and parakeratosis and proved to be resistant to therapy. It has
          also been detected in bowenoid papulosis lesions and condylomata
          acuminata. Of the latter, all four were biopsies taken from
          lesions in the vaginal vault of hysterectomized patients (de
          Villiers et al. Virol. 171, 248-253).
BASE COUNT 2242 a 1603 c 1860 g 2204 t
ORIGIN     101 bp upstream from beginning of E6 cds
           1 ttaataacaa ttgtactggt agtaaagggt gtaACCGAAA ACGGTccgAC CGAAAGCGGT
           -> E2 bind -> E2 bind
           61 acataTAAat tccaacccaa aaaacctgct ttggggccag tATGtctgca cggtgcgct
E6 orf start -> E6 cds ->
           121 cccaggccag gacctgtat gaactgtgtg accagtgcaa tattacattg cctacgttgc
           181 aaatgattg tgtgttttgc aagacggtcc taaaaacagc tgaggtactg gcctttgctt
           241 ttagagagtt atagtgtgtg tggcgcgacg actttccaca cgccgcatgt ccacgggtcc
           301 tggacctgca cggaaaagta aaccaataca gaaactttag atacgcagcc tatgcaccaa
           361 ccgtggaaga agagacagga ttaaccattt tacaagtaag gattagatgc tgcagtgcc
           421 acaagccttt gtctcccgtg gaaaaaacca accatattgt aaagaagacg caattcttTA
           481 Aattaaaga ttcgtggaca gggactgtc tacattgctg gaagaaATGc atggagaaag
E7 orf -> E7 cds ->
start
           541 gccacgctc ggagacattg tgtTAAacct gcacctgaa cctgtatgtc taaactgcaa
           <- E6 end
           601 cgagcaatta gacagctcag actcagaaga tgaccatgaa caggaccaac tagacagctt
           661 acacagtaga gagcgtgagc aacccacgca acaggacctg caagtaaatt tgcaatcatt
           721 taaagtagta actcgggtgtg tattttgtca gtgtttggtg cgcttagcag tgcatgttc
           781 catcacTGAT ataacacagt tccagcagtt gctgatgggc acattacata tagtgtgccc
E1 orf start ->
           841 caactgtgca gctacagagT GACAacaATG gcagactctc caggtacaga ggacgggggg
           E1 cds ->
           <- E7 end
           901 gctgggtgct caggatggtt tgtagtagaa gctgtagtgg ataacaaac gggggatgct
           961 gtatcggaag atgaggatga ggaggacata gaggatagtg gatttgatat gatagatttt
1021 attgataata gtgttgtggc agaggaacat gtagaactaa gtaatgcaca ggcactttta
1081 catgtacagc agacatgtgc agatgctgct gacctgtgcg agttaaacyc aaagtacatt
1141 agcccatatg taagtcctat acaatactca gaacctcta tagacgggga cctaagtcca
1201 aggctgcatg caatacggct tggcggcggc caaaaggcta aacggcgggt gttccagcgt
1261 gtggagcaaa gggacagtgg ctatggctat tctgaagtgg aaacaacaga gagacaggta
1321 gagacagaac atggcggacc ggaagatACC GTGGGGGGTa gtgggagggt gaccacagat
           -> E2 bind
           1381 gaggcggaag cagtagaggt tgtggaagac ggcagtcatg ttatagacca ctgtagtccg
           1441 cgcacacaac taatagagct gtttaaatgc aaggacctaa atgctaagct gtagtgtaag
           1501 tttaaagagc tttatggagt ggggtttgga gacctggtaa gacagtttaa aagtgataaa
           1561 tccacgtgta ccgattgggt gtagtctgtg ttcgggggta atcccacat agccgagggc
           1621 tttcatcac tgctgaaaag gcaggcatta tatttacata cccaatggac gtcatgcaaa
           1681 tggggatagg tgttgcttgc attgtgtaga tataagtggt gtaaaaatag ggaacagtt
           1741 gttagacagc tatccaaaat gttaaatgta cctgacaacc agatactggt acaaccgctt

```

```

1801 aaattacaaa gtccgctgc agcattatth tggtttagag caggaatggg taatgggagc
1861 gaggtgtccg gcacaacacc ggaatggata gctaaacaaa ctatgtaga acacagcttt
1921 gctgacacac agtttagcct aacagacatg gtgcagtggg catatgataa cgggcataca
1981 gatgagtgtg aaatagcata ctattatgca caaagagcag acgtggatgc aaatgcagca
2041 gcattcctaa aaagtaacaa tcaggccaaa tacgttaggg attgtgcatc catgtgcaaa
2101 cactataggt tagcagaaat gaggcgcacg tcaatggctg agtggataaa gcatagaggg
2161 gagaagtgtg atgaaggcga ctggaaacct atagttaa atattacgcta tcaacatata
2221 gatataatag ttttttagc tgcattaaaa aaatggctac agggatatacc taaaaaaaat
2281 tgcattttaga ttgtgggtcc tccagacaca ggaagtcac gttttggaat gaggcctgatg
2341 cactttatgc aagggtacaat aatatcatat gttaatcctc gtaccattt ttgggtgcaa
2401 tcattggcag atgcaaaagg agctatgcta gatgatgta cagcggcatg ctgggggtat
2461 atggatacac acatgaggaa cttattagat ggtaacccaa ccagcataga tagaaaaacat
2521 aaaccgttag cagtaattaa gtgcctcccg ttactgttaa cgtccaatat aaatataaca
2581 caggacagta agtaccataa tttacaaaag agggttcaag tgtttgaatt tccaaatcca
2641 tttccatttg acagcaacgg caatgctgtg tatgaattaa atgatgcaaa ttggaactcc
2701 ttttttaaaa ggttggcatc cagttTAGag ctgcaaacac ctggggacga ggATGgagaa
                E2 orf start ->                E2 cds ->
2761 tctagccagg cgcctagatt tgtgccagga acagttgta gaactgtaTG Aacaaaacag
                <- E1 end
2821 caaagagcta cagcaacata tattgcactg gaaatatata cgttatgaaa gtgcaatata
2881 ttatacagca agacaaatgg gcattaaaca tttaggccac caggtgggtc caagtttaga
2941 tgtttcaaaag gctaaagccc atgcagcaat tgaatgcaa atgtgttag aatccttgca
3001 aaacactgaa tataatgtag agccatggac gctgcaggac acaagtcaag aattatggct
3061 tgcagaacct aagaaatggt ttaaaaaagg tggcaaaaaca gtggaagta gatttgactg
3121 caatgaaaca aatgcaatgc attatacact gtggaccaca gtatatgtac aggtggatga
3181 tgcattgaca aaggtaaaag gccaggtgga ctacaaaagg ctctcatata cagtgcacgg
3241 gtgcacaaca tattatgTAG actttgcaaa ggaagcaca acgtATGgga aaacaaatag
                E4 orf start ->                E4 cds ->
3301 gtggactggt attgtgggtt cacacgttat atgttctcct agtactatcg aggaacacgg
3361 actaccattg gttgagactg ctgacgccag acccccgacc actgccaccg acacccccga
3421 cgcccccgcc acagcgacca ccgaaacggg cggccccgcc caggcaccgc ccagaaagcg
3481 acgaagaaac ggacacctgc ccatcaccac tactgtgggc aaatcactcg gaggagagta
3541 cgtggacact gcagacagaa cacgcacgcc tgaccctgaa agcaacaacg ggcacagggaa
3601 ctgtgggtgga ggttcttcta cacctaTAAt acaattagaa ggtgaagcca attgcttaaa
                <- E4 end
3661 gtgttttaga tatagactag gaaaagtgtc acatttattt tgtaattcct caactacatg
3721 gaggtggacc actgaaatcca gaaccgagaa aaatgctata ataacgtaa catatagtag
3781 tgtacagcaa cggctcgtact tcttagctat tgtaaaaaata ccaaaaacaa taaaacatag
3841 cttgggtatg ttaacactga tgTAAtatat gtatatatgt atatatatgg aacaccatct
                <- E2 end
3901 gtgaagggta ttgtctagca cgtaccacac aaccagccaa tctgctgcta ttattgttaa
3961 tagagtctac gcttggcata ctagtagtct atattgtttt ttacttatt gcaactgtta
4021 tctgtgtgtg tgcctcaat ctgtatgtgt ggagctctat actgtctgtc atctccattt
4081 tgtgctttag aacatggggg gcactaacat ctattactaa cctgtttatc ctaatactct
4141 tagtgtggta tttgcctgcg gtggctcctc acacatttat tatatatcac atacagaacc
4201 aattgtaaca tgttgacctg tcaattggat gatggtgata cctggatggc attatggatg
4261 ttacttgcgt ttataactgt gttgttactg ttattgggtg ttcatgttag aaccttatat
4321 ttataTAGgt ataccaagta atactttgtg tagtaaaaaa acattttttt atattacagc
L2 orf start ->
4381 gtcttgtaAT Ggtgtccagc agggccccgta ggcgcaagcg ggcgtcagct actcagttat
                L2 cds ->
4441 atcaaacatg caaggcggcg ggcacctgtc cacctgatgt tgttcataag gtggagcaaa
4501 caaccgttgc agatcaaatt ttaaaatggg ggagcatggg agtatttttt ggocggcttg
4561 gcataggttc cggctctggc acagggggca gggctggcta tgtgcctttg tctacaggtt
4621 cccgggctgt cctcctaaa tctttggtag ctgacgttgt tgctaggcca cctgtgggtg
4681 tggatactgt tgcaccctcc gatccgtcca ttgtatcttt aatgaggaa agtagtatta
4741 ttcagtcagg tgccccgtcc cttactattc caacagaggg tgggttttca gtaacgtcct
4801 ccggtacaga tgtgcctgcc attttagatg tgtcgtccac aaataccgtg catgtgacag
4861 ccaccacaca tcacaatcct gtctttactg atccctcggg tgtgcagccc atcccactg
4921 tggaggccgg tggctgcctc atagtgtcgc attccactat tactactagt gctgctgagg
4981 aaataccttt ggacacgttt gttgtccata gtgatccatt gtccagtaca cccgttctctg
5041 gtacgtctgt acggcccagg ttgggactgt acagtaagge cttgcagcag gtggaaattg
5101 tggaccctgc ctttctgtct accctcagc gtttaattac atatgacaat ccagtgtttg
5161 agaacgtgga tgatacattg cagtttgagc agccatccat acatgacgct cgggaccctg

```

HPV40

