Group B1 Sequences
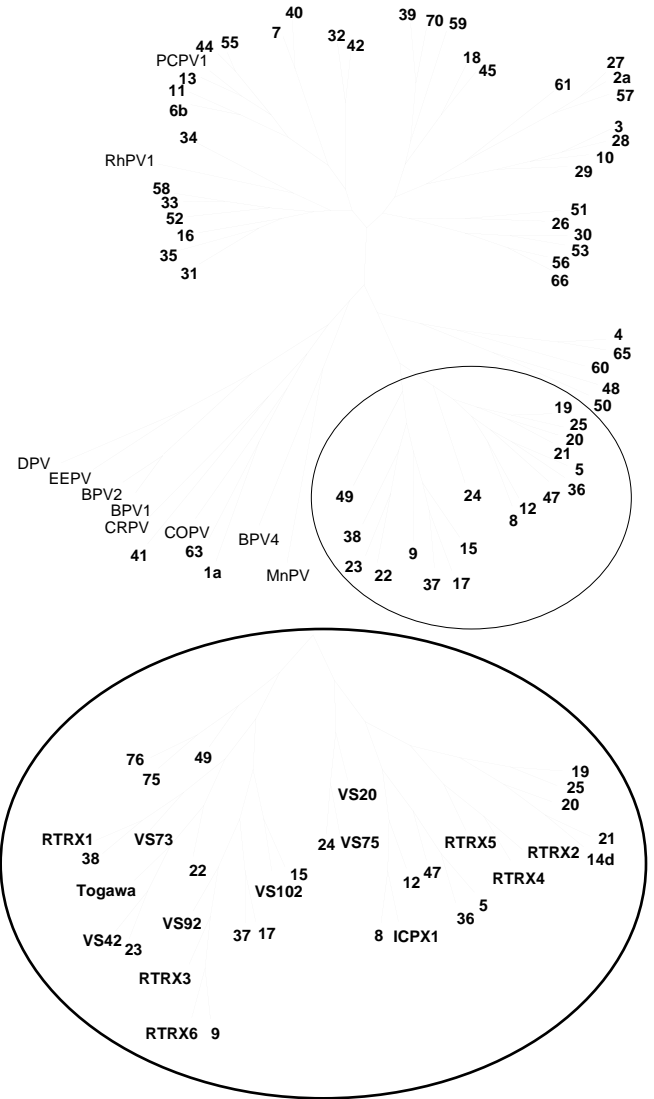
HPV5	HPV8
HPV9	HPV12
HPV14d	HPV15
HPV17	HPV19
HPV20	HPV21
HPV22	HPV23
HPV24	HPV25
HPV36	HPV37
HPV38	HPV47
HPV49	HPV75 (VS40)
HPV76 (CR148)	HPVICPX1
HPVRTRX1	HPVRTRX2
HPVRTRX3	HPVRTRX4
HPVRTRX5	HPVRTRX6
HPVVS20	HPVVS42
HPVVS73	HPVVS75
HPVVS92	HPVVS102
HPVTogawa	

INTRODUCTION

Group B1 viruses composed the old group H and are primarily associated with the multifactorial disease Epidermodysplasia Verruciformis (EV). Recent work suggests they are also commonly associated with immunosuppressed renal transplant recipients [1,2]. Several isolates which appear to constitute new types have been found in skin lesions of renal transplant patients [5,6]. Association of EV-related HPV types with squamous cell carcinomas (SCC) of the skin, and with SCCs of the esophagus has recently been suggested [3,4,5,7,27]. One potential new type was isolated from an immunocompetent patient [5].

Patients with EV tend to have depressed cell-mediated immunity [8]. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas [9]. Benign wart scrapings tend to be multiply-infected, with as many as six different viral types. However, in contrast, EV carcinomas tend to harbor only a few types, specifically HPV-5 and HPV-8, and less frequently HPV-14, HPV-17, HPV-20 and HPV-47 [9,10]. These types are rarely detected in lesions afflicting the general population. A key to host restriction of these viruses may be in part due to the unusual organization of the LCR in these viruses. The LCR of these viruses is short compared to the viruses in other groups; most of the types in this group also contain two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding [10,11,12]. Study of variants of HPV5 and HPV8 has revealed a higher level of sequence diversity within these types [13,14,15] than has been observed in primarily mucosal types such as HPVs 6, 11, 16 and 18 [16,17,18], suggesting that the EV-related types may accumulate mutations at a higher rate. One possible explanation for this is that the cutaneous tropism of the EV types could lead to additional mutations through UV-induced DNA damage; however, sequencing of variants of primarily cutaneous HPV2 show levels of variability comparable with that seen in the mucosal types [19].

This group forms two major branches based on phylogenetic analysis, each of which can be subdivided into two minor branches. These clusters have been designated as a_1, a_2, b_1, and b_2.

This phylogenetic classification is compatible with other classifications based on hybridization [21], transforming activity of the E6 gene [20], and conservation of the M33 and M29 regions and E2 binding sites in the LCR [11]. In addition, HPV-24, HPV-49, and several of the new sequences seem to form isolated branches, which may be related with their detection in immunosuppressed, non-EV patients.

Cluster a_1 consists of HPV-5, HPV-8, HPV-12, HPV-36 and HPV-47; the available sequence of HPVICPX1 suggests that it too is a member of this cluster. Both HPV-5 and HPV-8 are associated with macular lesions which frequently progress to malignancy [22,23,24]. Yabe et al. studied the characteristics of HPV-5 in lesions of differing severity. In a primary carcinoma, HPV-5 was present in an episomal state with a 40% subgenomic segment amplified. In the metastatic tumor, only the 40% subgenomic region was present, but integrated into the host genome [24]. The segment was determined to be the entire sequences of E6, E7, and the noncoding region and portions of E1 and L1, with no mutations present [25]. In addition, amplifications of the LCR have been reported in HPV-5 associated carcinomas [26]. HPV-5 and HPV-8 have also been found in significant numbers in squamous cell carcinomas of renal allograft patients. Barr et al. detected either HPV-5 or HPV-8 in nearly 60% of the cases surveyed in Scotland [27]. HPV-47 is primarily associated with benign lesions; however, it has also been detected in cases of malignancy [20]. HPV-12 induces benign macular and flat wart-like lesions [28]. HPV36 was isolated from two patients with actinic keratosis. HPVICPX1 was isolated from an immunocompetent patient [5]; other information is not currently available.

Cluster a_2 consists of HPV-14, HPV-19, HPV-20, HPV-21 and HPV-25. HPV types forming this cluster produce benign macular or flat wart-like lesions and malignant lesions in isolated cases. Both HPV-19 and HPV-25 induce macular lesions, which are benign in character [21,20,29]. HPV-14, HPV-20 and HPV-21 induce flat-wartlike lesions; HPV-20 and HPV-14 have been detected in carcinomas [20,29].

Cluster b_1 includes of HPV-9, HPV-15, HPV-17, and HPV-37; available sequence indicates that
HPVRTRX3, HPVVS92 and HPVVS102 are also members of this cluster [5,6]. HPV-15 was isolated from a benign flat wart-like lesion [29]. HPV-17 was isolated from benign macules and subsequently from squamous cell carcinomas and the malignant melanoma of an immunosuppressed patient [29,30]. HPV 9 DNA induces both macular and flat wart-like lesions, however it has also been identified in a keratoacanthoma [28,31]; HPV-37 was found in the same keratoacanthoma. HPVRTRX3, HPVVS92, and HPVVS102 were isolated from a squamous cell carcinoma, skin wart, and dysplastic wart respectively in renal transplant patients [5,6]. Also possibly belonging to this cluster is HPVRTRX6, although its position in phylogenetic trees is rather unstable. HPVRTRX6 was isolated from an SCC in one renal transplant patient [5].

Cluster b_2 includes HPV-22, HPV-23, and HPV-38; available sequence indicates that HPVRTRX1, the Togawa isolate, HPVVS42 and HPVVS73 are also members of this cluster. HPV-22 and HPV-23 were isolated from macules of EV patients [29]. HPV-38 was isolated, along with HPV17a, from a superficial spreading melanoma in an immunocompromised patient [31]. The Togawa isolate was found in multiple SCCs of the esophagus in nonimmunocompromised patients [7]. HPVRTRX1, HPVVS42 and HPVVS73 were isolated from an SCC, a verrucous biopsy and a skin wart biopsy in renal transplant patients [5,6].

Isolated types Several EV-related types or potential new types seem to be relatively unrelated to the clusters defined above, and, for the most part, to each other. HPV-49 was isolated from the flat warts of a Polish renal transplant patient. Favre et al. screened benign and malignant lesions from the general population, EV patients and transplant patients for the presence of HPV-49. In the survey, HPV-49 was not detected in any of the patients with EV but was detected in two additional cases of flat warts in renal transplant patients [32]. Related to HPV49 are HPV-75 (VS40) and HPV-76 (CR148), from a dysplastic wart biopsy and a skin wart biopsy, respectively, from renal transplant patients [6].

HPV-24 was isolated from macules in an EV patient [29]. HPVVS75 and HPVVS20 appear to be relatively closely related to HPV24, and were isolated from skin wart biopsies of renal transplant patients [6].

HPVRTRX2, HPVRTRX4, and HPVRTRX5 appear to form their own cluster, and were all isolated from cutaneous SCCs of renal transplant patients [5]. HPVRTRX2 and HPVRTRX5 were each isolated from one SCC in each of two patients, and HPVRTRX4 was isolated once, out of 53 SCCs from 26 renal transplant patients.

HPV-5, HPV36 and HPV-47 are close enough to each other to be considered "close types"—sequences that qualify to be distinct types under the criterion of ten percent dissimilarity at the nucleotide level, but between which most of these changes are "silent", causing no difference at the amino acid level (Part III). Also qualifying as close types are HPV-19 and HPV-25, and HPV-14d, HPV-20f and HPV-21.

What's new?

The complete genomes of HPV-20 and HPV-21 have been released since last year's publication and are given on the following pages. In addition new complete genomes for HPV types 22, 23, 24, 36, 37, and 38 are available. The newly designated types HPV-75 and HPV-76 are represented here by the sequences of a fragment of L1. Sequences of fragments of L1 from fourteen potential new types, drawn from renal transplant patients, immunocompetent patients and SCC of the esophagus, are included as well. Please refer to *Human Papillomaviruses 1994* pp. I-H-4 et seq., for the sequences of other members of this group.

References

[1] Gassenmaier A; Fuchs P; Schell H; Pfister H. Papillomavirus DNA in warts of immunosuppressed renal allograft recipients. Archives for Dermatological Research. Archiv fur Dermatologische Forschung, 1986, **278**(3):219–23.

[2] Blessing K; McLaren KM; Morris R; Barr BB; Benton EC; Alloub M; Bunney MH; Smith IW; Smart GE; Bird CC. Detection of human papillomavirus in skin and genital lesions of renal allograft recipients by in situ hybridization. Histopathology, 1990 Feb, **16**(2):181–5.

[3] Bouwes Bavinck JN. Epidemiological aspects of immunosuppression: role of exposure to sunlight and human papillomavirus on the development of skin cancer. Human and Experimental Toxicology, 1995 Jan, **14**(1):98

[4] Hartevelt MM; Bavinck JN; Kootte AM; Vermeer BJ; Vandenbroucke JP. Incidence of skin cancer after renal transplantation in The Netherlands. Transplantation, 1990 Mar, **49**(3):506–9.

[5] Berkhout RJ; Tieben LM; Smits HL; Bavinck JN; Vermeer BJ; ter Schegget J. Nested PCR approach for detection and typing of epidermodysplasia verruciformis-associated human papillomavirus types in cutaneous cancers from renal transplant recipients. Journal of Clinical Microbiology, 1995 Mar, **33**(3):690–5.

[6] Shamanin V; Glover M; Rausch C; Proby C; Leigh IM; zur Hausen H; de Villiers EM. Specific types of human papillomavirus found in benign proliferations and carcinomas of the skin in immunosuppressed patients. Cancer Research, 1994 Sep 1, **54**(17):4610–3.

[7] Togawa K; Rustgi AK. A novel human papillomavirus sequence based on L1 general primers. Virus Research, 1995 MAY, **36**(2-3):293–297.

[8] Lutzner,M.A. Epidermodysplasia verruciformis. An autosomal recessive disease characterized by viral warts and skin cancer. A model for viral oncogenesis. *Bull Cancer (Paris)* **65**: 169–82 (1978)

[9] Gassenmaier,A., Lammel,M., and Pfister,H. Molecular cloning and characterization of the DNAs of human papillomaviruses 19, 20, and 25 from a patient with epidermodysplasia verruciformis. *J Virol* **52**: 1019–23 (1984)

[10] May,M., Helbl,V., Pfister,H., Fuchs, P.G. Unique topography of DNA-protein interactions in the non-coding region of epidermodysplasia verruciformis-associated human papillomaviruses. *J Gen Virol* **72** (Pt 12): 2987–97 (1991)

[11] Ensser,A. and Pfister,H. Epidermodysplasia verruciformis associated human papillomaviruses present a subgenus-specific organization of the regulatory genome region. *Nucleic Acids Res* **18**:3919–22 (1990)

[12] Krubke,J., Kraus,J., Delius,H., Chow,L., Broker,T., Iftner,T., and Pfister, H. Genetic relationship among human papillomaviruses associated with benign and malignant tumours of patients with epidermodysplasia verruciformis. *J Gen Virol* **68**(12): 3091–103 (1987)

[13] Deau MC; Favre M; Jablonska S; Rueda LA; Orth G. Genetic heterogeneity of oncogenic human papillomavirus type 5 (HPV5) and phylogeny of HPV5 variants associated with epidermodysplasia verruciformis. *Journal of Clinical Microbiology*, 1993 Nov, **31**(11):2918–26.

[14] Deau MC; Favre M; Orth G. Genetic heterogeneity among human papillomaviruses (HPV) associated with epidermodysplasia verruciformis: evidence for multiple allelic forms of HPV5 and HPV8 E6 genes. *Virology*, 1991 Oct, **184**(2):492–503.

[15] Kawase,M., Favre,M., Pehau-Arnaudet,G., Durand,E., Majewski,S., Jablonska,S., and Orth,G. Unpublished data presented at the 14th International Papillomavirus Conference, Quebec City, July 23–28, 1995. Abstract title: Phylogeny and antigenic properties of Human Papillomavirus type 5 (HPV5).

[16] Heinzel,PA., Chan,SY., Ho,L., O'Connor,M., and others. Variation of Human Papillomavirus type 6 (HPV-6) and HPV-11 genomes sampled throughout the world. *Journal of Clinical Microbiology*, 1995 Jul, **33**(7):1746–1754.

[17] Ho L; Chan SY; Burk RD; Das BC; Fujinaga K; Icenogle JP; Kahn T; Kiviat N; Lancaster W; Mavromara-Nazos P; et al. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *Journal of Virology*, 1993 Nov, **67**(11):6413–23.

[18] Ong CK; Chan SY; Campo MS; Fujinaga K; Mavromara-Nazos P; Labropoulou V; Pfister H; Tay SK; ter Meulen J; Villa LL; et al. Evolution of human papillomavirus type 18: an ancient phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic groups. *Journal of Virology*, 1993 Nov, **67**(11):6424–31.

[19] Chan,SY., R bben,A., Grussendorf-Conen,E., Chew,SH., and Bernard,HU. Unpublished data presented at the 14th International Papillomavirus Conference, Quebec City, July 23-28, 1995. Abstract title: Molecular variants of Human Papillomavirus types 2, 24 and 57 found in cutaneous and genital lesions.

[20] Kiyono,T., Hiraiwa,A., and Ishibashi,M. Differences in transforming activity and coded amino acid sequence among E6 genes of several papillomaviruses associated with epidermodysplasia verruciformis. *Virology* **186**: 628–39 (1992)

[21] Pfister, H. Papilomaviruses: general description, taxonomy, and classification. In *The Papovaviridae 2* (N. P. Salzman and P.M. Howley, Eds.), pp. 1–38. Plenum, N.Y.

[22] Zachow,K.R., Ostrow,R.S. and Faras,A.J. Nucleotide sequence and genome organization of human papillomavirus type 5. *Virology* **158**: 251–4 (1987)

[23] Orth,G. Epidermodysplasia verruciformis: a model for understanding the oncogenicity of human papillomaviruses. *Ciba Found Symp* **120**: 157–74 (1986)

[24] Yabe,Y., Tanimura,Y., Sakai,A., Hitsumoto,T., and Nohara, N. Molecular characteristics and physical state of human papillomavirus DNA change with progressing malignancy: studies in a patient with epidermodysplasia verruciformis. *Int J Cancer* **43**: 1022–8 (1989)

[25] Yabe,Y., Sakai,A., Hitsumoto,T., Kato,H., and Ogura,H. A subtype of human papillomavirus 5 (HPV-5b) and its subgenomic segment amplified in a carcinoma: nucleotide sequences and genomic organizations. *Virology* **183**: 793–8 (1991)

[26] Deau,M.C., Favre,M., and Orth, G. Genetic heterogeneity among human papillomaviruses (HPV) associated with epidermodysplasia verruciformis: evidence for multiple allelic forms of HPV5 and HPV8 E6 genes. *Virology* **184**: 492–503 (1991)

[27] Barr,B.B., Benton,E.C., McLaren,K., Bunney,M.H., Smith,I.W., Blessing,K., Hunter,J.A. Human papilloma virus infection and skin cancer in renal allograft recipients. *Lancet* **1**: 124–9 (1989)

[28] Kremsdorf,D., Jablonska,S., Favre,M., 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)

[29] Kremsdorf,D., Favre,M., Jablonska,S., Obalek,S., Rueda,L.A., Lutzner,M.A., Blanchet-Bardon,C., Van Voorst Vader,P.C. and Orth,G. Molecular cloning and characterization of the genomes of nine newly recognized human papillomavirus types associated with epidermodysplasia verruciformis. *J Virol* **52**: 1013–8 (1984)

[30] Yutsudo,M and Hakura,A. Human papillomavirus type 17 transcripts expressed in skin carcinoma tissue of a patient with epidermodysplasia verruciformis. *Int J Cancer* **39**: 586–9 (1987)

[31] Scheurlen,W., Gissmann,L., Gross,G., and zur Hausen,H. Molecular cloning of two new HPV types (HPV 37 and HPV 38) from a keratoacanthoma and a malignant melanoma. *Int J Cancer* **37**: 505–10 (1986)