```

5221 cgtttatgga catcattacc ttgcacaggc ctgcattaac gtctcggcgg ggtgtcatac
5281 gcttttagcag ggtgggtcaa cggggcacca tgtacacacg gcgagggacc cgcattgggg
5341 gtcgtgtaca cttttttagg gacattagtc ctattggtgc agctgatgat attgaattgc
5401 accctctagt ggctcggca ccacatacac tggagacacc acatacacta gagacaccac
5461 tggacactac tgatgcctg tttgatgtgt atgcagacat ggatactata gatgatgatg
5521 cagcatatgc tacatttcoa ttacatcctg ccgattctac tcgtatatct aacacatcca
5581 tacctcttgc cacggtttct gacacattat taacatctgg tcctgacata gtgtttcctt
5641 ctattcctgc aggtacacca tttttgcctg tgtcaccttc tatacctgcc atatctgtac
5701 taattcacgg tacagattat tttttgcatc ctgcatacta ttttaaaaaa cgccgaaaagc
5761 gcatatTAGc acatcagtAT GtggcaactT AAtgaaaatc aagtatattt accaccgcca
L1 orf start -> L1 cds ->
5821 acgctgttg ctaccattgt tagcacagat gagtatgtgc aacgcaccag tttatattat
5881 catgctggta gtgccaggtt actgactata ggacatccat actttgagtt aaaaaaacc
5941 aatggtgaca tttcagtgcc taaggtttct ggacatcaat acagggattt tagggtacgt
6001 ttgctgacc cgaataagtt tggtttatct gacacctcct tgtttaattc tgaaacgcag
6061 cgccttgtgt gggcatgtgt ggggtgtggag gtcggccgtg gccagcccct aggggttgg
6121 gttagtggcc atccatactt taataaggat gaggatgtgg aaaactcatc tgcctatggc
6181 acaggtccgg ggcaggatag tagggaaaat gtagctatgg attataaaca gacacagtta
6241 tgtatgttgg gctgcacacc cccaattggg gaatattggg gtaagggcac tccgtgcaat
6301 gcttctcggg taacccttgg ggactgtcct gtattagaat taaaaactga ggttattcag
6361 gatggcgaca tgggtggata tggctttggg gctatggatt ttgcttccct gcaggccaat
6421 aaaagtgatg tgccattgga tttatgcaca tctattagta aatatccaga ttatttggga
6481 atggctgcag aaccgtatgg aaatagttta ttttttttc tacgcagggg acaaatgttt
6541 gttaggcact tttttaatag ggcaggtact actggtgata gtgtcccaac tgacttatat
6601 ataacaggta catctggtcg gactcctatt gcaggcagta tttattactc cacaccaagt
6661 ggatccttgg ttacctctga ttctcagata ttaacaagc cattgtggat acaaaaggcc
6721 cagggccata acaatggcat atgttttggc aatcagttat ttgttacagt ttagacacc
6781 actcgtagca ctaatttaac cttatgtgct gccacacagt cccccacacc accccatat
6841 aataacagta atttcaagga atatttgcgt catggggagg agtttgattt gcagtttatt
6901 tttcagttat gtgtaattac cttaaagca gaggttatga catatattca tgcfaatggat
6961 cctacgttgt tggaggattg gaactttaa attgctcctc cagcctctgc atccttagag
7021 gatacatata ggttccttac caacaaggct attgctgtc agcgcgatgc gcccccacag
7081 gtacgggagg atccatataa aaaatataaa ttttgggatg tcaatttaac agaagattt
7141 tcttcccaat tagatcaatt tccattagga cgtaagttcc ttatgcagge tgggtgacgt
7201 gcagggccta ggtttaaac caggaagcgc cctgcccctt cctcgtcttc ctcttctaag
7261 ccagtcaccc ccaaacgtaa aaaaacaaag cgaTAGtat ggctgcatat gtgtattctg
<- L1 end
7321 gggatacac ccttgactc ccgtatgtgt tggactttg gaatgtatgt atgtattgtt
7381 tgtgtatgtt ttgtgtttgt tgtgtatgtg tccgtatggt actgtgtatg tattgtttgt
7441 atgctgtata cctacgttgt ggttgtgtgt ttaataaagc tatatgtgtg gtgtgggtgtg
7501 gtatggcagt gtcatggcgt cttgttcagt gttgaccatg tactttttgt gtctaaatcc
7561 tccattttgt actgcgcgAC CGTTTCGGT cgcggttggc acacacatac atttgtcagc
-> E2 bind
7621 aaatttttat tgcatgttac acattgttag gtcaacatgc octgcaaaa atgtagtatt
7681 agggtgaggt ttggcacacc tttgattgca ctttctctgt actgttactc atcttattgt
7741 atgcacttcc aaacatgctt ttaggcacat agttttgcct aataaaaagt tagctaatag
7801 cagttttggc aaaacataac acttggggtt aacatagtac atattactca tgggtcacac
7861 ctgcggACCG CTTCGGTtg ctacatgttt tattactttc tacttatta
-> E2 bind

```

LOCUS HPV42 7917 bp ds-DNA VRL 21-JAN-1992
 DEFINITION Human papillomavirus ORF E6, ORF E7, ORF E1, ORF E2, ORF E4, ORF E5, ORF L2, and ORF L1 genes, complete cds.
 ACCESSION M73236
 KEYWORDS vacular papilloma.
 SOURCE Human papillomavirus type 42 DNA.
 REFERENCE 1 (bases 1 to 7917)
 AUTHORS Philipp,W., Honore,N., Sapp,M., Cole,S.T. and Streeck,R.E.
 TITLE Human papillomavirus type 42: New sequences, conserved genome organization
 JOURNAL Virology 186, 331-334 (1992)
 COMMENT HPV42 was originally isolated from biopsies of confluent, flat papillomas of the vulva of a 32-year-old patient. It has subsequently been found only in benign genital lesions, usually showing no cell atypia. Beaudenon et al. (Virology 161, 374-384) surveyed 513 benign and malignant genital lesions for the presence of HPV-42 DNA. Positive detection was reported for 17 specimens. All but two of these showed the histologic features of condylomas or flat papillomas, while the remainder were pigmented papules exhibiting some features of mild or moderate dysplasia. Lorincz et al. (Obestet Gynecol 79: 328-33) have classified HPV-42 as a "low-risk" virus.

HPV-42 shares the general genomic organization of all other sequenced papillomaviruses. Its closest relative is HPV-32, while its similarity to other sequenced types is quite low. Philipp et al. report that HPV-42 has not yet been associated with carcinoma, although it displays several features that have been associated with malignant progression. These include the conserved splice donor and acceptor sites potentially leading to an mRNA coding for a spliced E6 protein, and the conserved cell division motif in the E7 protein that has been linked to transformation capability. HPV-42 lacks the glucocorticoid response element found in most "genital" HPVs.

BASE COUNT 2433 a 1478 c 1647 g 2359 t
 ORIGIN