[32] Favre,M., Obalek,S., Jablonska,S., and Orth,G. Human papillomavirus type 49, a type isolated from flat warts of renal transplant patients. *J Virol* **63**: 4909 (1989)

```
LOCUS       HPV20          7757 bp    DNA              VRL        18-JUL-1995
DEFINITION  Human papillomavirus type 20, complete genome.
ACCESSION   U31778
KEYWORDS    .
SOURCE      Human papillomavirus type 20.
REFERENCE   1  (bases 1 to 7757)
  AUTHORS   Delius,H.
  JOURNAL   Unpublished, Sequenced by Hajo Delius, Deutsches
            Deutsches Krebsforschungzentrum, Angewandte Tumorvirologie,
            I.N.F. 506, W-6900 Heidelberg, Germany
REFERENCE   2  (sites)
  AUTHORS   Kremsdorf,D., Favre,M., Jablonska,S., Obalek,S., Rueda,L.A.,
            Lutzner,M.A., Blanchet-Bardon,C., Van Voorst Vader,P.C. and Orth,G.
    TITLE   Molecular cloning and characterization of the genomes of nine newly
            recognized human papillomavirus types associated with
            epidermodysplasia verruciformis
  JOURNAL   Journal of Virology 52 (3), 1013-1018 (1984)
REFERENCE   3  (sites)
  AUTHORS   Gassenmaier,A., Lammel,M. and Pfister,H.
    TITLE   Molecular cloning and characterization of the DNAs of human
            papillomaviruses 19, 20, and 25 from a patient with epidermo-
            dysplasia verruciformis
  JOURNAL   Journal of Virology 52 (3), 1019-1023 (1984)
REFERENCE   4  (sites)
  AUTHORS   Kiyono,T., Hiraiwa,A. and Ishibashi,M.
    TITLE   Differences in transforming activity and coded amino acid sequence
            among E6 genes of several papillomaviruses associated with
            epidermodysplasia verruciformis
  JOURNAL   Virology 186 (2), 628-639 (1992)
  MEDLINE   92124737
REFERENCE   5  (bases 1 to 7757)
  AUTHORS   Farmer,A.D.
    TITLE   Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     HPV20 was originally isolated from skin warts of epidermodysplasia
            verruciformis (EV) patients [2,3]. It has additionally been
            detected in a squamous cell carcinoma from another EV patient,
            although the association is not frequent. Hybridization assays
            and phylogenetic reconstructions based on DNA sequences indicate
            that HPV20 is most closely related to HPV21 and HPV14, and then to
            HPV19 and HPV25. This grouping agrees with assays of the degree of
            transforming activity of the E6 protein (these related HPV types
            had relatively low transforming activity as compared to HPVs 5, 8,
            and 47), and clustering of similarity of amino acids in the second
            zinc finger domain of E6 [4].  The E6 gene of HPVs 14, 21, and 25
            can enhance the induction of anchorage independent growth of 3Y1
            cells by the HPV16 E7 gene, although again less effectively than
            that of HPVs 5, 8, and 47. HPV20 was cloned via AvaI restriction.
            But contrary to the assumption that type 20 had only one AvaI site
            (Kremsdorf et al., 1984) the sequence analysis of the clone showed
            the presence of two additional AvaI fragments of 16 and 176 nt,
            respectively, at the cloning site (position 1158 in the final
            sequence) in opposite orientation relative to the large AvaI
            fragment containing the major part of the viral genome. The segment
            between the AvaI sites at position 1142 and 1334 is inverted in the
            pBR322 clone. This inversion leads to disrupted E7 and E1 ORFs in
            the clone. The sequence has been fixed to yield colinearity with
```

```
                    the closely related HPV types.
BASE COUNT     2431 a    1510 c    1698 g    2118 t
ORIGIN      200 bp upstream from beginning of E6 cds
        1 tcgggcgcgg tcatacatta ctcatttggt agttgttgtt gccagctacc atcaagcata
       61 gcatgttttt gcctgtaacg ttatcggcac agtgattaat atatatatat atatatatat
                      dinucleotide "TA" repeat region ->
      121 atatatatat atatatatat atatatatat agatacatat agacagatat catagagcTA
      181 Atgcagagag tgcaggcacA TGgctacacc tccttcttca gaagacagcg ctgatgaagg
    E6 orf ->         E6 cds ->
     start
      241 accatctaat attggagagg caaaacctcc aatcttagag ccaccattgc ctgcaacaat
      301 ctgtggccta gcaaaacttt tagaaatacc gctagatgat tgtttgatac cttgtaactt
      361 ctgcggtaat ttccttacac atttagaagt ttgtgagttt gatgagaaga agcttacttt
      421 aatttggaaa gatcatttgg tttttgcatg ctgtcgtgtt tgctgctcgg caacagcgac
      481 atatgagttt aatcaatttt atgagagtac tgtttttaggc agagacatag agcaagtaac
      541 aggcaaatct gtttttgata tagatgtcag gtgctacacc tgtatgaaat ttttagactc
      601 aattgaaaag ctagacatct gtggcagaaa gcgtccattt tatttagtga gaggctcttg
      661 gaaaggaatc tgtaggctgt gTAAgcattt tcaaTAATGa ttggtaaaga ggtcacattg
                      E7 orf start ->          <- E6 end
                                    E7 cds ->
      721 caagatattg tgctggagtt aaatgaattg cagcctgagg ttcaaccagt tgacctgttt
      781 tgtgaagagg agttaccgaa cgagcagcag gagagagagg aggagcctca gattgaaaga
      841 gcctcataca aagttgttgc accttgcggc tgctgcaagg tgaaacttcg catctttata
      901 agcgctacag aatttgctat tagaagcttt caacaattgc TGAttgacga gctgcagctg
                                       E1 orf start ->
      961 ttgtgtcctg actgtcgcgg gaactgcaaa cATGgcggat ccTAAaggta gtacatcaa
                           E1 cds ->            <- E7 end
     1021 agacgggttg gatgattggt gtattgttga agctgaatgt agcgatgtag acaatgattt
     1081 ggaagaatta tttgacagag atacagactc agatatttca gaattattag atgataatga
     1141 cctcgagcag ggcaattctc gggaactatt tcatcaacaa gagtgtaagg acagcgagga
     1201 gcaattacaa aaactaaaac gaaagtacat aagtccaaaa gctattgcac agcttagtcc
     1261 gcgacttgaa agtatttcac tgtcaccaca gcagaagtca aaacgaaggc ttttttgcaga
     1321 gcaggacagc gggctcgagt taactcttac aaatgaagct gaagatgttt cttctgaggt
     1381 ggaggaggta ccggccctag actctcagcc ggttgctgag ggacacttag aacagtagaa
     1441 cattcattat acagaattat tgcgtgccag taaccataag gcaattttgt tggcaaaatt
     1501 taaggaggct tttgggatag ggtttaatga tttgacacgt caatttaaaa gttacaaaac
     1561 ctgctgtaat gattgggttc tatctgtgta tgcagttcat gaggatcttc ttgaaagctc
     1621 aaagcagtta ttgcaacagc attgtgatta tatatggatc cgtgggatag cagcaatgtc
     1681 attgtttcta ttgtgtttta aagcaggaaa aaatcgtggg actgtgcata aattaatgac
     1741 atcaatgttg aatgtgcatg aaaagcaaat attgtctgag cctccaaaat taagaaatgt
     1801 tgctgctgct ttattttggt ataaaggtgc aatggggtcc ggagcatttt ctcatggtcc
     1861 atatcctaac tggatggcac agcaaactat tgttggtcat cagagcacag aagccagtgc
     1921 ttttgacttg tctgaaatga ttcagtgggc atttgaccat aattatctag atgaggctga
     1981 tatagccttt cagtatgcta agctagcacc agaaaatagt aatgctgtag catggcttgc
     2041 acataataAC CAAGCAAGGT ttgttagaga atgtgcatca atggtcaggt tttataaaaa
      E2 binding ->
     2101 aggtcaaatg aaagaaatga gcatgtcaga atggatttat gccagaatta atgaagtaga
     2161 aggcgaagga cattggtcat ctattgctaa atttcttaga tatcagcaag taaatgttat
     2221 aatgtttta gctgctttga aagatatgct gcattctgta cctaaacata actgtatatt
     2281 aatacatggc ccacctaata ctggaaaatc tgcattcact atgtcattga tacatgtgtt
     2341 aaagggaagg gtattgtcct ttgtaaattc taaaagccaa ttctggttac aaccaatgtc
     2401 agaaactaaa atagcattaa ttgatgacgt aactgatcct tgctgggttt atatggatac
     2461 atatttaaga aatggcttag atggacatta tgtctcacta gattgcaagc ataaagcacc
     2521 aattcaaaca aaatttcctg cattactgct tacctctaat attaatgttc ataatgaagt
     2581 taactataga tatttacata gtagaattaa aggatttgaa tttccaaatc catttccaat
     2641 gaaaccagac aatacccctg agtttgagct tactgaccaa agctggaaat cttttttttac
     2701 aaggctttgg aagcaattag agcTGAgtga ccaagaagac gagggagaaa ATGgagaatc
                      E2 orf start ->                 E2 cds ->
```

```
2761 tcagcaagcg tttcaatgct ctgcaagatc agctaatgaa catttaTGAg tctgcaccag
                                                          <- E1 end
2821 acactcttga gtcgcaaatt gagcactggc aaaccctgcg aaaagaagct gtgctactat
2881 attttgctag gcaacatggt atcagcaggg ttggatatca acctgtgcct gtattagctg
2941 tgtcagaagc caaagctaaa caggctatag gaatggtatt aaggttacaa tcattgcaaa
3001 aatctgaata tggaagtgaa ccatggtctt tggtagatgc aagtgcagag acatttagaa
3061 gcccgccaga aaatcacttt aaaaaaggtc cgatttcagt agaggtcata tatgacaaag
3121 ataaagacaa tgccaatgct tataccatgt ggagatttgt ttattaccaa gatgatgacg
3181 acaagtggca caaaagtgct agtggtgtta accaaacagg catatatttt atgcaaggaa
3241 catttagaca ctactatgtt ttgtttgctg atgatgcgag tagatatagt acaactggac
3301 aatgggaagT GAaagttaat aaggaaactg tgtttgctcc tgtcaccagc tccacccccc
     E4 orf start ->
     NH₂ terminus unknown
3361 ccgactcacc aggaggacaa gcagactcaa acgcctcctc ccagacccccc gccaccacca
3421 ctgactccac gaccagacag tcgcccagaa aacagtcaca acaaaccaac accaaaggga
3481 gaaggtacgg acggagacct tccagtagga caaggcgaac aacccaaacg cgccagaggc
3541 gacggtccag gtcaaagtcc aagtccaagt ccaggtcgcg gtcgaggtcg cggcACCGGT
                                                          E2 binding ->
3601 CTCGGTctcg gtctcggtct gaatcgccgc gccggcggtc tcggtaccga tcacgatccg
3661 gatccagagg gagagtcgcc ctccgcgcca ttaccaccac caccacaacc accaccagac
3721 gggcaggtgg agggtcaccc acctccacct cctccaccac ctcacaacgg tcgcgacagc
3781 tgcggggagg gggccgtggg gggagcagac aaaagagcaag gggaaggcga tcatcatcca
3841 cctcccccac cccctcaaaa cggtcacgag gggagtcaga gtctgttagg caacatggca
3901 tctctccttc tgacgtggga acagcagttt acacagttag ttcaagacat acaggaagac
3961 ttggaagatt actggatgaa gctctcgatc ccccagTGAt tttagttagg ggagagccta
                                                 <- E4 end
4021 atacgcttaa gtgctttcgc aatagggcca aacaaagata tacagggctg tataagtctt
4081 ttagcacggc ctggtcgtgg gtggctggag atggcacgga gcgtctaggc aggtccagaa
4141 tgctcattag cttttatatcc ttcagtcaaa gaaaagattt tgatgagact gtgaaatatc
4201 cgaaggggggt tgaccggtcg tttggttcat ttgacagctt aTAGcaacct aaccttctaa
                                                    <- E2 end
4261 ccactgcatg ctactaacac actaacattt tttaattttt atTAAtattt tttatttgct
                                       L2 orf start ->
4321 ATGgcgcgcg ctaagcgagt caagcgggac tctgctacta acatatacag aacctgcaaa
     L2 cds ->
4381 caagcaggta cttgtcctcc tgatgttata AATAAAgtgg aaagcacaac tattgctgat
                            signal ->
4441 aaaattttgc agtatggtag tgctggtgtt tttttttggggg gattaggcat aagcactgga
4501 aaaggtacag gaggaaccac aggttatgtg cctttgggag aaggcccatc ggtgcgtgtt
4561 ggtggtacac ctacagtcat acgACCTGCT TTGGTcccag acaccatcgg cccctccgat
                E2 binding ->
4621 attatacctg tggacacctt aaatccggtg gagccttcta cctcttctat tgttccactt
4681 acagaatcca caggaccaga tcttttacct ggtgaagtgg aaactattgc agaaatacat
4741 ccaggccccct caaggccacc aactgataca ccagttacat ctactaccag tggttctagt
4801 gcagttctag aggtagcacc agaaccaaca cctccagctc gtgtcagagt cagccgcacc
4861 cagtatcata acccatcatt tcaaataata actgaatcaa caccaacatt ggggggaaagc
4921 tcattagcgg atcatatagt agtgacatct ggttctgggg gccaagcaat tggggggatg
4981 acacctgaac ttatagagct tcaggatttc ccatcaaggt attcatttga aatagaagag
5041 ccaaccccctc ctagaagaac tagcacacct atgcaaagac ttcaaaatgt gttcaggcgt
5101 agaggaggcc ttactaacag aagattagtt caacaagtgc ctgtagacaa tccattattt
5161 ttgacacaac cttctagatt ggtccggttt cagtttgata cccggttttt tgaggaagaa
5221 gttactcaaa tatttgaaca agatttagac actttttaatg agcccccaga cagagacttt
5281 ttggatgttc agagtttagg caggcctcaa tactcagaaa ctcctgcagg ttatgtgcgg
5341 gtcagccgtg caggtcaacg aaggactatc agaactcgtt ctggagcaca aataggggtct
5401 caagtgcact tttatagaga tctcagtagt attgatacag aagatcctat tgaactgcag
5461 ttgttgggggtc agcattctgg cgatgcaact attgtccaag gtccagtaga aagcactttt
5521 gttgatatca atgtagatga aaacccactt tcagaaatca gtgcatattc tgatgattta
5581 cttttagatg aagctaatga agactttagt ggctctcagt tagttgtagg gggaaggcgt
5641 tctacatcta catacactgt tcctcacttt gaaactacta gatctagctc ttactatgta
```

```
5701 caagatacaa aggggtatta tgtagcatat cctgaagata gagatgttag taaggacatt
5761 atttatccta atccagattt accagtggtc attattcaca catatgacac aagtggagat
5821 ttttatttac atccaagtct tactaaaaga ttaaaaagaa aaaggaaata tttgTAActt
                                                              <- L2 end
                                                      L1 orf start ->
5881 tttctttttgc agATGgcagt ttggcaagca gctagtggta aggtgtacct tccaccatct
              L1 cds ->
5941 acaccagttg ccagggtcca aagtacggat gaatatgtgc aaaggactaa catatactat
6001 catgcataca gtgatcgcct actaactgtt ggtcatccat attttaatat atatgacatc
6061 caaggcacta agataaaagt ccctaaggtt tctggaaatc agcacagagt gtttaggtta
6121 aaactaccag atcccaacag atttgcatta gcagatatgt ctgtgtataa cccagataaa
6181 gaaagattgg tctggggctg tagaggtata gaaataggta gaggacagcc attaggcgtt
6241 ggaagtgtag gtcatccatt atttaataaa cttggtgaca cagaaaaccc taattcatat
6301 aaagggaatt caactgatga tagacaaaat gtatcttttg accctaaaca actacaaatg
6361 tttataatag gctgtgcccc atgtttagga aacattggg acagggcttt accatgtgca
6421 gacgacgttc caaacccagg ttcatgccct ccaatagaat taaaaaatac agcaatacaa
6481 gatggcgata tggcagatat aggatatggc aacctaaatt ttaaagcatt acaagaaaac
6541 agagcagatg taagtttgga tattgttaat gagacctgta aatatccaga ctttttaaaa
6601 atgcagaatg atgtttatgg agattcctgc tttttttatg ctcggcggga acaatgttat
6661 gctagacact tttttgtacg tggggggcaaa acaggagatg atatacctgc aggacaaatt
6721 gatgaaggta gcatgaagaa tgcattctac attccacctg tgaataatca ggcacagaac
6781 aacctaggta attcaatgta tttcccaact gtcagtggct cattggtgtc tagtgatgct
6841 caattgttta ataggccatt ttggctgcag cgcgcacagg gccacaacaa tggcatctgc
6901 tggttcaatc aactatttgt tactgtagta gataatactc gaaatacaaa ttttagcata
6961 tcagttcatt cagaaaacac tgatgtttct aaaattcaaa attatgattc tcagaaattt
7021 caagaatatt taagacacgt agaagaatat gaaatttcat taattttaca gctctgtaaa
7081 gttcctttaa cagctgaagt tttagctcaa attaatgcta tgaattcaaa tatattagag
7141 gagtggcagt taggattcgt tcctgcaccg gataatccta tccacgatac atacagatat
7201 attaattctg cagctactag atgtcctgat aaaaatcctc aaaagaaag agaagatcct
7261 tacaaggatc taaacttttg gaatgttgac ctatcagaaa gattatcctt agaattggat
7321 caatattctt taggacgcaa attcttattt caagcaggtt tacaacaagc gACCGTAAAC
                                                            E2 binding ->
7381 GGTacaaaaa ctgtatcttc aaagttatct actagggggcg tcaaacgaaa acgcaaacaa
7441 TAAacccgAC CGTTTTCGGT acAATAAAgt caactttac acggtattca aggaatgttt
        <- L1 end         signal ->
  E2 binding ->
7501 atttactctg actaactaag ataccaaccg cacccgacac ataaaggtga gttgtgtgcc
7561 aaatgaggtg agttgtgagc cagaagagat cacagccaag tcaggcttga gccagatcag
7621 atacactgcg tgccagagtt ggctcaaact tcatcgtccc aacacgttcg gaacaggagg
7681 aaatgtaagg ctgccaacgc ttttggctct cttttttggc acagcagaag ACCGTTAACG
                                                            E2 binding ->
7741 GTaagttttt atttgta
//
```

```
LOCUS       HPV21           7779 bp    DNA             VRL        18-JUL-1995
DEFINITION  Human papillomavirus type 21, complete genome.
ACCESSION   U31779
KEYWORDS    .
SOURCE      Human papillomavirus type 21.
REFERENCE   1  (bases 1 to 7779)
  AUTHORS   Delius,H.
  JOURNAL   Unpublished, Sequenced by Hajo Delius, Deutsches
            Deutsches Krebsforschungzentrum, Angewandte Tumorvirologie,
            I.N.F. 506, W-6900 Heidelberg, Germany
REFERENCE   2  (sites)
  AUTHORS   Kremsdorf,D., Favre,M., Jablonska,S., Obalek,S., Rueda,L.A.,
            Lutzner,M.A., Blanchet-Bardon,C., Van Voorst Vader,P.C. and Orth,G.
  TITLE     Molecular cloning and characterization of the genomes of nine newly
            recognized human papillomavirus types associated with
            epidermodysplasia verruciformis
  JOURNAL   Journal of Virology 52 (3), 1013-1018 (1984)
REFERENCE   3  (sites)
  AUTHORS   Kiyono,T., Hiraiwa,A. and Ishibashi,M.
  TITLE     Differences in transforming activity and coded amino acid sequence
            among E6 genes of several papillomaviruses associated with
            epidermodysplasia verruciformis
  JOURNAL   Virology 186 (2), 628-639 (1992)
  MEDLINE   92124737
REFERENCE   4  (bases 1 to 7779)
  AUTHORS   Farmer,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     HPV21 was originally isolated from skin warts of an
            epidermodysplasia verruciformis (EV) patient [2]. Hybridization
            assays and phylogenetic reconstructions based on DNA sequences
            indicate that HPV21 is most closely related to HPV14 and HPV20, and
            then to HPV19 and HPV25.  This grouping agrees with assays of the
            degree of transforming activity of the E6 gene (these related HPV
            types had relatively low transforming activity as compared to HPVs
            5, 8, and 47), and clustering of similarity of amino acids in the
            second zinc finger domain of E6 [3]. The E6 gene of HPVs 14, 21,
            and 25 can enhance the induction of anchorage independent growth of
            3Y1 cells by the HPV16 E7 gene, although again less effectively
            than that of HPVs 5, 8, and 47.
BASE COUNT     2426 a   1518 c   1680 g   2155 t
ORIGIN      200 bp upstream from beginning of E6 cds
        1 acggtaagtt atgcACCGGG TGCGGTcgaa ttattactca ttcgatagtt gttgttgcca
               E2 binding ->
       61 gctaccattt aggacagcat gtttttgcct gTAAcgttat cgacacatac tcacaccata
                               E6 orf start ->
                                       dinucleotide "TA" repeat region ->
      121 tatatatata tatatatata tatatatata tatatatata tattcaTATA TAcatactag
                                                  signal ->
      181 ggaagatgcc ctagtactcA TGgctgactc ttcaacagac agtgctgacg aaggtccttc
                    E6 cds ->
      241 tcctaagcgt agacatttag aagaagaaaa tacatctagc tttttagagc caccattacc
      301 agctacaatt cgtgacctag ccaatctgtt agagatacca ttggatgatt gtttagtacc
      361 ttgtaacttt tgcggtaatt ttcttactca tttagaagtt tgtgagtttg atgagaaaaa
      421 gcttagttta ctttggaaag atcattgtgt gtttgcctgt gtcgtgtttg ttgcgcagc
      481 aacagcgaca tatgaatata atgaatttta tgaatctact gttgtaggta gagatataga
      541 agaaataaca ggcaaatcta tttttgatat tgatgtcagg tgctacaatt gcatgaaatt
```

```
  601 tttagactca atagaaaagc tagacatttg tggtaggaag ttttttttttc ataaagtgag
  661 aggctcttgg aaaggaatct gtaggctgtg TAAgcatttt caaTAATGat tggtaaagag
                             E7 orf start ->          <- E6 end
                                               E7 cds ->
  721 gtcacattgc aagatattgt tctggagtta aatgaattgc agcctgaggt acaaccagtt
  781 gacctgtttt gtgaagagga gttaccgagc gagcagcagg aaacagagga ggagctacca
  841 gaaaggaccg cgtacaaagt tgttacacct tgcggctgct gcaaggtcaa gcttcgcatc
  901 tttgTAAacg ctacacaatt tgctattaga acatttcaga atctgctgtt tgaagaattg
E1 orf start ->
  961 cagctgttgt gtcctgagtg ccgcggaaac tgcaaacATG gcggatccTA Aaggtagtac
                                             E1 cds ->          <- E7 end
 1021 atctaaagaa gggttggagg attggtgtat tgtggaagct gaatgtagcg atgtagaaaa
 1081 tgatttggaa gaattatttg acagagatac agactcagat atttcagaat gttagatga
 1141 taatgacctg gagcaggaa attcgcggga actatttcac cagcaagaga gtaaggaaag
 1201 cgaggagcaa ttacaaaaac taaaacgaaa gtacttaagt cctaaagctg tcgcacagct
 1261 cagtccgcga ctcgaaagta taacgctgtc acctcagcag aagtctaaac gaaggctctt
 1321 tgcagagcag gacagcgggc tcgagtgtac tcttacaaat gaagaagatg tttcttctga
 1381 ggtggaggta ccggctctag actctcagcc ggttgctgag gcacaattag gaacagtaga
 1441 cattcattat aaagagttat tacgtgccag caacaataag gcgattctta tggcaaaatt
 1501 taaagagttt tttggggtag gatttaatga tctgacacgc aatttaaaaa gttacaaaac
 1561 ctgttgtaat gcttgggttc tgtctgtata tgcagttcat gatgatcttc ttgaaagctc
 1621 aaagcagtta ttgcaacagc attgtgatta tatatggata cgtgggatag gagcaatgtc
 1681 attgttttg ttatgtttta agttggaaa aaatcgtggg actgtgcata gttgatgac
 1741 tgcaatgtta aatgtgcatg aaaagcagat catatctgag ccaccaaaat taagaaatgt
 1801 tgctgctgca ttgttttggt ataaggggtgc gatggggtct ggagcattta cttatggacc
 1861 ttatcctgat tggattgccc agcaaacaat tgttggtcat caaagtacag aagccagtgc
 1921 atttgatatg tctgcaatgg ttcaatgggc gtttgataat aactatttag atgaagctga
 1981 tatagcctat caatatgcta agctagcacc agaagatagt aatgctgtag catggcttgc
 2041 acataataat caggccagat atgttagaga agttgcatct atggtaagat tttataaaaa
 2101 aggacaaatg aaagaaatgt ctatgtcaga gtggatacat actagaatta atgaagtaga
 2161 aggagaagga cattggtcaa ctatagcaaa gttccttaga tatcagcaag taaattttat
 2221 aatgtttcta gcagcattaa aagacatgct acattcagtt cctaaacgta attgtatatt
 2281 gatttatggt cccccctaaca ctggaaagtc agcatttact atgtctttaa ttcatgtact
 2341 aagagggagg gtgctatcat ttgtgaattc caaaagccag ttttggctgc agccaatgtc
 2401 agaatgtaaa atagcattaa ttgatgatgt gacagatcca tgctggatat atatggatac
 2461 ttatttaaga aatggcctag atggtcatgt tgtatcatta gactgcaaac ataaagcacc
 2521 gatgcaaacc aaatttcctg cattactact tacatctaat atcaatgtgc ataatgaagt
 2581 taattataga tatttgcata gcaggattaa aggctttgaa ttcccaaatc catttcccat
 2641 gaaagcagac aatacccctg aatttgagct tactgaccaa agctggaaat ctttttttac
 2701 aaggctttgg aatcaattag agcTGAgtga ccaagaagac gagggagaaa ATGgagaatc
                      E2 orf start ->                      E2 cds ->
 2761 tcagcgatcg tttcaatgtt ctgcaagatc agctaatgaa catttaTGAg tctgcagcaa
                                                        <- E1 end
 2821 acactattga gtcgcaaatt gagcattggc aaacactgcg aaaagaagct gtgctgcttt
 2881 attttgctag gcaaaagggt gtgacacggc ttggatatca atatgtacct ccattagcag
 2941 tttcagaatc aagagctaaa caggctatag gatgatgct gcagttgcaa tcattgcaaa
 3001 aatctgaata tgcaaaggaa ccatggtcac tggTAGatac cagtgcagag acatttagaa
                         E4 orf start ->
                         NH2 terminus unknown
 3061 gccctcctga aaatcatttc aaaaaagggc cagtgtcagt tgaggttatt tatgataacg
 3121 ataaagacaa tgctaatgct tacaccatgt ggagatatgt ttattacgtg gatgatgacg
 3181 accaatggca taaaagtcca agcggtgtca accacacagg catatatttt atgcaaggaa
 3241 cttttagaca ctactatgtt ttatttgctg atgatgcaag tagatatagc agaactggac
 3301 attgggaagt taacgttaat aaggaaactg tgtttgctcc tgtcaccagc tccaccccac
 3361 ccgactcacc aggaggacaa gcagactcaa acacctcctc cacgaccccc gccaccacca
 3421 ctgactccac gtccagactc tcgtccacca gaaaacagtc acaacaaacc aacaccaaag
 3481 gaagaaggta cggacgagga ccgtccagta ggacccggcg aacgacccaa acgcatcaaa
 3541 ggcggcgatc gaggtccaag tccaggtcgc ggtcgcggtc gcggtcgcgg ctccgatccc
 3601 gatcccggtc ccgatcccgg tcctattccc ggtcccggtc tcaatcgtct gaccagccgc
```

```
3661 aataccgatt cagatccgga gggcaagtgt ccctcatcac taccgccacc accaccacca
3721 ccaccgcaac caactactcc accagagggt cagggcgagg gtcatcctcc acctcctcct
3781 ccacctccaa acggccacga cggccacgag gaggggccat tggagggagc agtgggaggg
3841 ggagacggtc atcctccacc tcccccagcc cctccaaacg gtcacgagga aagtcagagt
3901 ctgttaggca acgtggcatc tctcctgacg acgtgggaaa gtctcttcaa tcagttagta
3961 caagaaatac aggtcgactt ggaagattac tggacgaagc tctcgatccc ccagTAAtct
                                                          <- E4 end
4021 tagtcagggg ggaacccaat acgctaaaat gctttcgcaa tagagccaag cttaaatacg
4081 cagggttgta taaggctttc agtacggcct ggtcgtgggt ggctggagat ggtactgagc
4141 gtctaggcag gtccagaatg ctcattagct tcttctcctt tgagcaaaga aaagatTTTG
4201 ataagactgt taaatatccg aaaggtgttg accggtcgta tggttccttt gatagcctaT
4261 AGcagccttT AAcatactaa ctatagctct gctactaaca tattaacact ttttgatTAT
     <- E2 end                                                  signal ->
     L2 orf start ->
4321 ATATtttttt tttatttttta tttttatgct ATGgcgcgtg ctaagcgagt caagcgagac
                                L2 cds ->
4381 tctgctacta atatttacag aacctgcaaa caagcaggca catgtccccc tgatgttatt
4441 AATAAAgttg aaagcacaac tattgctgat aaaatattgc agtatggtag tgctggtgtt
signal ->
4501 tttttcgggg ggctgggcat aagcactgga aaaggtacag gcggtaccac aggttatgtg
4561 cctttgggag aaggtcctgc agtccgtgtt ggcaatgctc ctacggtcat tagACCTGCA