```

1 cttatTATAA ActaCAATcc tggctttgaa aaaTAAGGGA GTAACCGAAT TCGGTtcaAC
  signal -> signal -> -> <- promoter element
                                E2-bind -> <- ->
61 CGAAACCGGT acaTATATAA accacccaaa gtagtgggtcc cagtTAAggc agaATGtcag
  E2-bind <- -> signal                                E6 cds ->
                                E6 orf start ->
121 gtacatctgc ctcatcacag ccacgcacat tataccaatt gtgtaaggaa tttgggctga
181 ctttcgagaa tttacagatt tcctgcattt ggtgcaaaaa gcacttaaca ggcgcagagg
241 tgctcgcgta ccattttaaa gatttggtag tgggtgggag gaaggacttt ccatatgctg
301 catgtgcatt ttgtttagaa tttaattcta aaatttgtgc actgcgacac tacgaaagat
361 cagcattttg gtatacagtg gagaagaaa ctggactact tttagaagaa caacaaatta
421 gatgtgcctt gtgtcaaaag ccgttatcac agagcgaaaa aaaccatcat atTGAtacag
                                E7 orf start ->
481 gtacaagatt tcaatttata ttgtgtcagt ggacgggtcg gtgtacgcat tgcagaggac
541 aATGcgtgga gagacgccta cccTAAagga cattgttttg tttgacatac caacgtgtga
E7 cds ->                                <- E6 end
601 gacaccatt gacctgtatt gctatgaaca attggacagc tcagatgaag atgaccaagc
661 caaacaggac atacagcgtt acagaatact gtgtgtgtgt acacagtggt acaagtctgt
721 TAAactcggt gtgcagtgta cagagggcga cataagaaac ctgcaacaga tgcttttggg
E1 orf start ->
781 cacactggat attgtgtgtc ctttgtgtgc ccgcgtggag TAActgcaAT Gcgggatgat
                                E1 cds ->
                                <- E7 end
841 acaggtacag aggaggggct aggggtttct ggatggtttt gtgtagaagc tatagtagac
901 aaaacaacag aaaatgctat ttcagatgac gaggacgaaa atgtagacga tagtgggtta
961 gatcttggg atttttaga taatagtaca gtaatacata caaagcaggt acatgcacaa
1021 gccttattAA ATAAacaaca agcacatgca gatcaggagg cagtacaggc actaaaacga
  signal ->

```

HPV42

```

1081 aagctattag gcagtccata tgaagccct gtcagtgatt cacagcacag catagacaac
1141 gaactaagtc ctaggcttgg cggtttaacg ctatgtcggg ggtcccaagg ggccaacga
1201 cgattattcc agtcactgga aatcgagac agtggatatg gctattctga agtggaaagta
1261 cagcagacac aggtagaaca cggacatggc gccgtacatg ggactatggg taacgggggg
1321 gcagtgggta gtgaacttgg ggtgcaggaa aatgaagaag gtagtactac aagtacgcct
1381 acaacaaggg tggtagaatt acttaagtgt aagaacctgc atgcaacatt gttaggtaag
1441 tttaaagaat tgtttggagt gtcatttggc gatttagtaa gacagtttaa aagtgacaaa
1501 agcagttgta cagactgggt tattgcagca tttgggggta atcatagtat tgcagaaggg
1561 tttatacat taattaaagc agattcacta tatacacata tacaatggct aacctgtacg
1621 tggggcatgg tgttattaat gctaattaga tttaaatgtg gaaaaaatcg tactacagtg
1681 tccaaaggcc ttagtaaatt attaaacata cctacaaatc aattattaat agagccacct
1741 cggttacaaa gtgtggctgc cgccatatac tggtttagat caggaataatc taatgctagc
1801 attgtaaccg gagacacacc agagtggatt caaagacaaa caattttaga acattgtttt
1861 gcagatgccc aatttaattt aacagaaatg gtgcaatggg catatgataa tgatattact
1921 gaagacagtg acattgcata tgaatatgca caacgggcag acagggatag caatgctgct
1981 gcatttttaa aaagtaactg ccaggcaaaa tatgtaaaag attgtggcgt catgtgcaga
2041 cattataaaa aagcacaatg gagacgtatg tctatgggtg catggataaa acatagaagt
2101 gccaaagatag gggatagtgg agattggaaa cctatagtaa aatttattag atatcaacaa
2161 attgattttt tagcatttat gctcgcattt aaaaagtttt tacataatat acctaaaaaa
2221 agttgtttag tgttaattgg tcctccaaat acaggaaaat cacagtttgg aatgagttTA
                                                    ->
2281 ATAAActtct tagcaggaac tgtaatatca tttgtaaatt cacatagcca tttttggctg
signal <-
2341 cagccattgg acagtgcaaa aatagctatg ctggatgatg caactccacc atgttgagca
2401 tatttagata tatatttaag aaatttatta gatggcaatc catgcagtat agatagaaaa
2461 cataaagcat taacagttgt taagtgccca ccattactta taacatcaaa tacagatatt
2521 agaacaaatg acaaatggaa atacctatac agcagagtta gtttatttga atttccaaat
2581 ccattttccat tagatacaaa tggaaatcct gtatatgaat taaatgacaa aaattggaaa
2641 tcattttttc aaaggtttggt gtccagctTA Gaatttcaag aatcagagga cgaggaagac
E2 orf start ->
2701 tATGgagaga ctggccaac gttttagatgc gtgccaggaa cagttgttag aactgtaTGA
E2 cds ->                                                    <- E1 end
2761 ggaaaatagt agggatttac aaaaacatat tgaacattgg aaatgtttac gtagggaggg
2821 agtggatttg tataaggccc gtgaaatggg ctttgcaaat ataggacatc aaatagtacc
2881 aacattggaa acatgtagag ccaaggccca catggcaatt gaaatacact tggcattaga
2941 gacattattg cagtcctcgt atggtaaaga accatggaca ttgcaagaaa caagtaatga
3001 actgtggcct acgaatccta aaaaatgttt taaaaaacia ggacgtaccg tggaggttat
3061 attgattgga aaacaggaca atgcaatgca ttatacagca tggacatata tatatataca
3121 aactgtgcaa ggtacatggt gtaaagtaca aggacacggt tgccatgcag gactatatta
3181 tattgtggaa aatatgaac agttttattg taattttaa gaggaggcaa aaaaatatgg
3241 ggtaacagac caatgggagg tacatgatgg caatcaggTG Attgtttctc ctgcacccat
E4 orf start ->
NH2 terminus unknown
3301 atctagcacc acatccaccg acgcagagat accctctact ggatctacta agttggtaca
3361 acaagtgtgc accacaaacc cattgcacac cacaacgtcc attgacaacc accacgcaga
3421 ctgtacagac ggaacagcat acaacgtgcc catccaaacc tcaccgccac gaaaacgata
3481 cagacagtggt ggacagtgc catcacagca cctgcagcac tcaaaccoca gcatccccag
3541 catccccagc gcatccgtgg accctggatt gtgtggggtc agaactaaca gtgaaaactg
3601 taacaagcga cggaaacct gtggaagtca ggctacgcct gTAAAttcatt tacaaggtga
<- E4 end
3661 ccctaattgc ctaaaatgcc tacgatttag gctaaaaaga aattgttcac atttatttac
3721 acaggtgtca tctacatggc atttaacaga aaatgattgt acacgtgaca ctaaaactgg
3781 tataataaca atacattatt atgatgaagc acaaagaaat ttatttttaa atactgtaaa
3841 aataccttct gggataaaaat cctgtattgg atatatgtct atgttacagt ttataTGAtt
<- E2 end
3901 agttgtatat gtgtATAAc agttatagga cttcaatact gtgactccac aacgtgtggg
E5 orf start ->
NH2 terminus unknown
3961 acaaccggcc agaactgct gcttttattg tttatagttg ttggtgcgtg tgttgtgtgt
4021 gtgtggatta gtttacaaaa ttatccatat cctgtatggg cctcctgcct tgctagctac
4081 ctaacattgg tgcattatc atgggttcag gtactaacat actttgacta ttttttcta
4141 tgtttaatca ttcttggat tcctctgtc ttactaacat tactaataca ttagcaata
4201 caaTAAcaca tattagttta ggtgtgtgtg tgtgtgtgtc atgtgatttg tacatggttg
<- E5 end

```



```

4261 tacatatata ataccaatta ttgtttggct actattttca tttatagcca cactgctggt
4321 ttgcatattg gtattacaaa cataTAAact gttaccatac gtatatacag tgctgtaAAT
                L2 orf start ->
4381 AAActtttgt tatattgtgt gtacttcttt tgtgctatta caATGccacc acaacggctcc
signal <-
                L2 cds ->
poly-A signal
4441 cgcagacgaa agcgggcctc tgccacacaa ttatatcaaa cgtgtaaggc ctcagggaca
4501 tgtcctccag atgttattcc caaagttgaa ggaaccacat tggcagataa aattttacaa
4561 tggggtagtt taggcgtggt ttttgggggg ttgggaattg gcaactggtg aggtacgggt
4621 gggcgcacgg gctatgtgcc tctgggaaca aggcctcctg taattgctga accaggacct
4681 gcagtacgc caccaatagc tgttgacacc gtggggccat ctgatcctc tattgtttcc
4741 ttattagaag agtcatcagt tattgatgca ggaataacag tacctgatat tacttctcat
4801 ggaggtttta atattactac atctactggt gggcctgcct caacgcctgc tatattagat
4861 atctcccctc ccactaatac tatacgtgtc acaacaacta catctaccaa tcctttatat
4921 attgatcctt ttacattgca gccgccattg ccagcagagg ttaatggggg cctattaata
4981 tctactccta ccatcacacc ccaactcatat gaagaaatc caatggcac gtttgttgta
5041 tctacagata caactaacac atttactagt actcccattc ctggccctcg tgcgtctgca
5101 cgctgggggt tatattctag agcaacgcaa caacgtccag ttactaccag tgcattttta
5161 acatctcctg cacggttggt tacttatgac aatccagcct atgaaggact tacggaggat
5221 acattagtat ttgaacatcc atccattcat actgcacctg accctgattt catggatata
5281 gttgcattgc atcgtcctat gttatcatcc aaacagggta gtgtacgtgt tagtagaatt
5341 ggacaaaggc tgtctatgca gacacgtcgc gggaccctgt ttgggtcacg tgtacacctt
5401 tttcatgacc ttagccctat tacacactct tcagaaacta ttgaattaca gcctttatct
5461 gcttcttcag tatctgcagc ctccaatatt aatgatgggt tatttgata ttatgttgat
5521 actagtgatg taaatgttac aaataccact tcctctatac ctatgcatgg ttttgctacc
5581 ccccgtttgt ccactacatc tttccctaca ttacctagca tgtctacaca ttctgccaat
5641 accaccatac ctttttcggt tcctgccact gtgcatgtgg gccctgattt atctgttgg
5701 gaccaccat gggacagtac cccaacgtct gtaatgcctc agggtaactt tgTAATggta
                L1 orf start ->
5761 tcaggatggg attttatatt gcctcctagt tatttttggc gtaggcgccg taaacctgta
5821 ccataatttt ttgcagATGt ccgtgtggcg gccTAGtgac aacaaggttt atctacctcc
                L1 cds ->
                <- L2 end
5881 tcctcctggt tccaagggtg tcagcactga tgaatatgtg caacgcacca actactttta
5941 ccataccagc agttctagcc tattggttgt tggcaccct tattactcta ttacaaaag
6001 gccaaataag acatctatcc ccaaagtgtc tggtttacag tacagagtat ttagagttag
6061 gctccctgat cctaataagt ttacattgcc tgaactaat ttatataacc cagagacaca
6121 gcgcatgggt tgggcctgtg tggggctaga agtaggtcgt ggacagcctt tgggcgttgg
6181 tattagtggc catccattat tgaataagtt ggatgatact gaaaatgcgc ctacatattg
6241 tggaggccct ggtacagaca atagggaaaa tgtttctatg gattataaac aaacacagtt
6301 gtgttttagt ggctgtaaac ctgccatagg ggagcactgg ggtaaaggta ctgcctgtac
6361 accacagtcc aatggtgact gccaccatt agaattaaat aatagtttta ttcaggatgg
6421 ggatatgggt gatgtagggt ttggggcact agattttggg gctttacaat cctccaagc
6481 tgaggtaact ttggatattg taaattcaat tactaaatat cctgattact taaaaatgct
6541 tctgaggcc tatggtgaca gtatgttttt ctttttaagg cgagaacaaa tgtttgtctg
6601 tcatttgttt aatagggtg gcgcaattgg tgaacctgta cctgatgaac tgtataccaa
6661 ggctgctaata atgcatctg gcagacataa tttaggtagt agtatttatt atcctacccc
6721 tagtggttct atggtaacat ctgatgcaca actatttAAT AAACCATATT GGTTacaca
                E2 bind ->
                signal ->
6781 agcacaagga cacaataatg gtatatgttg gggaaatcag ctatttttaa ctgtggttga
6841 tactaccctg agtactaaca tgactttgtg tgccactgca acatctggtg atacatatac
6901 agctgctaat ttttaaggaat atttaagaca tgctgaagaa tatgatgtgc aatttatatt
6961 tcaattgtgt aaaataacat taactgttga agttatgtca tatatacaca atatgaatcc
7021 taacatatta gaggagtgga atgttgggtg tgcaccacca ccttcaggaa ctttagaaga
7081 tagttatagg tatgtacaat cagaagctat tcgctgtcag gctaaggtaa caacgccaga
7141 aaaaaaggat ccttattcag acttttgggt ttggggagta aatttatctg aaaagttttc
7201 tactgattta gatcaatttc ctttaggtag aaagttttta ctgcaggccg ggttgcgtgc
7261 aaggcctaaa ctgtctgtag gtaaacgaaa ggcgtctaca gctaaatctg tttcttcagc
7321 taaacgtaag aaacacaca aaTAGatgta tgtagtaatg ttatgatata tattttatgt
                <- L1 end
7381 atttatttgt gtactgtggt AATAAAactac tttttatatg ttgtgtgttc tccattttgt
                signal ->
7441 tttttgtact ccattttggt tctagACCGA TTTTCGGTgt atctggcctg ttaccagggtg
                E2-bind ->

```

HPV42