                                                        E2 binding ->
4621 TTGGTccctg acaccattgg cccgtctgat attattcctg tggacacctt aaatccagtg
4681 gagcccacaa cttcctctat tgttccactc acagactcta caggcccaga tctgttacct
4741 ggagaagtgg aaactattgc agaaatacat cctggtccga ccaggcctcc acctgacact
4801 gcagtcacta ctagtacaaa tggttctagt gctgttttag aagtagcacc agagcctacc
4861 cctccttctc gtgttagagt aaccagaaca caatatcata atccatcttt tcaagtaata
4921 actgaatcaa ctcctactac aggcgaaagt tctttagcag atcatatatt agtaacatca
4981 gggactgggg gacaaactat aggggggcagt acacctgaac tcatagaact ccaggacttt
5041 ccttctagat attcatttga aattgaggag ccaacacctc ctagaagaac tagtacaccc
5101 attcaaagaa ttcaaaatat tataaggaga aggggtggcg ggctcacaaa taggcgtttg
5161 gttcaacagg ttaatgtaga gaatcctttg tttgtatcca ggccttctag attagtgcag
5221 tttcaatttg ataaccctgc atttgaagaa gaagtgacac aaatatttga gcaagatatt
5281 gatactttca atgaaccacc agatagagac ttttttagata ttaaaacact tggtaggcct
5341 caatactcag aaACCCCTGC AGGTacgtg agagttagtc gtcttggtaa acgaggaact
          E2 binding ->
5401 attcgtactc gttcaggaac acaaattggt tctcaggtcc attttttacag ggaccttagc
5461 accattaaca cagaggaccc tattgaactt caattattgg gtgagcattc tggcgatgct
5521 acaattgtcc agggtccagt tgaaagcaca tttattgata ttaacgttga tgaaaaccct
5581 ctttctgaag attttagtgc acattcagat gatttacttc tagatgaggc aaatgaagat
5641 tttagtggtt cccaattagt ggttggaggc cgccgctcca cttcttctta tactgttcca
5701 cgttttgaaa ctactagatc tggttcttat tacgtgcagg acaccaaggg ctattatgta
5761 gcctatcctg aagatcgaga cactagtaca gatataatct atccaacacc agatttgcca
5821 gttgtaatca tacacacatt tgatacaagc ggtgattttt acttacatcc gagtcttagc
5881 agaaaattta agagaagaag gaaatatttg TAAccttttc ttttgcagAT Ggcagtttgg
                                     <- L2 end  L1 cds ->
                                 L1 orf start ->
5941 caagcagcta gtggtaaggt ttaccttcca ccgtctacac cagttgccag ggtccaaagc
6001 acggatgaat atgtacaaag aacaaacatc tactatcatg catatagtga tcgcttatta
6061 actgttggtc atccatattt taatgtctat gacgtcaata gtgctaagat aaaagtacct
6121 aaagtatctg gaatcaaca cagggtattc agactcaaat tgccagatcc taatagattt
6181 gcacttgcag atatgtctgt atacaatcca gacaaggaaa gattagtttg ggcctgcaga
6241 ggtatagaaa taggaagagg gcaACCCTTG GGGGTgggaa gtgtaggtca ccctttattt
                 E2 binding ->
6301 aataaagttg gggacacaga aaatcctagt tcatacaaaa ctcaaccaaa ttctactgat
6361 gatagacaaa atgtatcatt tgatcccaaa caactacaaa tgtttataat aggctgtgca
6421 ccttgcttag gagaacattg ggataaagct atcccatgtg caactgacaa tccacctcca
6481 ggatcgtgcc ctccgattga attaattaat tcagcaatac aagatggcga tatggcagat
6541 ataggatatg gcaatctaaa tttcaaagcc ttacaacaaa ataggtctga tgttagttta
```

```
6601 gacatagtta atgaaacgtg taagtatcca gacttcttaa aaatgcaaaa tgatgtgtat
6661 ggagattcat gtttctttta tgcacgcaga gagcaatgtt atgccagaca cttctttgtt
6721 agaggggggca aaacaggaga tgacataccc gcaggacaaa ttgatgaggg tagtatgaag
6781 aatgcatatt acattccacc aatgaatgat caagcacagt acaagattgg taactccatg
6841 tatttcccaa ctgtcagtgg ctcattggtg tctagtgacg ctcaattgtt taacaggcca
6901 ttttggctac agcgtgcaca aggccataat aatggcatat gttggtttaa tcaattattt
6961 gttacagtag tagacaacac tcgtaacaca aactttagta tttcagtaaa tcctgagaat
7021 gcagacgtgt ctaaaattga aaattataaa gccgagagct ttcaagaata tttaagacac
7081 gttgaagaat atgaactttc tttaattta caattatgta aagttccttt aacagcagaa
7141 gtcttagctc aaattaatgc aatgaatgca atatttttag aagaatggca gttaggattt
7201 gttcctgccc cagacaatcc tattcatgat acatatagat acattgactc tgcagctact
7261 agatgtcctg ataaaaaccc tccaaaagaa cgagaagatc cttataaaaa tatgaaattt
7321 tgggatgtag atttaacaga acggttgtct ctagacttag atcaatattc tcttggaaga
7381 aaatttttat ttcaagcagg tttgcagcag acgACCGTTA ACGGTacaaa gacactttct
                     E2 binding ->
7441 tcaagggtat ctaccagagg aattaaacga aaacgcaaaa atTAGacatg ACCGTTTTCG
                                             <- L1 end
                                        E2 binding ->
7501 GTacAATAAA gtcaacttt acacagtatt caaggaatgt ttatttactc tgactaagca
  signal ->
7561 aaataccaac cgcgcccgac acataaaggt gagttgtgag ccaaatgagg tgagttgtaa
7621 gccaaaagag gtcagagcca agtctgttct gagccagatc agatactacg cgcgccagag
7681 ttggatcaca tctcgttgtt ctaacacgct aaggactcaa ggaaatgtaa gtctgccaat
7741 cgattttggc tcgtgttttg gcagaagtta ggaccgtta
//
```

```
LOCUS       HPV22       7368 bp    DNA              VRL       18-JUL-1995
DEFINITION  Human papillomavirus type 22, complete genome.
ACCESSION   U31780
KEYWORDS    .
SOURCE      Human papillomavirus type 22.
REFERENCE   1  (bases 1 to 7368)
  AUTHORS   Delius,H.
  JOURNAL   Unpublished, Sequenced by Hajo Delius, Deutsches
            Deutsches Krebsforschungzentrum, Angewandte Tumorvirologie,
            I.N.F. 506, W-6900 Heidelberg, Germany
REFERENCE   2  (sites)
  AUTHORS   Kremsdorf,D., Favre,M., Jablonska,S., Obalek,S., Rueda,L.A.,
            Lutzner,M.A., Blanchet-Bardon,C., Van Voorst Vader,P.C. and Orth,G.
  TITLE     Molecular cloning and characterization of the genomes of nine newly
            recognized human papillomavirus types associated with
            epidermodysplasia verruciformis
  JOURNAL   Journal of Virology 52 (3), 1013-1018 (1984)
REFERENCE   3  (bases 1 to 7368)
  AUTHORS   Farmer,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     HPV22 was originally isolated from macules on the chest of an
            Italian epidermodysplasia verruciformis (EV) patient [2]. The HPV22
            genome, like that of HPVs 9, 15, 17a/b, 23, 37, 38, is smaller than
            most PV genomes at approximately 7.4 kb. Phylogenetic
            reconstructions based on DNA sequences of established types
            indicate that HPV22 is most closely related to HPVs 23 and 38, and
            then to 15, 17, 37 and 9. Although Kremsdorf et al [2] found
            substantial cross-hybridization between HPV22 and HPV19, nucleotide
            sequence comparison fails to support a close relationship between
            these two types.
BASE COUNT     2315 a   1352 c   1614 g   2087 t
ORIGIN      149 bp upstream from beginning of E6 cds
        1 ccgccaaagc tttgccaggt cttggcagaa catttgctgg caaagactgc ACCGATAACG
                                                               E2 binding ->
       61 GTaagaactt ttaattttta ACCGTAGGCG GTtatttgtt attcgtagca acaattgtgg
                    E2 binding ->
      121 tTAAcaacaa tctcctgcca gaatatacAT GcaACCGCTT GTGGTaattt atgcactgct
 E6 orf start ->                  E6 cds ->
                                             E2 binding ->
      181 tgcatattat ttaagtagga tgggctgcta ttctgtattc atggctttgc aaagaccact
      241 gacagtacag caacttagtg ataagttgac tgtacctgta gtagatcttt tgctaccttg
      301 tagattctgc agtaggtttt taacctattt ggaattgcgg caatttgatt ataagaattt
      361 gcaattaatt tggacagacg aggactttgt gtttgcatgt tgcagcggct gtgcctacgc
      421 ttcagcccaa tttgaatttc agcagtatta tcaagttact ttgtatggtc gtgaaattga
      481 gcaagaagaa caacgacctg taggccaaat ttatatgaga tgtcaatatt gcttgaagtc
      541 tcttgatttg ctagaaaagt tagatatctg ctgttccaat caaccatttc acaaggttag
      601 agatcattgg aagggaaggt gcaggcactg TAAagcaata gaATGAttgg gaaacaagct
                                          E7 orf start ->        <- E6 end
                                                       E7 cds ->
      661 actctgtgtg atatagttct tgaagagctt gtcctgccca ttgacctgca ttgccacgag
      721 gagctgcctg aacttccaga agagttagaa gaatcagtgg tagggagga gcctgagtac
      781 actccttaca agattgtagt atattgtggg ggttgtgata caaagctgaa gctgtatata
      841 ctagcaactc tctctggaat tcgcgacttt caaacatctc tacttggacc tgTAAaactt
                                                             E1 orf start ->
      901 ttgtgtccca cctgtcgaga agagattcgc aATGgacgac gaTAAaggta ctgacacaac
                                          E1 cds ->         <- E7 end
```

```
 961 tgatgctaaa gaaggatgta gtggttggtt tatgttagaa gctgcgtgct cagatgatag
1021 tgacttagat aatagtttgg aaaagttatt tgaagatggt acagagtcag atgtatctga
1081 tttaataaat gatgatgata ctgctgctca gggaaattcc cgcgaattgc tatgtcaaca
1141 gcaaagtgag gaatgtgagc agcagattca atatctaaaa cgaaagtatt tcagtccaaa
1201 ggctgttcag cagctaagtc cacgtctgca gtctatgaat atttcgcctg ggcataaatc
1261 taaaaggaga ttatttgtgg agcacgacag cggactggag tgttccctaa atgaagctga
1321 agatcttact gaagaggtgg aggtaccggc gagcgctcca gcgccggcag cacagggtgg
1381 tgtagggtcg ggacattaca ccagtttgtt aagatgcaac aatgtaaagg cagtattgct
1441 gggaaaattt aaagacgcat ttggagtgag ctataatgag ctgactagac aatttagaag
1501 taataagact tgctgtaagc attgggtatt ggccatatat gctgctaaag atgaattaat
1561 agatgcgtcc aaacaattgt tacaacagca ttgtacctat ttgtggttgc aaacattctc
1621 acccatgtca ttatatttat gttgtttttaa tgttgggaaa agtagagaaa ctgtgatgcg
1681 attgttatct tccatgttac aagttaatga gaatcatatt ttatcagaac ctccaaaaat
1741 cagaagtatg atagctgctt tattttggta taaaggtagt atgaatccaa atgtctatgc
1801 atttggagag tatcctgagt ggattatgac acagactatg atacatcacc aaactgctga
1861 cagtgtacaa tttgacctgt ctgaaatgat acaatgggct tatgatcaag attatgttga
1921 tgaatgtact attgcatacc agtatgctag attggctgat agtaatagta atgccagagc
1981 attttttagct cataatagtc aagccaaata tgttagagaa tgtgctcaaa tggttagata
2041 ttataaacgt ggagaaatgc gagatatgtc aatttctgca tggatacatc attgtatatc
2101 aaagatagaa ggcgatggtc actggcaaga tattgttaaa tttttgcgat accaagggtt
2161 aaattttata gtgttttttag ataaatttag aacatttcta aaaaattttc caaagaaaaa
2221 ttgtttgtta atatgtggtc ctccggatac aggaaaatct atgtttagca tgtcattaat
2281 gaaagcatta agaggacagg tagtttcatt tgcaaattct aaaaagtcatt tttggctaca
2341 gcctttagca gatgcaaaac tggctttatt agatgatgct acagaagttt gctggcaata
2401 tattgatgct ttcttaagaa atggattaga tggtaacatg gtatctttag atatgaaaca
2461 tagagctcca tgtcaaatga aatttccacc tcttattata acatctaaca ttagtttaaa
2521 aaaagaaaaa aaatttccct atttacatag tagaatatat gaatttgagt ttcctaacaa
2581 atttcccttt gacgcaaatg atacacctct gtttaaactt actgaccaaa gctgggcgtc
2641 ttttttttaaa aggctttgga cacaattaga acTGAgtgat caagaagaag agggagaaaA
                         E2 orf start ->                    E2 cds ->
2701 TGgagaaact cagcgaacgt ttcagtgcac tacaagagaa gttaatggac ttataTGAat
                                                                <- E1 end
2761 caggtgtaga ggatcttgaa acccaaattc aacattggaa attattaaga caagaacaag
2821 tgttatttta ttatgcaagg agacatggga tattgcgttt ggggtaccaa ccagtaccca
2881 ctctggcaac ttcagagagt aaagcaaaag atgctatagc catgggacta ttgctggaaa
2941 gcttacaaaa atcacaatat gcagaggaac cgtggacctt agtagaaact agtttggaga
3001 cagttaaaag ccctccagca gactgttttta aaaaaggacc taaatctgtg aagtgtact
3061 ttgatggaga tcctgaaaat gtaatgtctt atacagtgtg gtcatacatt tattatcaga
3121 ctgatgatga gtcatgggaa aaggtggaag gtcatgtgga ctatacagga gcttactata
3181 TAGaagggac ctttaaaacc tattatatta aatttgaaac agatgctaaa cgatatggta
E4 orf start ->
NH2 terminus unknown
3241 caacaggaca ttgggaggtg catgttaata aagatactgt gtttaccccct gttaccagtt
3301 ctacgccgcc agttggagtc gcctcccaga actccgcacc cgaaccggca tccacctccg
3361 actccccaca acggtcatca caagtcaccc accgatacgg ccgaaaagca tctagtccta
3421 caatcaccac catcaggagg caaaaaaggc gagagagaca aagacaagaa accccaacaa
3481 ggcgaagaaa aaccagatca aggtcccgaa gcaccgagca gcggggaggg agggccacca
3541 gacgatccct ctccagagaa tccgcagaat cccccaggcg gggagggaga ggtggagggg
3601 gcccctcac caggtcccgc tcaaggtcgc gatcccgtac acgagagtct gttgacgggg
3661 gtggcgtcgc gcctgacgaa gtgggagcaa cacttcgatc aattggtaga cagcatagtg
3721 ggcgacttgc gcaattactg gacgcagcta aagacccccc agTAAttctg ctacgcggtg
                                                   <- E4 end
3781 cagcaaatac attaaaatgc tatcgctata gatttagaaa gaaacatgct ggaagcttcc
3841 aatttattag tacaacgtgg tcctgggtag gggggcatac aaccgataga atcgggcgct
3901 ctaggatact aatatcattt catacagata gggaaagaga gaagtgcttg caacaaatga
3961 aacttccttt aggtgtagaa tggtcatatg ccagtttga tgatttaTAA actgcttttt
                                                   <- E2 end