```
7501 caTTGGCCat gtttcctaac attttgcaaa cctattcact ttttaaattt ataaatgcaa
    -> NF-1 bind
7561 tatgtgctgc caactgtttt atggcacgta tgttctgcc aactacactc cctaattcct
7621 ttacataaca cacacgcctt tgcacaggca tgtgcacaaa ggTTGGCAaa ggtagcata
    NF-1 bind ->
7681 tctctgcagt taccatttc ctttttcctt tttttatgt atgagtaact taattggtat
7741 atgtAATAAA aaagctttta ggcacatatt ttcagtTGTG GCAtacacat ttacaagtta
    signal -> NF-1 bind ->
7801 ccTTGGCTta aacaagtaaa gttatttgtc actgttgaca cattactcat atatataatt
    -> NF-1 bind
7861 tgtttttaac atgcaggtgg caACCGAAAC CGGTacataa atccttctta ttctttt
    E2-bind ->
```

LOCUS HPV43E6 591 bp ds-DNA VRL 15-SEP-1989
 DEFINITION Human papillomavirus type 43 (HPV-43), E6 region.
 ACCESSION M27022
 SOURCE Human papillomavirus type 43 DNA recovered from a vulvar biopsy with hyperplasia.
 REFERENCE 1 (bases 1 to 591)
 AUTHORS Lorincz,A.T., Quinn,A.P., Goldsborough,M.D., Schmidt,B.J. and Temple,G.F.
 TITLE Cloning and partial DNA sequencing of two new human papillomavirus types associated with condylomas and low-grade cervical neoplasia
 JOURNAL J. Virol. 63, 2829-2834 (1989)
 COMMENT HPV-43 was classified by Lorincz et al. (Obestet Gynecol 79: 328-337) as a "low-risk" virus. Prevalence studies indicate that HPV-44 and HPV-43 have been found in 4% of cervical intraepithelial neoplasms, but in none of the 56 cervical cancers tested.

During an analysis of approximately 1000 anogenital tissue samples, two new HPV types, HPV-43 and HPV-44, were identified. The complete genome of HPV-43 was recovered from a vulvar biopsy and cloned into bacteriophage lambda. The biopsy was taken from a woman living in the Detroit Michigan area. The DNA consisted of two fragments: a 6.3 kb BamHI fragment and a 2.9 kb HindIII fragment. The total quantity of unique DNA was 7.6 kb. Only the E6 region of the cloned sample has been sequenced, although all positions of the ORFs have been deduced and are consistent with the organization of DNA from HPV-6b. A possible feature of HPV types associated with malignant lesions is the potential to produce a different E6 protein by alternative splicing. This potential has been found in types HPV-16, HPV-18, and HPV-31. HPV-43 has both the potential E6 splice donor site at nt 233 and the potential splice acceptor at nt 413.

BASE COUNT 189 a 110 c 133 g 159 t
 ORIGIN 106 bp upstream from beginning of E6 cds
 1 attactaaca attattatac ttgtagttTA AgggtgggAC CGAAAACGGT ccgACCGAAA
 E6 orf start -> -> E2 bind -> E2 bind
 61 GCGGTacata tataaaccac ccaaaaacca tagcttgtgg ggcataATGt ctgcacgtag
 E6 cds ->
 121 ctgctcccaa aacgcacgga ctatatattga gttgtgtgat gaggtaaca taactttgcc
 181 tactctgcaa attgggtgca tattttgcaa gaagtgggta cttaccacgg aaGTattatc
 5' sj /\
 241 gtttgcatth agagatttaa gggttgtgtg ggcgcacgga tatccgtttg ctgcatgctt
 301 ggctgtccta cagtttcatg gaaaaataag tcaatatagg cactttgact acgcagcata
 361 tgcagatact gtagaagaag aaacaaagca aacagtgttt gattttgca ttAGatgctg
 3' sj /\
 421 taagtgcac aagccattat caccagtgga aaaagtacag catattgtgc aaaaggcaca
 481 attctttaa atacatagcg tgtggaaagg atactgctta cattgctgga aatcatgcat
 541 gaaaaaacgc cgacgatcag agactatgtg cTAAActatgc aaccagaacc t
 <- E6 end

//

HPV43MY911

LOCUS HPV43MY911 455 bp ds-DNA VRL 16-OCT-1994
DEFINITION Human papillomavirus type 43 (HPV-43), partial L1 cds, My09/My11 region.
ACCESSION U12504
SOURCE Human papillomavirus type 43 DNA recovered from a patient with vulvar hyperplasia.
REFERENCE 1 (bases 1 to 455)
AUTHORS Bernard,H.-U., Chan,S.-Y., Manos,M.M., Ong,C.-K., Villa,L.L., Delius,H., Peyton,C.L., Bauer,H.M., and Wheeler,C.M.
TITLE Identification and assessment of known and novel human papillomaviruses by PCR amplification, restriction fragment length polymorphisms, nucleotide sequence, and phylogenetic algorithms
JOURNAL J. Infect. Dis. (1994) In press
COMMENT HPV-43 was first isolated from a patient with vulvar hyperplasia. The cloned DNA was provided by Dr. A. Lorincz and was subsequently sequenced by Dr. H. Delius over the L1 MY09/MY11 segment. HPV-43 and the several other types recently sequenced over this region by Dr. Delius were used as type-specific probes to screen DNA for novel genital HPV types. The screened DNA was obtained from four recent epidemiological studies. Primer regions are annotated in the sequence; information in this region is not accurate due to primer degeneracy.

BASE COUNT 134 a 93 c 89 g 139 t
ORIGIN
1 gccacaggac ataataatgg catttgtttt gggaatcagt tgtttgttac agtggtagat
L1 cds ->
-> MY11 PCR primer <-
61 accactcgta gtacaaactt gacggtatgt gcctctactg accctactgt gcccagtaca
121 tatgacaatg caaagtttaa ggaatacttg cggcatgtgg aagaatatga tctgcagttt
181 atatttcaat tatgcataat aacgctaaac ccagagggtta tgacatatat tcatactatg
241 gatccacat tattagagga ctggaatttt ggtgtgtccc cacctgcctc tgcttctttg
301 gaagatactt atcgcttttt gtctaacaag gccattgcat gtcaaaaaaa tgctcccca
361 aaggaacggg aggatcccta taaaaagtat acattttggg atataaatct tacagaaaaa
421 tttctgcac aacttaccba gtttccctta ggcgc
L1 cds ->
-> MY09 PCR primer <-

LOCUS HPV54MY911 452 bp ds-DNA VRL 16-OCT-1994
 DEFINITION Human papillomavirus type 54 (HPV-54), partial L1 cds, My09/My11 region.
 ACCESSION U12501
 SOURCE Human papillomavirus type 54 DNA.
 REFERENCE 1 (bases 1 to 452)
 AUTHORS Bernard,H.-U., Chan,S.-Y., Manos,M.M., Ong,C.-K., Villa,L.L., Delius,H., Peyton,C.L., Bauer,H.M., and Wheeler,C.M.
 TITLE Identification and assessment of known and novel human papillomaviruses by PCR amplification, restriction fragment length polymorphisms, nucleotide sequence, and phylogenetic algorithms
 JOURNAL J. Infect. Dis. (1994) In press
 COMMENT HPV-54 was first isolated from a patient with condyloma acuminata. Cloned HPV-54 DNA was obtained from the Papillomavirus Reference Center, Heidelberg and subsequently sequenced by Dr. H. Delius over the L1 MY09/MY11 segment. HPV-54 and the several other HPV types recently sequenced over this region by Dr. Delius were used as type-specific probes to screen DNA for novel genital HPV types. The screened DNA was obtained from four recent epidemiological studies. Primer regions are annotated in the sequence; information in this region is not accurate due to primer degeneracy.
 BASE COUNT 134 a 87 c 94 g 137 t
 ORIGIN
 1 gcacagggtc ataataatgg tatttggttg ggcaatcaat tgtttttaac agttgtagat
 L1 cds ->
 -> MY11 PCR primer <-
 61 accacccgta gtactaacct aacattgtgt gctacagcat ccacgcagga tagctttaat
 121 aattctgact ttagggagta tattagacat gtggaggaat atgatttaca gtttacattt
 181 cagttatgta ccatagccct tacagcagat gttatggcct atattcatgg aatgaatccc
 241 actattctag aggactggaa ctttgggtata accccccag ctacaagtag tttggaggac
 301 acatataggt ttgtacagtc acaggccatt gcatgtcaaa agaataatgc cctgcaaag
 361 gaaaaggagg atccttacag taaatttact tttggactg ttgaccttaa ggaacgattt
 421 tcacttgacc ttgatcagta tccccttggg cg
 L1 cds ->
 -> MY09 PCR primer <-

HPV57

LOCUS HPV57 7861 bp DNA VRL 13-APR-1992
 DEFINITION Human papillomavirus type 57 (HPV-57), complete DNA.
 ACCESSION X55965
 SOURCE Human papillomavirus type 57 DNA.
 REFERENCE 1 (bases 1 to 7861)
 AUTHORS Delius,H.
 JOURNAL Unpublished (1990)
 REFERENCE 2 (sites)
 AUTHORS Hirsch-Behnam,A., Delius,H. and De Villiers,E.M.
 TITLE A comparative sequence analysis of two human papillomavirus (HPV) types 2a and 57
 JOURNAL Virus Res. 18, 81-98 (1990)
 COMMENT HPV-57 was originally isolated from the biopsy of an inverted papilloma of the maxillary sinus of an 81-year-old patient. This later developed into an invasively growing tumor which caused the death of the patient. Screening of a large number of biopsies from benign and malignant lesions from different tissues subsequently revealed it to be present in both genital and oral lesions, as well as in two cases of verrucae vulgares from immunosuppressed patients. In none of these samples was there any evidence for integration of the viral genomes. HPV-57 has infrequently been detected in cases of cervical intraepithelial neoplasia. It thus seems to have a preferential, but not exclusive, tropism toward mucosal tissues (de Villiers, et al. Virol. 171, 248-253).

HPV-57 is most closely related to HPV-2a and HPV-27. These three viruses have the highest G/C content of all sequenced papillomaviruses, while HPV-57 has the highest G/C content among this group.

The ORFs of the HPV-57 genome are generally homologous to those of all other sequenced papillomaviruses. The authors of [2] report that none of its ORFs show any definite homology to the E5 ORFs of other papillomaviruses, but propose several possible candidates in the region following E2 and preceding L2. In addition, they note the presence of the following potential regulatory elements: polyadenylation signals following both the early and late sets of genes; a number of direct repeats both in the LCR and in the non-coding region between the E2 and L2 genes; E2-binding sites located in the LCR; an Sp-1 binding site is located in the region following the E2 gene, and NF-1 binding sites are located in the LCR on both strands of the genome.

BASE COUNT 1973 a 1867 c 2069 g 1952 t

ORIGIN