4021 tactaacaca cTAAcattgc ctatttatac taacctattt gcttgctact aacaaaATGg
        L2 orf start ->                              L2 cds ->
```

```
4081 cgcgagcgcg aagaacaaag cgagcgtcag taactgacat ttataaaggc tgtaaggcct
4141 ctgggacttg tcccccctgat gttattAATA AAgtggaaca aaatacactt gctgataaaa
                         signal ->
4201 ttttaaagta tggcagtgtt ggtgtgtttt ttggtggtct tggtataagt acaggtaagg
4261 gtaccggtgg tcctacaggc tatattcctt taggtcaagg tcctggagtg cgtgtgggcg
4321 ccactcccac agtggtccgc cccggggtca tacctgaaat aattggacca actgaattaa
4381 taccagttga ctcagtaaca ccaattgacc ctgcagcacc atccatagtg acattaacag
4441 acagtagtgc aggtgctgac cttttacctg gtgaagttga aactattgca gaagtacatc
4501 cggtcccaat agacaatgtg gaacttgaca caccttttagt ttctggggac cgtcacgcca
4561 ttttggaggt gactgatgct aatccccctt ttaggcgcac ggttacccgg acacaatatc
4621 ataatcctgc ttttgaaatt atttcagagt ctacaccatt aataggtgaa tctacaccct
4681 ctgaccatgt ttttgttttt gaaggctcgg gaggtgtaca ggtaggggat gctaatgaaa
4741 gcattgaatt ggatactttt ccttctagat atagttttga cattgaggag ccaacccctc
4801 ctcgtagagt tagtacacca attgaaagaa tcagtcagga atttagaact ttaagaagag
4861 ccttatacaa cagaagatta acagaacagg tccaagtaag agacccccttg tttattcgat
4921 ccccgtccag gcttgtgaga tttcaatttg ataatccagt attcgatgag gaagttacac
4981 aaatatttga aagagatgta gctgcagtag aagaaccacc agacagggat ttttttagata
5041 ttgaaagact tggaaggcct atactaacag aaactgcaga aggccgtgtt cgtgtcagca
5101 ggttagggca acgtgcatcg ctgagcacac gcagcggcgc acgtgtaggt gctagagtgc
5161 atttctttac agatattagc actattaatg cagaagagcc cattgaatta gaattattag
5221 gtgagcattc tggcgacagc tctgtagtac aagaaccatt tgaaagcaca atattggatg
5281 tcaatattga caacatacct gaaagtttgg atacaaacat agcagaaaca tctgtagact
5341 atgattctgc tgatttgtta ttagacaacg gtgtggagga ctttagtagg tcacaattgg
5401 taataggtcc ttcagataga tcacttccat ctattactgt tccacaattt gaatcccccta
5461 gagaaaccat tgtgtacata caagacatag agggtaatac agttgtatat cctaaatatg
5521 aagaaaggcc aactattata ttacctacac cctcggggcc tgctaTAAtt caatcaccta
                                                    L1 orf start ->
5581 cacattcctc ctttgactat tatttacatc ctagcttgcg aaggaaaaaa cgcaaacgca
5641 aatatttaTA Atgttttttca gATGaccctc tggcttccaa cttcgggtaa gatatatttg
              <- L2 end
                    L1 cds ->
5701 cctcctacgc caccggtagc ccgagtacaa aacacggacg agtatgtgga gaggactgac
5761 atctattacc atgctataag tgaccgtttt attaactgtag gacatcctta ctttgatgtt
5821 agatcatcag atggagcaaa aatagaggtc cctaaagtgt ctggaaatca gtttagggct
5881 tttagagtaa catttccaga tcctaacaaa tttgctttgg gagatatgac aatccatgat
5941 cccgaaaggt atagattagt atgggcttgt aaagggttag aaataggaag aggacagccc
6001 ttaggtgtag gtaccacagg tcatccatta tttaataaat tacatgatac tgaaaacccct
6061 actgaacgcc aggaaggaac atcagatgat agaagaaatg tttctttttga tcctaaacag
6121 gttcaaatgt ttatcattgg atgtataccg tgtttaggtg aatattggga taaagctcct
6181 gtttgtgaag atgcaggcag tcaggtagga ttatgtcctc cactagaatt aaaaaatggt
6241 gttatagagg atggagatat gtttgatata ggatttggaa atataaataa taaaacacta
6301 tcatttaata gatctgatgt aagcttagac attgtaaatg aaatctgtaa atatcctgat
6361 tttcttacaa tgtcaaatga tgtctatggc gactcatgct ttttttgtgc acgtagggag
6421 caatgttatg cacgacacaa ttttgtacgt ggtggtcttg ttggtgatgc ataccagat
6481 gatgcagttc aacaagatca taaatattac ttgcctgcag cttcacagac tgctttagaa
6541 aactccactt actttccaAC CGTTAGTGGT tctttagtaa cctctgatgc ccaactattc
              E2 binding ->
6601 aacaggcctt tttggttgaa gcgcgcgcag ggccataata atggtatttt gtggaacaac
6661 caaatgtttg taacagtagc tgataatacc cgtaacacta attttttctat tagtgtggca
6721 agtgacggca ccacagttaa ttatgatgct aaaaaaatca gagaatttat gcgccatgtg
6781 gaagaatacc aattatcctt tattttgcag ctatgtagaa taccattaga agcagaggta
6841 ttaactcaaa ttaatgccat gaatcatggc atttttagaaa attggcaact aggctttgta
6901 cctacaccag acaattctgt ccatgatact tataggtatt tacaatctaa agctacaaaa
6961 tgtcctgatg ctgtacctga cacacaaaag gaagatcct ttggtcaata tactttttgg
7021 aatgtagaca tgtctgaaaa gttatcattg gatttagatc agtatccact gggtcgtaaa
7081 ttttttattc aatctgggtt acaacgtgca agggccagtg ccaggtcag tgtgaaacgt
7141 tctgctacgc ggaaaacgtc taaaactgta aaacgaagga aacttacctc tTAACCGTTT
                                                    <- L1 end
                    E2 binding ->
```

```
7201 TCGGTtgctt tAATAAAatc tattaactaa tctggtatgt gaagcatttt ttgaccacct
              signal ->
7261 ttgtgactaa accgaacaag tcaacaccag caACCGCACC CGGTttttac attataaatt
                                E2 binding ->
7321 cctcgaggta agataaccat cagtagatac catcggcacc tggagcaa
//
```

```
LOCUS       HPV23        7324 bp    DNA              VRL        18-JUL-1995
DEFINITION  Human papillomavirus type 23, complete genome.
ACCESSION   U31781
KEYWORDS    .
SOURCE      Human papillomavirus type 23.
REFERENCE   1  (bases 1 to 7324)
  AUTHORS   Delius,H.
  JOURNAL   Unpublished, Sequenced by Hajo Delius, Deutsches
            Deutsches Krebsforschungzentrum, Angewandte Tumorvirologie,
            I.N.F. 506, W-6900 Heidelberg, Germany
REFERENCE   2  (sites)
  AUTHORS   Kremsdorf,D., Favre,M., Jablonska,S., Obalek,S., Rueda,L.A.,
            Lutzner,M.A., Blanchet-Bardon,C., Van Voorst Vader,P.C. and Orth,G.
  TITLE     Molecular cloning and characterization of the genomes of nine newly
            recognized human papillomavirus types associated with
            epidermodysplasia verruciformis
  JOURNAL   Journal of Virology 52 (3), 1013-1018 (1984)
REFERENCE   3  (bases 1 to 7324)
  AUTHORS   Farmer,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     HPV23 was originally isolated from macules on the forearms of a
            Polish epidermodysplasia verruciformis (EV) patient [2]. The HPV23
            genome, like that of HPVs 9, 15, 17a/b, 22, 37, 38, is smaller than
            most PV genomes at approximately 7.4 kb.  Phylogenetic
            reconstructions based on DNA sequences of established types
            indicate that HPV23 is most closely related to HPVs 22 and 38, and
            then to 15, 17, 37 and 9. Strong hybridization was observed between
            HPVs 22 and 23.
BASE COUNT     2331 a   1328 c   1592 g   2073 t
ORIGIN      200 bp upstream from beginning of E6 cds
        1 agcagatacc atcagcacct ggagcgaccg ccaagacttc gccaacttgg cagaacattt
       61 gttggcaaga aaagagcACC GATAACGGTa agaacttta tttttttgacc gtaggcgttc
                       E2 binding ->
      121 atttactaac cttggcaaca attgtggtta acaacaatca taagccaata atacatgcaA
                                                                E2 binding ->
      181 CCGCTTGTGG TAAtttattA TGcagactgt gcattattta agtaggatgt gctacaccaa
            E6 orf start ->
                         E6 cds ->
      241 attattgatg gactcgacgc gACCACTGAC GGTacagcaa cttagtgata agttgacagt
                         E2 binding ->
      301 accagtggta gatctcttgc taccttgcag attttgttct aggtttctta cctatttaga
      361 gttgcgagaa tttgattata aacatttgca gttaatctgg acagaagaag attttgtatt
      421 tgcatgctgc agtggctgtg cttatgcttc tgctcaattt gaaattcaac aattttatca
      481 gctaactgtg tatggtcgtg aaattgagca ggaggagcaa cgacctatag gccaaatttg
      541 tattaggtgt cagtattgtt tgaagtctct cgatttgata gaaaagctag atatctgtag
      601 ttttaatcaa ccatttcaca aggttagaaa tcattggaag ggaaggtgca ggcattgTAA
                                                                E7 orf start ->
      661 ggaaatagaA TGAttgggaa acaagctact cttcgtgata tagttcttga agagcttgtc
            E7 cds -> <- E6 end
      721 cagcccattg acctgcattg ccacgaggag ctcactgaag aggtagaaga agcagtcgta
      781 gaggaggagc ctgaatacac tccttacaag atcatcgtag tttgtggagg ctgtgagaca
      841 cagttaaagc tttacgtgct agccacagat tttggaattc gctcgttcca agcatctttg
      901 ctagaaaacg TGAagctggt gtgtcctgcc tgtcgagaag acattcgcaA TGGacgacga
            E1 orf start ->                                E1 cds ->
      961 TAAaggtact gatactgcta aagaaggctg tagtacttgg tgcttattag aggctgcttg
            <- E7 end
```

```
1021 ttctgatgat agtgacctag atgatagttt ggagaaatta tttgaagaga atgcagagtc
1081 agatgtgtct gatttaataa atgatgatga taatgctgct cagggaaatt cccgcgaatt
1141 gctatgtcaa caggagagtg aggaatgcga gcagcaaata caatacctaa aacgaaagta
1201 taatatcagt ccagaggctg ttcagcagct tagtccacgt ctacagtctt tgaatttgtc
1261 gcctgggcat aaatctaaaa ggagattgtt tgtggagcaa gacagcggac tggagttatc
1321 tctaaatgaa gttgaagatt ttactcaaga gttggaggta ccggcgagcg ctccagggcc
1381 ggcagcccag ggtggagtag ggctgggaca tattgaaagt ttgttaagat gtaaaaatgc
1441 taaagcagtg ttgctacata aatttaagga aggttttgga attagttata atgagcttac
1501 cagacagttt aaaagcaata agacctgctg taaacattgg gtattggcca tatatggtgc
1561 aaaagaagag ctcatagatg cgtctaagca attgttacaa cagcactgtt cttatatttg
1621 gttgcagaca tacacaccta tgtcacttta tttatgttgc tttaatgttg caaaaagtag
1681 agaaacagtt gtaaaattat tgatttctat gctgcaaata catgaaaatc atatattatc
1741 agaacctccg aaaaacagaa gtgtacctgt agctttattt tggtataaag gcagtatgaa
1801 ccctaatgta tatgcatttg gtgagtatcc tgagtggatt gtgacacaaa ccatgataca
1861 acatcaaact gctgacagta tacaatttga tttgtctcgt atgattcaat gggcctacga
1921 taatgatcat cttgacgaat gtagtattgc ttataactat gcaaaattgg ctgacacaga
1981 cagcaatgca agagcttttt tagctcaaaa tagccaagca aaacatgtaa gagattgtgc
2041 acagatggtt aagcattata aaagaggtga aatgcgagaa atgactattt ctgcatgggt
2101 acatcattgc atatctagaa ttgaaggtga tggacaatgg caagatattg ttaaattttt
2161 gcgctatcag ggattaaact tcattgtatt tttagataaa tttagaacgt ttttacagaa
2221 ttttccaaaa aaaaattgtt tgttaatata tgggcctcca gacacaggca aatcaatgtt
2281 tactatgtct ttaatgaaag cactaagagg tcaagtaata tcgtttgcaa attctaaaag
2341 ccaattttgg ctgcaaccat tagctgatgc aaagatcgcc ttattagatg atgcaacaga
2401 agtttgttgg caatatattg atatgtttct tcgaaatgga ttggatggta atgtagtgtc
2461 gttggatatg aaacatagag caccatgtca aatgaaattt ccaccattaa ttattacatc
2521 taatattagc cttaagaaag aaaagaagtt tccttacttg catagtagaa tatatgaatt
2581 tgaatttcca aacagatttc catttgattc agatgataaa cctttgtttta aacttactga
2641 ccaaagctgg gcgtcttttt ttaaaaggct ttggatacaa tTAGgactca gtgaccaaga
                              E2 orf start ->
2701 ggacgaggga gaggATGgaa gcactcagcg aacgtttcag tgcactacaa gacaagttaa
              E2 cds ->
2761 tggacctgta TGAatcaggt ttagaggatc ttgaaactca aatacagcat ggaaactct
                     <- E1 end
2821 taagacaaga acaaatttta ttgtattatg ctcgaaaacg tggaattatg cgtttggggt
2881 accagccggt acctcctctg gcaacatcag aaattaaagc aaaagatgct atagcaattg
2941 gaatttttgct ggaaagttta caaaaatcca aatatgcaga tgagccatgg acattagttg
3001 agactagctt ggagacaatt agaagtccac cagTAGattg ctttaaaaag ggacctaaaa
                              E4 orf start ->
                              NH2 terminus unknown
3061 cagtggaggt gtattttgat ggagatcctg aaaatgttat gccatataca gtatggtctt
3121 atatttacta tcaaactgat gaggacactt gggaaaaggt tgaaggacat gtggattata
3181 caggagctta tttttatgag ggccaactta aaaactatta cattaaattt gaagcagatg
3241 caaagcgctt tggtactaca ggaatgtggg aagtacatgt aataaagat actgtcttta
3301 cccctgttac tagttctacg ccgccagttg gagacgcctc caacaacgcc gttcccgaag
3361 catctaccac ctccttgtcc tccccacaac ggtcaccatc caccaaccgc cgatacggcc
3421 gaaaagcatc tagccctaca gccaccacca ggaggcaaaa aagacaagga aaagaaaccc
3481 tcaccaggcg aagaaaaacc agatcaaggt cccggagcag agagcaacgg gggggagggg
3541 aaacccaaag atcctcctcc agaggagcct caaaatcccc ccggcgggga gggagaagtg
3601 gaggggggcc cctcacccgc tccagatcaa gatccagatc accagagtct gttacagggg
3661 gtggcgttgc acctagtgaa gtgggagcgt cacttcgatc agttagtaga cacagtagtg
3721 gaagacttgc gcaactattg gatgcagcta aagacccccc agTAAtattg ctgcgcggcg
                                                  <- E4 end
3781 gtgcaaatac attaaaatgc tatcgctata ggtttagaaa aaagcatgct ggtaaatttt
3841 attatgttag cacaacgtgg tcatggattg ggggtcattc tactgataga gtagggcgtg
3901 caaggatgtt aatagcattt cattctaatc atgaaaggga aaaatgtatt caagaaatga
3961 agttaccttt aggagtagat tggtcctatg gacaatttga tgatttaTAA cctgctttt
                                                  <- E2 end