```

1 taatatataa ctataatcct tcattcaaaa aaTAGggcgt aACCGAAAAC GGTcagACCG
                                     E6 orf start -> -> E2-bind -> E2-
61 AAAACGGTcg TATAAAaaca ggagcgccat gtacaagggc agggATGtct gaagaaaatc
   bind -> signal E6 cds ->
121 catgccctag gaacatcttt ctgctgtgca gagagtatgg tttggagcta gaggatttga
181 gaatactgtg cgtgtattgc aagcggccgt tatcagacgc tgatgtgctg gcatttgcag
241 taaaggaact gttttagtagt tggagaaagg gattccctta tggagcatgt gaaaaatgct
301 taattgcagc agcaaaactt agacaataca ggcaactggca ttactcatgc tacggagaca
361 cagtggagac cgagacagga ataccatAC CTCAGCTGTT tatgagatgc tataatctgccc
                                     -> E2-bind

```

```

421 aTAAgcccct gtggtgggag gagaaggagg cactactggt cggaacaag cggttccaca
E7 orf start ->
481 agatatcagg ccagtgacc ggacattgca tgaactgtgc gccaaAGTgc atggagaacg
E7 cds ->
541 ccccgacctt gaggacatca cacTAAatatt ggaagaaata cccgaaattg ttgacctaca
<- E6 end
601 ttgagcagag caatttgaca actcagaaga agataactaac tatcaactga cagaaccagc
661 tgtgcaggcc tacggggtgg taacgacgtg ctgtaagtgc cacagtacag tccggctggt
721 ggttgagtgc ggagcggcgg acataaggca tctggagcag ctgttcctga atacgttgac
781 caTAGtgtgc ccccgtgctg taTAGcgtcA TGgaggattc ggaaggtacc gacgggaccg
E1 orf start ->
E1 cds ->
<- E7 end
841 atgaggacgg gtgccgggca ggggggtggt tccatgtgga agccataata actcatggcc
901 agagtcagggt atccagtgac gaggatgagg atgaaacaga aacaaggagg gatttagact
961 tcatagacaa tagggttccc ggagatgggc aggaagttcc ctgagcagta tatgcacaac
1021 aaatcgccca ggatgacgaa gcaacagtgac agggcctaaa acgaaagtgt ttggcgagtc
1081 ctttgtctgc atgctcatgc atagagaatg atttgagtcc cagattagat gcaatctcgc
1141 taaacagaaa gtcagaaaaa gcgaagagag gcttattcga gacagagcca ccagacagtg
1201 ggtatggcaa tacgcagatg gttgttggaa cgccagaaga ggtaacaggg gaagacaaca
1261 gtcagggggg gcggccgggtg gacgttaggg aggaggagcg tcaagggggg gacggagagg
1321 cagatctaac tgtacacact ccacagtcag gaacagacgc ggcggttagc gtgctgacgt
1381 tactgaaaag tagcaatctg aaggcgacgt tactgagtaa gttcaaggag ctatacgggg
1441 tgggatatta cgaactggtc agacagttca agagcagcag gacagcatgc gcagactggg
1501 tagtatgtgc ctttcgtgtg tattatgcag tggcagaggg tataaaacag ctgatacagc
1561 cacacacgca atatgcacac atacagatac aaaccagttc ctggggaatg gtggttttca
1621 tgttgttcg ctacaactgt gcaaagaaca gggaccaccg ctccaagaac atgagcatgc
1681 tgttaaacat tcccgagaag catatgctca tagaaccacc aaaactaaga agtacacctg
1741 ctgccttgta ctggtacaag acatccatgg gtaatgggag tgaggctat ggagagacac
1801 cagaatggat tgtgagacag acactgatag gacacagtat ggaggatgag cagttcaaat
1861 tatctgttat ggtgcagtac gcataatgacc atgacataac ggacgagagc gcactggcat
1921 ttgagtacgc acagctggcg gacgtggacg ccaatgccc agcatttcta aacagcaatt
1981 gccaggccaa atacctaaaa gacgcgggtg caatgtgcag aactacaag cgtgcagaga
2041 ggaacaaat gagtatgagc cagtgatca cattcagagg gtagtaagata tcagaggaag
2101 gagattggaa gcctatagtg aagtttttaa ggcacaggg ggtagagttc gtgtcattcc
2161 ttgcccctt taagtcatth ctaaaggcg tgcccaagaa gaattgcata gtgtttatg
2221 gacctgcaga tacaggcaaa tcataattht gcagagttc tttgagttc ctagggtggc
2281 ctgtaataatc atatgccaat tccagtagcc acttttggct gcagccctta gctgacagta
2341 agataggggt actggacgat gcgacggccc agtgctggac gtatattgat acatatctta
2401 ggaacctact agatggcaat ccctcagca tagacaggaa acataagacc ctgctgcaga
2461 taaaatgtcc tccactgatg ataacaacca acataaatcc tttagaggag gacaggtgga
2521 agtatttgcg cagcagagta acactgttta agtttaccac cccatttccc ttcgcaagtc
2581 ccggggagcc cttataccct ataaataatg caaactggaa atgctttttc caaaggtcgt
2641 ggtcccgcct agaccTAAac agtccagagg atcaggaaga caATGgaaac actggcgagc
E2 orf start ->
E2 cds ->
2701 cgttttagatg cgtgccagga gacgttgcta gaactgtaTG Aaaaagatag caacaaactt
<- E1 end
2761 gaggacaaaa ttaaaccattg ggcgcaggtc cggctagaaa atgtaatggt gttaaaggct
2821 cgggaatggt gaatgacacg agtcggctgt acaacagtgcc ccgccctcac cgtgtcgaaa
2881 gctaaggctt gtcaggccat agaggttcag ctggcattac agacattgat gcagagtgca
2941 tatagcacgg aggcattggac cctgcgagac acgtgcctgg agatgtggga agcacctcca
3001 aagagatgct ggaaaaagaa aggacaatca gtattagtaa agtttgatgg cagctgtgac
3061 agagacatga tatacacggg ctggggccat atatatgtgc aggacattaa cgatgatacc
3121 tggcataaag tgcccgggca ggtggacgaa ctgggactat tttatgtgca cgacggcgta

```

HPV57

```

3181 cgtgtaaatt atgtggactt tggaaatagag gcccTGACCT ATGGGGTTac tgggacgtgg
                                E4 cds ->
                                E4 orf start ->
                                -> E2-bind <-
3241 gaggtgcagg tcggggggcg tgttatttat catacatccg catctgtgtc tagtaccag
3301 gccgccacct cggacgacga cacacatcc cctcttagat ctgctgcggc cgcagtcaca
3361 gccacagcca cagccacaac agcagtcccg cccacactcc aggactccgc ccaggcgcca
3421 tcgagtccgc cacccaagcg ccagcgggtc atcgtcggac agcagtggca acagcccagc
3481 tctacgcgga aggtcagaga agggcaggtg gagtgtcaaa acgacaggag catccgtaac
3541 cctgacagca cagacccccg cggggggccac agtgaccttg acgctgtgcc tgTGATccac
                                <- E4 end
3601 ctgcaaggtg aagcaaactg tttaaagtgc ttcagataca gggtgcaaaa acataaagac
3661 gtaactgtttg tgaaggcatc ctcccactgg cattgggcgt gtgggaatgg tgacaagact
3721 gcctttgtaa catttgtgta caaaagttag gaacagcgtg cagaattcct tacaagggtc
3781 catctACCCA AAGGGGTgaa ggcactgcca gggatatatgt ctgcatttgt aTAAatgtat
                                -> E2-bind <-
                                <- E2 end
3841 agtctgtaac ataatacaca ggtccgcaac atttCACAGC Ctgtatcacc aaCACAGCCc
                                -> repeat region start
3901 tggactactt tctctgcgtg gtgtcagggg ggtcccatct gctgtactg ttgctcttcc
3961 tttggctctc tcaactaacc cctctggtag cctatctggt gttcttcttc tgtgtCTATA
                                signal ->
4021 TAGggctgtg gttgatattt ttgcaggcct tgtggtttct acaatagttg atattacatt
4081 gtgaccaact gctgctacc tgtatatacc tccccctgt atactgcaat gtatcctgct
4141 gtgtacaagg gcacGGGCGG gtcgtatccc attgtcctgt ggggcoctga tgatgtggat
                                Sp-1 bind ->
4201 tgcTTGTTGA TtattcttTT GTTGATtgcc attttgTTGT TGATgctgta cgtccgctcg
4261 ctgcagtagt gTACATACTC ATTgtgtttt tgcTACATAC TCATTtTGAT acatttttat
                                L2 orf start ->
                                repeat region end <-
4321 cgggtgtgcac ctgattgcct ttgtattgct acgtgtcAAT AAActtattg ccATGtcacc
                                signal ->
                                L2 cds ->
4381 acgtgcaaaag cgtcggaaag cgcctcccc cactgacttg tateggacat gcaagcaggc
4441 tggaaactgac ccccctgata tcatacctag ggtggaacag gacacattag ctgataggat
4501 actcaaatgg ggcagcctgg ggtcttttt cgggtggcctc ggtataggta ctggaagcgg
4561 cactggggga cgcacaggct acataccagt gggcaccaga ccaacaactg tcgttgatgt
4621 aggactggcg ccaaggccac ctgtagtaat agaacctgtg ggggcgtctg aaccatctat
4681 tgtaatttg gtggaggatt ctatgcatcat taatgctggg tcctctcatc caACCTTAC
                                E2-bind ->
                                E2-bind ->
4741 CGGTACTGGT gggtttgagg ttaccacctc aacggtgact gaccctgcag tcctggacat
4801 cactccttcg ggtaatgggg tgcaggttag cagcagtagc tttgtgaatc ctctcttcac
4861 tgaccctgct attattgagg ctccccaggc tggggagggt acagggcatg tgcttgtag
4921 cactgcccaca tcagggtccc acggcttcga ggaataacca atgcagacct ttgcgacctc
4981 tgttggggat ggaggagag ccataagtag cacacctgtc ccaggcgtgc gcagggttgc
5041 tgggccccgc ctttatagta gggctaataca gcaggtgccg gtccgggacc ctgcctttat
5101 tgaccgtcct gcggatttgg tgacatttga caacctgtc tatgaccccg aggaaactat
5161 aatatttcag catccaggct tgcattgacc accggaccga gacttctctg acatagtgctc
5221 actgcaccgc cctgccctca catccACCCG GCAGGGTact gtccgcttca gtcgggttggg
                                -> E2-bind
5281 acgccccggc acgcttcgca cgcgtagtgg taaacaaatt ggggctaggg tacacttcta
5341 tcatgatata agccctgtcg ctcccaggga attggagatg gagccgctgt tccccccac
5401 gtcggagccc ctatatgaca tatatgccga gtcggatttc ttgcaacct tagattcgga
5461 tgtccccgcg gccctcagag gtaccctttc cctggcagac actgcagtgt ctgcatccac
5521 cgcttctacg ttgcgggggg ccaccactgt tcccctgtca ggtgggtgtg atgtgcctgt
5581 gtataccggg cctgatattg acccgtctgt aggcctgggt atgggaccgc tgggtgcctgt
5641 gataccagcc ataccatcct ctgtgtacaT AGttgggggt gattactact tgctgccaag
                                L1 orf start ->
5701 ttATGttctg tggcoctaaac gacgtaaacg tgtgcaactat ttctttgcag atggctatgt
L1 cds ->
5761 ggcggccTAA tgaagcaag gtatacctgc ctccaacacc tgtctcaaag gtgctcagta
                                <- L2 end
5821 cggatgtcta tgtcacgcgg acgaaatgtt attatcatgg tgggagctct cggctcctca
5881 cagtaggcca tccatattat tctataaaaa aaagtggcaa taataagggtg tctgtgcccc
5941 aggtatcggg ctaccagtac cgtgtgttcc atgtgaagct gccggaccct aataagtttg

```