4021 attTAAcaca ctaacattgc ctattgctat tttttactta acttatattg ctatattgct
L2 orf start ->
```

```
4081 actaacattA TGgtacgggc gcaaagaact aagcgagcgt ctgttactga tatatacaaa
        L2 cds ->
4141 ggctgtaaag cctctgggac ttgtcccct gatgtactaA ATAAgtgga acaaaataca
                                       signal ->
4201 cttgctgata aaatacttaa atatggcagt gttggtgtgt tttttggtgg acttggaatt
4261 ggtacaggta agggtaccgg tggtgccacg gggtacgtcc cattgcgacc tggagtacga
4321 gtgggcggta ctcctacagt ggtccgccct gcagtcatac ctgaaataat tggaccaact
4381 gaattaatac cagttgactc aatagcacca attgaccccg aagcaccatc aatagtctca
4441 ttaacagaca gtggcgcagc tgctgacctt ttccccagtg aagcagaaac tattgcagag
4501 gtacatccta cacctgtaga cataggaatt gatacaccta ttgtagctgg aggccgtgac
4561 gccatttag aggtggtaga tactaatcct ccaacaaggt tcagtgtaac aagaacacaa
4621 tatgataatc catcttttca aataatttca gaatccacac ctatcacagg tgaggcatcc
4681 cttgctgatc atgtatttgt gtttgaaggt tctggaggtc agcacgtagg agcggtaact
4741 gaagagattg aattagatac atatccttcc agatattcct ttgaaattga ggaagctaca
4801 ccaccacgca gaactagtac tcccattgaa agaataagtc aggaattcag gaacctacgt
4861 agagcactgt ataacaggcg cttaacagaa caggttcaag taaaaaaccc tttattttta
4921 actactccat ctaaacttgt aagatttcaa tttgataatc ctgtgtttga tgaagaggtc
4981 acacaaatat ttgaaagaga tgttgctgaa gtggaggaac ctccagatag ggacttttta
5041 gatatagaca gattaggaag accattatta acagaatcca ctgaaggccg tattagatta
5101 agtaggttag gtcaaagggc ttccattcaa acacgcagtg aacacgtgt tggttcacgt
5161 gtacacttct atacagattt aagcactatt aatacagaag aacctataga attagaatta
5221 ttaggcgagc attctggaga tgcatcagtt attgaggaac ctctgcaaag cactgtaata
5281 gatatgaact tagatgatgt tgaggctatt caggatacta tagatactgc agatgattat
5341 aactctgcag atctttatt ggacaatgca attgaagaat ttaataattc tcaattagtg
5401 tttggcactt ctgatagatc ttcgtctgca tattctatac acggttga atcccctaga
5461 gaaacaattg tatatgttca agatatagaa ggtaatcagg taatttatcc tgggcccaca
5521 gaaaggccaa ctataatatt tccctacct agtgccctg ctgtagtcat acacacattg
5581 gacaagtctt ttgattatta cttacatccc agctTAAgaa agaaaaggcg caaacgcaaa
                                     L1 orf start ->
5641 tatttaTAAt gttttcagA TGaccctctg cttccagct tctggtaaga tatatttacc
            <- L2 end
                L1 cds ->
5701 tcctacgcca cctgtagccc gagtgcagag tacggatgaa tatgtggaaa gaactgacat
5761 ctattaccat gcaactagtg atcgattact aactgtaggc cacccatatt ttgatgttag
5821 atcaccggat ggtagtaaaa tagatgtacc aaaggtttca gggaatcaat tcagggcctt
5881 tagagttaca tttccagacc ctaataagtt tgcattagca gacatgacta tctatgatcc
5941 tgataaatac aggttggtgt gggcctgcgc aggacttgaa atcggccgcg gccaACCTTT
                                                                E2 binding ->
6001 AGGGGTcggc agtacaggac acccgctatt aataagctc cgtgatgcag aaaattctag
6061 tgaacgtcag gaaggtactg tagatgacag aagaaatatc tcatttgatc ctaagcaagt
6121 acagatgttt ataattggtt gcacaccgtg cttaggtgaa tattgggata cagctcctgt
6181 ctgtaaagat gcaggtagcc aactagggtt gtgtcctcct ttagaattaa aaaacagtgt
6241 tatagaagat ggggacatgt cgacattgg ctttggtaat atcaataata aaacattatc
6301 ctttaataga tcagatgtta gtttagatct tgtaaatgag gtttgcaaat atccagactt
6361 tttgactatg tcaaatgatg tatatggaga tgcctgtttt ttttgtgccc gaagagagca
6421 atgctatgcc aggcactatt ttgttcgagg cggtgtagta ggagatgcaa tacctgatgg
6481 tgcagttcaa caggatcaca atattattt acctgcagac caacaaaaca ctttagaaaa
6541 ctcactttat tttcctactg tcagtggatc tttggtaact tctgattctc aactttttaa
6601 tagaccattt tggttaaaac gtgctcaagg ccataacaat ggtatttat ggaacaacca
6661 gatgtttgtg actgtagcag ataatacacg taatacaaac tttagtatca gtgttaccaa
6721 tgacagcagt ttagaaaagt atgatgccac taaaattaga gagtttacaa gacatgttga
6781 agaataccaa ctttcttta tactacagtt gtgcaggata ccttaaaagg ccgaggtctt
6841 aacacaaatt aatgccatga attcagatat tttagagaat tggcagttag ggtttgttcc
6901 tacaccagat aatgcagttc atgacacata cagatatttg gcttcaaagg ccacaaaatg
6961 tccagatgca gtacctgaca cgcaaaaaga ggatccttt ggaaagtatt cattttggaa
7021 tgttgatatg acagaaaaat tgtctctaga cctagatcaa tatcccttag gccgtaagtt
7081 tctgtttcaa attggagtgc agcgtgtacg gtccggtacc aaacggcctg caactcgaaa
```

```
7141 agtgaccaaa actgtcaaaa ggaaaaaagt gcaattgTAA CCGATATCGG TcgccAATAA
                                            <- L1 end   signal ->
                              E2 binding ->
7201 Aatatgttaa ctaatctggt atgtgaagta tttttttaacc gtctttgtga ctaaaccgaa
7261 caagtcaaca ccagcaaccg cacccgtttc cacattataa attcctcgag gtaagattat
7321 gatc
```
//

```
LOCUS       HPV24         7452 bp    DNA            VRL        18-JUL-1995
DEFINITION  Human papillomavirus type 24, complete genome.
ACCESSION   U31782
KEYWORDS    .
SOURCE      Human papillomavirus type 24.
REFERENCE   1  (bases 1 to 7452)
  AUTHORS   Delius,H.
  JOURNAL   Unpublished, Sequenced by Hajo Delius, Deutsches
            Deutsches Krebsforschungzentrum, Angewandte Tumorvirologie,
            I.N.F. 506, W-6900 Heidelberg, Germany
REFERENCE   2  (sites)
  AUTHORS   Kremsdorf,D., Favre,M., Jablonska,S., Obalek,S., Rueda,L.A.,
            Lutzner,M.A., Blanchet-Bardon,C., Van Voorst Vader,P.C. and Orth,G.
  TITLE     Molecular cloning and characterization of the genomes of nine newly
            recognized human papillomavirus types associated with
            epidermodysplasia verruciformis
  JOURNAL   Journal of Virology 52 (3), 1013-1018 (1984)
REFERENCE   3  (bases 1 to 7452)
  AUTHORS   Farmer,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     HPV24 was originally isolated from macules on the breast of a Dutch
            epidermodysplasia verruciformis (EV) patient [2]. It hybridizes only
            weakly with other other EV-related HPVs [2]; phylogenetic analysis
            of DNA sequences indicates that HPV24 is clearly most closely
            related to EV-related types, but it is not closely related to any
            particular subgroup.
BASE COUNT     2385 a   1352 c   1629 g   2086 t
ORIGIN      200 bp upstream from beginning of E6 cds
        1 gcgccaagtt tcttggcact cgcacagcAC CGTTAACGGT aagtgttttg tatactgtac
                                   E2 binding ->
       61 cgggcgcgat accaaatatg ccacctgcgc aattgttgtt agcaactact ttcattacct
      121 aacaagtatt ttcatgcacc gtttccgtat cactgtgaat gtatttattg atTAGacaTA
                                                        E6 orf start ->
                                                                signal ->
      181 TAAAAaggaa gacatcttgA TGgctcaacc aggtaaacct cagtcagtgt tagaacttag
                    E6 cds ->
      241 tagattatta aatataccat tagacgattg tgttgtacca tgtaactttt gcaaaagatt
      301 tctaagttat acagagttaa ctgactttga caccaaatgc ttaagtttga tttggaaaga
      361 cgattttgtg tttgcatgtt gtagatattg ttgtgttgct acagcagcat ttgaatttga
      421 aaattatttt gtagagagtg tgataggttg ggaaatagaa caaaaagaaa atacacctct
      481 ttcggacatt attgtaaggt gtcaccattg cttaaaatta cttaatcaaa ttgaaaagct
      541 tgatatctgt ggaagatctg aattatttca taagtgagg aggggctgga aaggactctg
      601 taggcagtgT AAgcagataT AATGattgga aaagaggtca ctctacaaga cattgtctta
          E7 orf start ->       <- E6 end
                        E7 cds ->
      661 gagttgactg agccacagac tgttgatttg cactgtgaag aggagttgcc agagcaggat
      721 acggaggtgg agcctgaaag aagagcttac aaaataatac tttgttgcgg cggcggttgt
      781 ggaacccggc ttcgattatt tgtagcagca acacagtttg gaatacgtgg cctgcaagac
      841 ttactgcTAG aagaggtggt catttgtgt cccgactgcc gtaacagcga tctgcagcAT
        E1 orf start ->                                           E1 cds ->
      901 GgcggacaaT AAaggtagtg aaacagacgg gttgcatgaa tggtgcttat tagaagctga
                    <- E7 end
      961 ctgtagtgat attgaaaatg atttggacac attgcttgag caaaattcag attctgatgt
     1021 atcagatctt ataagtaatg atggggactt ggaacaggga aactcccgag aactgtttca
     1081 acagcaggag ttagaggaga gcaatgcttt gttgcaaagt ctaaaacgaa agtatattag
     1141 tcctaaagct gtattacagc tgagtccaca acttgaatcc atttcgctgt catcagacca
```

```
1201 taaaactaag aggaaattgt ttgctgagca ggacagtgga gtagaattga cccttacaaa
1261 tgaaactgaa gatgttacta cattggtgca gcaggaggaa gaggtaccag ctttagagac
1321 atcttccacc agtaacttaa ggaaggaaga taatgcacat tataaagaac ttatgcgatg
1381 cagtaactta aaagctacct tactatcaaa atttaaaaat gcttttggtg taagctttgt
1441 tgaactgact cgccagttca gaagtaataa gacttgttgt aatgattggg tggtagcaat
1501 atatggtgta aattatgatt tatttgaaag ttctaagcaa ttactacaac agcattgtga
1561 ttatatttgg gttacagaaa tgtctgcaat gttttttatat ctgttgtgtt tcaaggctgg
1621 caaaaataga caaactgtta taaggttatt agtgtctatg ttatatgtgg cagaggagca
1681 aattttatca gaacctccaa aattgcgaag tacagtgtca gctttatttt ggtataaagg
1741 tagctcaaat gctgctacct ttacacatgg gtcttatcct aaatggatta tagaacaaac
1801 actgatagga catcaaacag gagaagctgc aacgtttgac atgtccacaa tggtacaatg
1861 ggcttttgat aatgatctca cagaggaagc tgacatagct tttcaatatg caaagctggc
1921 acctgatgat gtaaatgcaa ctgcttggtt ggctcacaat aatcaggcgc gttttgtaag
1981 agaatgtgct aatatggtta gatattataa aaaaggtcag atgcgagaaa tgtctatgtc
2041 ggcatggatc cactttaaat tagagcaagt agaaggggaa gggcagtggt caactattgt
2101 aaaatttatt aggtatcagg gtataaattt tattagtttt ttaagtgctt taaaagattt
2161 tttacatggt aaacctaaaa agaattgttt attaatatat ggccctccaa atactggaaa
2221 atcagcattt actatgtcat taattaaagt gttgcatggc agagtaatat catttgtaaa
2281 ttctaaaagt cacttttggt tgcaaccaat gtctgaagcc aaaatagctt tacttgatga
2341 tgctactgat ccatgttgga tatatatgga tacatattta agaaatggtc ttgatggaca
2401 cttagtttca ttagattgta agcataaagc acctatacaa atacgctttc ctccattact
2461 tataacatct aatattaatg caatggcaga acctaactat agatatttac acagtaggtt
2521 agtagcgttt gaatttccta atcccttttcc tatgaaagat gacgataccc ctgaatttga
2581 acttactgac caaagctgga aatctttttt taaaaggctt tggagacaat TAGacctcag
                                             E2 orf start ->
2641 tgaccaagaa gacgagggag aggATGgaga acctgaaaaa gcgtttcgat gtgttgcaag
              E2 cds ->
2701 atctactaat gaacatttaT GAacaaggca gtgatacact agaatcacaa attgaacatt
                        <- E1 end
2761 ggcaggcctt gcgaagagag gcagtgttac tttattatgc tagacaaaat ggggtgctgc
2821 ggctgggtta cttaccagtt ccaccactag ctacctcaga agccaaagct aaacaggcta
2881 ttagtatggt gttgcagctg caatctttac aacaatctcc ttatggcaca gaaaagtgga
2941 cactggtgga cacaagcata gaaaccttta aaatactcc agagaaccat tttaaaaaag
3001 gacctataaa tgtggaggtt atatatgatg gtgatccgga caatgccaat ttgtatacta
3061 tgtggaaata tgtatattat atggatgata atgaccagtg gcaaaaaact gaaagtggcg
3121 ccaatcatac aggcatatat tatttaaTAG gggagtttaa acattattat gtgctgtttg
              E4 orf start ->
              NH₂ terminus unknown
3181 ctgatgatgc caatagatac agtaaatctg gccaatggga ggtcaggatt aataaggaaa
3241 ctgtgtttgc ccctgtcacc agctccacac cacccgactc cccaggaggg tccccgagaac
3301 tacccggatc caccgctaac tccaaggcct caagcccaac ccaacagcca caacaagcct
3361 gtagtgacga acaaccaag cggaagaggt acgggcgaag ggagtcaagc cccactgact
3421 ccagatgcag acgacgatcc tcgtccccggc aaaagaagca agggcgacga gcacggtccc
3481 gcacccggtc gcgctgcagc tccactcaaa ctcgatctag atccacctca aggaggtcca
3541 gatcaacctc caggggcaac aggaggtgta ggggagacac ccccagaggg caacgaggag
3601 tctcaacctc ctccagggga aggggaaggg gcagtagaag gtcatcctcc tcctcctccc
3661 ccaccccccag aaccaaagcc tcacaacggg gatgcgacac acggtctgtt agggacagtg
3721 gcatctctcc tggggacgtg ggaagaaagc ttcagacagt tagtggaaga aattcaggaa
3781 gacttggacg attattggag gaggctctcg atcccccagT AAttttactt cgtgggggtg
                                             <- E4 end
3841 ctaacacatt aaaatgtttt cgcaacaggg caaagcttag atatagagga cattataaag
3901 catttagcac atcttggtca tgggtagctg cagatggcac agagcgtcta ggcaggtcca
3961 gattgctcgt cagttttacc agttttaagc agcgaagtgg gtttctcgac ctagtaagat
4021 ttcctaaggg tgtagattgg tcactgggaa gttttgataa actgTAAgta ctaacatagc
                                             <- E2 end
4081 ttttgctacT AAcacacata taacctattt ttactttttg tatgctttgc aATGgtgcgt
     L2 orf start ->                                          L2 cds ->