```

6001 gtctgcctga tgccaacctc tatgatcccg acaccagcgc tctgctgtgg gctgtgtctg
6061 gcgttgagggt gggctcgtggc cagcctctgg gtgtagggat atcoggccac ccttattata
6121 acaaacagga tgatactgaa aattcacaca atcccagcgc agctgatgat gggagggagt
6181 atatatccat ggattataaa cagacacagc tgtttatfff gggttgcaag ccccctatag
6241 gtgagcattg gtccaagggc actacctgca gcgggtcttc tgctgttggg gactgtcccc
6301 ccctgcagtt tacaaatacc actattgaag atggggatat ggttgaacc gggttcgggg
6361 cgctggattt tgccgctcta cagtccaaca aatcagatgt ccccttggat atctgtacta
6421 acatatgtaa atatccagac tatctgaaga tggctgcaga cccttatggc gattctatgt
6481 tcttttcctc gcgcagggag caaatgttca ctccggcattt ttcaatcgg ggtgggtcga
6541 tgggtgacgc cctcccggat gagctatatg tcaagagttc taccgtccag acccccggta
6601 gttatgttta tacctccact cccagtggct ctatggatc ctctgaacag cagttatffa
6661 acaagcctta ctggctgagg agggcccagg gacataacaa tggcatgtgc tggggcaatc
6721 ggatcttctc aacagtgggtg gacaccacgc gcagcacaaa tgtctctttg tgtgccactg
6781 taaccacaga aactaattat aaagcctcca attataagga ataccttagg catatggagg
6841 aatatgattt gcagttcatt tttcaactgt gcaaaataac actcaccccc gagataatgg
6901 catacataca taacatggat gcgcgggtgc tagaggactg gaactttggg gtccccccac
6961 ccccgctccg cagcctgcag gacacctaca ggtatttgc atcccaagc ataacatgtc
7021 agaagcccac acccctaag acccctactg atccctatgc aacctagaca ttctgggatg
7081 tggatctcag tgaaagtttt tccatggatc tggaccaatt ccccctggga cgcaagtttt
7141 tattgcagcg gggggccacc cccactgtgt ctgaaaaacg cgcogctgca actgcagcgg
7201 cgcccactgc taaacgcaaa aaggtcaggc gaTAGtgatt ctgtgtctgc ctcaattcct
        <- L1 end
7261 ttgctctact tttgtatatg tacatatggt tcagtgttgt ctgtgtgttt gtgtttgtgt
7321 gtgttgtctg ttattatggt tgcatGTACA CATGTCgGTA CACATGTctg gtatgtattc
        -> repeat region start
7381 ctcccatatg AATAAAcgtg tgtcatgtgt tgtgtgtcct gttgcaactc gtaattgtcc
        signal ->
7441 ccgctgcatg gtttgacac tgtggcctgt atgtagcccc ctggtagata cgACCGTTTT
        -> E2-bind
7501 CGGTtgctg cagTTTCGGT cggcgctgct gccagcacac tcatatcctt taatccttta
        repeat region end <-
7561 attgctttaa tcctttcact tttttactgt gccaaactaaa atgattttgc tttttgattg
7621 ttttgtgtct gcattaatgc agtttttctt tttccagtgc cagaccgcgt gtgggcgtgc
7681 acattcctac atagattatc ttctgtgtgt ggcggggatt ttccctgcgt ctgcagaaaa
7741 acctggccac acagcacctt gggcgcgtcg tttttgcagc caactttccc ttgccaagtt
7801 gtcttgcgcg gcattccaag aaacacACCT ATTCCGGTcg caatctctac tatgtggttt
        -> E2-bind
7861 a

```

HPVCP8061

LOCUS HPVCP8061 452 bp ds-DNA VRL 16-OCT-1994
DEFINITION Human papillomavirus, isolate CP8061, partial L1 cds, My09/My11 region.
ACCESSION U12479
SOURCE Human papillomavirus, isolate CP8061.
REFERENCE 1 (bases 1 to 452)
AUTHORS Peyton,C.L. and Wheeler,C.M.
TITLE Identification of five novel human papillomaviruses in the New Mexico triethnic population
JOURNAL J. Infect. Dis. (1994) In press
COMMENT Data kindly provided prior to publication by Dr. C. Wheeler, University of New Mexico, School of Medicine, New Mexico Tumor Registry, 900 Camino del Salad NE, Albuquerque, NM, 87131-5306.

Five novel HPV sequences were identified in a study in which 3655 cervical specimens were screened against known genital HPV DNA [1]. The specimens were obtained from clinical investigations conducted at the University of New Mexico. The study subjects included Native Indians, Hispanics, and non-Hispanic whites. The viral DNA was PCR amplified using the L1 consensus primer MY09/MY11 pair, which can hybridize to a broad spectrum of HPV types. Resultant fragments range from 449 to 458 nucleotides in length. The amplification products were initially screened against 2 sets of type-specific probes and a generic probe. If hybridization to the generic probe and not to the type-specific probes occurred, the samples were further analyzed by restriction fragment length polymorphisms. RFLP patterns which did not match reference patterns were considered to be derived from novel HPVs. The five novel samples which were identified in this study include CP8304, CP6108, CP8061, CP141, CP4173. Peyton et al. also identified two HPV45 subtypes and one HPV56 subtype. They conclude that since the existence of subtypes appears to be relatively rare, it suggests that HPV45 and HPV56 are more divergent than many HPV types. It should be noted that CP141 (U12476) is almost identical to LVX160 (U12486) and HPV1LAE1 (U01535) and that CP4173 (U12477) is almost identical to LVX100 (U12485). Both LVX160 and LVX100 were identified by Ong et al. in a 1994 study which examined Amazonian Indian subjects (Ong et al., J. Infect. Dis., 1994, in press). Primer regions are annotated in the sequence; information in this region is not accurate due to primer degeneracy.

In a subsequent study Bernard et al. evaluated ten novel genital HPV types, including the five identified in the Peyton et al. study, and other known genital types to determine phylogenetic relationships. They observed that the genital types CP6108, CP8304, CP4173 and CP8061 form a branch with HPV types 61 and 62. This emergent minor branch is positioned between two others which contain cutaneous types. Bernard et al. speculate as to whether other low-risk genital types have escaped detection because of considerable sequence divergence from the common genital types (Bernard et al., J. Infect. Dis., 1994, in press).

Bernard et al. also assessed the linear correlation coefficients for the MY9/MY11 fragments against the rest of L1 (.851) and against the E6 gene (.888). Since these values are close, the authors suggest that the evolutionary distance information obtained for the primer pair region should be comparable to that available from the other regions of the genome (Bernard et al., J. Infect. Dis., 1994, in press).

BASE COUNT 127 a 90 c 93 g 142 t
ORIGIN

1 gcacaggtgc ataacaatgg catttgttgg ggcaatcagc tttttgtaac agttgtggac
L1 cds ->

```
-> MY11 PCR primer <-  
61 acatcacgta gtacaaatat gtccatctgt gctacaaaa ctggtgagtc tacatataaa  
121 gcctctagtt tcatggaata tttgagacat ggagaagaat ttgatttgca atttatattt  
181 caactatgtg ttattaattt aacagctgaa attatggcct acttacatcg catggatgct  
241 acattactgg aggactggaa tttttggttc ttaccacctc ctaactgctag tcttggatgat  
301 acctaccgct ttttacagtc tcaggccata acctgtcaga aaaacagtcc tctctctgca  
361 gaaaaaaagg acccctatgc agatcttaca ttttgggagg tggatttaaa ggagcggttt  
421 tcaactagaat tggatcagtt tcccctggga cg  
L1 cds ->  
-> MY09 PCR primer <-
```