4141 gctaaaagaa caaaacgaga ttctgctact aatatttaca gaacatgcaa gcaggcgggc
```

```
4201 acatgtccac ctgatgttat tAATAAAgtg gaacagtcaa ccattgctga taatatactt
                        signal ->
4261 aaatatggaa gtgctggtgt cttttttggt ggcctgggta ttagtacagg ccgagggact
4321 ggaggcacca caggatatgt gccattgggt gaaggtactg gagttcgtgt gggcagtaca
4381 cccacagtcg ttcggcctgc ccttgtgcct gaagtaatag gtcctgctga cttattacct
4441 gttgatacaa tagcccctgt tgacccagca tcgtcatcta tagttcctct tactgaatca
4501 tcaggtgtag accttttacc tggtgaaata gaaacaatag cagaagtaca tcctatacct
4561 gatgtgccca catttgatac tccagtagtg acaacaagca aaggctctag tgccatttta
4621 gaagtagctc ctgagcctac tccaccgaca cgtgttcgtg tcagcagaac acaataccat
4681 aatccagcat ttcatattat tacagaatct acaccaagtc aaggtgagag ttcattgtca
4741 gacgaaataa ttgtcgcctc tggtgcaggt ggacaatcag taggcgtttc tgaaaatata
4801 gaactgcagg atttatcaaa tagatattct tttgaaatag aaacaccaac accaccaagg
4861 cgtagcagta ccccattaca aagagctaca caagcattta gacaaagatc acttacaaat
4921 aggagactgt tacaacaagt gcctgttgaa gaccctttgt ttttaactca accatctaaa
4981 ttggttagat ttgcatttga gaatccagct tttgaagagg aagtaacaca ggtatttgaa
5041 caagaccttg caggtttttgt agagcctcct aacagagatt ttttagatat tgcagaactg
5101 ggaaggccta gattttctga aacacgcgag ggttatgtta ggttaagcag attgggtcgt
5161 agagcaacta ttagaacaag ggcaggaaca caaataggag cacaagtaca ttttttataaa
5221 gatctaagtt ctattaatac tgaagctcct attgagttgg atcttttagg gcagcattct
5281 ggggatgcaa ctatagttca tgggactgta gaaagcacat ttatagatac taatatagag
5341 gaaaatcctt tagctgaaca aatggagttg gaaattgata cttatcctga agctcattca
5401 tttgatgctt tgttagatga agcaacagac gatttttagtg gttcacagtt agttataggc
5461 aatagaagat ccactacatc atatactgtt cctagatttg aatccccaag aaattcttct
5521 tattatgtac aggatttgca gggatattat gtagcctatc ctgaatctcg cgataaaata
5581 gaacttattt atccctcacc cacattacct gcagttgtca tacatacaga agatagtagt
5641 ggggacttttt atttacatcc tagtttattg caaaggcgca gacgcaaacg aaaatatttg
5701 TGAtttttac agATGtcggt gtggttgcca gccagtggta aggtatattt gccaccatca
            <- L2 end
L1 orf start -> L1 cds ->
5761 acacctgttg cgaaagttca aagcacggat gaatacatac agagaacaaa catcttctat
5821 catgcttata gtgaccgcct attaactgta ggacacccat atttcaatgt ttacaacaat
5881 gatggcacag tattagaggt ccctaaggtt tcaggaaatc aacacagagt ttttttaggctt
5941 aaattaccag accctaatag atttgcttta gcagatatgt ccgtatataa tccagagaag
6001 gaaagattgg tatgggggttg cagaggagta gaaataggta gaggacaacc attaggtgtt
6061 gggacaagtg gacatccatt atttaacaaa gtgaatgaca cagaaaaccc tgtatcatat
6121 aggacacaag catcgtccac agatgataga caaaatacct catttgatcc taaacaaatt
6181 caaatgtttta ttataggttg tgcaccctgc ataggagaac attgggaagt agctgagagg
6241 tgtgctggtg ataataatga tgctggtaga tgtccaccta ttaagttagt aaattcagta
6301 attcaagatg gtgatatggc agatattggt tatgggaatt aaatttttag aacactacaa
6361 caaagtagat cagacgtaag tttggatatt gtaaatgaaa cctgtaaata tccagacttt
6421 ttaaaaatgc aaaatgatgt ttatgggagat tcttgttttt tctttgctcg ccgtgagcaa
6481 tgttatgcaa gacattttttt tgtacgagga ggtaaaccag gggatgacat acctggtgag
6541 caaattgatg cagggaccta caaaaatgac ttttacattc cagcagccac aggtcagaca
6601 caaaaaaata ttggcaactc aatgtatttc ccaacagtca gtggctcatt ggtatctagt
6661 gatgctcaat tatttaatag accattttgg ttacagcgtg cacaaggtca taataatgga
6721 atttgttggg ctaatcaatt gttcattacg gtagtagata caccaggaa caccaatttt
6781 agtataagtg tatacactga aaatggcaaa gtaacagata ttaacgagta tgatgctaat
6841 aaatttaggg aatatcagag acatgtagaa gagtatgaaa tttcacttat actgcagcta
6901 tgtaaaatcc ctttaaaagc agatgtgtta gcacagatca atgctatgaa tccatctcta
6961 ttggaagagt ggcagttagg atttgtgcct gcaccagaca atcctcttca aagtacctac
7021 agatatatag aaagtttggc aacaccttgt ccagataaag tagctccaaa agaaagagaa
7081 gatccttatg cccccttatac gttttggaat gtagatttat ctgaacgctt atctttggaa
7141 ctggatcaat attccttagg acgaaagttc ctgtttcaag cgggtcttgt acaaaaaaca
7201 tctaaaaaaa catctaatgt atccaagggg accaaacgaa aacgcacgTA AAATAAAggc
                                                        <- L1 end
                                                    signal ->
7261 tgtttACCGT TTTCGGTaca atatatgtgt attccaagaa tgcttgttat tcatgtgcgt
    E2 binding ->
```

```
7321 gactaatttg aactttggct tgaatgtcaA CCGCACCCGG Tacaagtaga taaaatcttg
                             E2 binding ->
7381 ctaccacagc agacacaaag gaatcactgt cggtctcaac acgctcggat ttggcgcatt
7441 caaccgtttt tg
//
```

```
LOCUS       HPV36          7722 bp    DNA              VRL       18-JUL-1995
DEFINITION  Human papillomavirus type 36, complete genome.
ACCESSION   U31785
KEYWORDS    .
SOURCE      Human papillomavirus type 36.
REFERENCE   1  (bases 1 to 7722)
  AUTHORS   Delius,H.
  JOURNAL   Unpublished, Sequenced by Hajo Delius, Deutsches
            Deutsches Krebsforschungzentrum, Angewandte Tumorvirologie,
            I.N.F. 506, W-6900 Heidelberg, Germany
REFERENCE   2  (sites)
  AUTHORS   Kawashima,M., Favre,M., Jablonska,S., Obalek,S. and Orth,G.
  TITLE     Characterization of a new type of human papillomavirus (HPV)
            related to HPV5 from a case of actinic keratosis
  JOURNAL   Virology 154 (2), 389-394 (1986)
  MEDLINE   87020745
REFERENCE   3  (bases 1 to 7722)
  AUTHORS   Farmer,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     DNA of HPV36 was extracted from biopsied lesions of 2 patients with
            actinic keratosis (AK) and molecularly cloned; 22 other patients
            with AK were HPV-negative by blot hybridization to a range of
            HPV-types. The HPV36-positive patients were an 89 year old man
            (patient A) and a 61 year old man (patient B) with one lesion on
            the head and multiple lesions on the backs of hands and fingers
            respectively. Patient B also had a squamous carcinoma on the
            forehead. In patient B, HPV DNA was detected in only two of seven
            lesions of the hand and not from the sqamous carcinoma. HPV36 DNA
            was not detected in many other skin lesions in the general
            population (22 samples of Bowen's disease, 12 squamous cell
            carcinomas, 21 basal cell carcinomas, 8 keratoacanthomas). In
            contrast, HPV36 was found in 7 of 18 epidermodysplasia
            verruciformis (EV) patients known to be positive for HPV5 and at
            least one additional EV-associated type. The low rate of detection
            of HPV DNA in AK lesions in the general population allows for the
            possibility that HPV36 was not a causative factor in those AK
            lesions where it was detected. Nevertheless, HPV36 was the first
            example of an EV-associated type to be detected in an
            immunocompetent non-EV patient. Blot hybridization unde stringent
            conditions revealed significant cross-hybridization only with
            certain EV-associated types (HPV 5,8,12,14,19-23,25); no
            cross-hybridization was detected with the other EV-specific HPVs
            (9,15,17,24), nor various non-EV cutaneous types (1-4, 7, 10,
            26-29, 34), nor variousmucocutaneous and mucousal types (6, 11, 13,
            16, 18, 30-33,35). Phylogenetic analysis of DNA sequence indicates
            that HPV36 is most closely related to HPVs 5, 47, 8 and 12.
BASE COUNT    2337 a   1497 c   1773 g   2115 t
ORIGIN      200 bp upstream from beginning of E6 cds
        1 gtaagttatt taatttatgt ACCAGGTGCG GTactggaat ttcacaataa taattgttgt
                  E2 binding ->
       61 tgccaactac cattgctata ttcaagtttt tgcctgtatc gttttcgtat catgtgaata
      121 atatactgta tagTATAAAt aaataaataa atatatatat atatgcttca aaggttgggt
                  signal ->
            dinucleotide "TA" repeat region ->
      181 tttttaaTAA ttaaggcaaA TGgcagagca agcctccgaa cagcagaata ttacagaaaa
       E6 orf start -> E6 cds ->
```

```
 241 agaaaaagaa cagctgcctt taactattaa gggcctgtca gaatcattag gcattccgtt
 301 tgtagactgt ctaatacctt gtaacttttg tggcaaattt ttagattatt tagaagcttg
 361 tgagtttgag gttaaaaagc ttagtttaat ttggaaggat tactgtgtat ttgcctgctg
 421 tcgtgtttgt tgtggtgcaa cagccacata tgagtttaat cagttttatc agcagacagt
 481 attaggaaga gatattgaat tggctgcagg tcgctccata tttgaaattg acattaggtg
 541 tcagacgtgc ttagcttttc ttgacataat tgagaagtA Gattgttgtg gaagaggact
                                  E7 orf start ->
 601 tcccttttcat agagtcagaa acgcctggaa gggaatctgt aggcagtgta agcattttta
 661 taATGattgg TAAagaggtc accgtgcaag atattgttct ggagctcagt gaggtgcagc
E7 cds ->          <- E6 end
 721 ctgaagtact accagttgac ctgtttttgtg aagaggaatt accaaacgag gatacggagg
 781 aggagcttga caccgaaaga atcgtttttca aagtcattgc accgtgtggt tgcagccact
 841 gtgaggtcaa gctccgcgtt tttgtccaag ctacagaatt tggcatcaga gcatttcaac
 901 agttgcTGAc cggtgacctg cagcttctgt gtcccgagtg tcgtgggaac tgcgaacATG
E1 orf start ->                                                E1 cds ->
 961 gcggatccTA Aaggtagtac atctaaagaa gggtttggtg attggtgtat tttggaagct
               <- E7 end
1021 gactgtagtg atatagaaaa tgatatggaa caattatttg aaagagatac agattctgat
1081 atttcggact taatagatga ttgtgacctg aacagggaa attctttgga actatttcac
1141 caacaggagt gtaagcagag cgaggagcaa ttacaaaaac taaaacgaaa gtgtcttagt
1201 ccaaaagctg tcgcacagct tagtccgcga cttcagtcaa tttcattgtc acctcagcag
1261 aagtctaagc gaaggctctt tgcagagcag acagcggag tcgagctgac cttaaacaat
1321 gaagctgaag atgttactac tgaggtggag gtaccggcta tagactctcg gccggatgac
1381 gagggaggat cagggatgt agatatacat tatctttcac tgttgcgttc cagcaacaaa
1441 aaagccacat taatggcaaa atttaaagca gcgtttgggg taggctttaa tgaattgaca
1501 cgtcaattca aaagtcacaa aacctgctgt aatcattggg ttgtctctgt ttatgcagtc
1561 catgatgatc tatttgaaag ttcaaagcag ttgttgcaac agcattgtga ctatttatgg
1621 gttcgtggga tagatgcaat gtcattatat ctattgtgtt ttaaggcggg aaaaaatcgc
1681 gggacagtgc ataagttgat gacttcaatg ttaaatgtgc atgaacagca gattttgtct
1741 gagcctccaa agttaagaaa tactgctgct gcattgtttt ggtacaaagg ctgtatggga
1801 tcggggggtgt tcagttatgg gccatatcct gattggattg cacaacagac tatattaggt
1861 cacaataatg ctgaagcaag caccttttgat ttttcacaga tggtacaatg ggcctttgat
1921 aatcagttag ttgatgaagg agatattgca tatcgatatg caaggcttgc accagaagat
1981 gccaatgctg ttgcatggct tgcacataat agtcaagcta aatttgtaag agagtgtgca
2041 gcaatggtgc gctttttataa aaaaggtcaa atgagggata tgtccatgtc ggaatggatc
2101 tatactaaaa tacatgaggt agaaggtgaa ggtcactggt cagatatagt aaaatttttta
2161 agatatcagg aggtaaactt tataatgttt ttggctgcat tcaaagattt ttttgcactca
2221 aagccaaaaa aaaattgtat tttaattcat ggccctccta attcaggaaa gtcatcattt
2281 gcaatgtcat taataagggt attaaaaggc agagtattat cttttgtaaa ttcaaaaagc
2341 caattttggt tgcaacccct ttctgaatgt aaaatagcat tgattgatga tgtaactgAC
                                                            E2 binding ->
2401 CCCTGTTGGT tatatatgga taattatcta agaaacggtt tagatgggca ttatgtctca
2461 ttggattgta aatacaaagc tccaatgcaa acaaagtttc ctccattatt actaacatct
2521 aatataaatg tgcatgagga agctaattac agatacttac acagtagaat taaaggattt
2581 gcatttccaa atccattccc aatgaaatca gacgatacac ctcagtttga gttaactgac
2641 caaagctgga aatctttttt tgaaaggctt tggacacaat TAGaactcag tgaccaagaa
                                  E2 orf start ->
2701 gacgagggcg aaaATGgaga atctcagcga gcgtttcaat gctctgcagg atctgctaat
              E2 cds ->
2761 gaacatttaT GAagctgcag aacagacact tgaggcacag ataaaacact ggcaaacctt
              <- E1 end
2821 gcgacaagaa gctgttttgc tctactttgc taggcagaga ggtgtgacaa ggcttggata
2881 tcaacctgtg cctgtaaaag ctgtatctga agcaaaagct aaagaggcta tagcaatggt
2941 gctgcagctt cagtctctac aaacgtctga atatgcatct gaaacttgga cattagttga
3001 caccagtata gaaactttta gaagcgctcc agacggtcac tttaaaaagg gtccagtgcc
3061 tgtagaagtg atttatgaca atgatgcaga taatgccaat ttgtacacta tgtggacata
3121 cgtgtattat atggaagacg atgtgtggca taaagccaga agtggggtca atgagactgg
3181 catatattat ttacaaggaa catttaaata ttactatgta ctatttgctg acgatgcacg
```

```
3241 taaatatagt caaactggac aatgggaagT GAaagttaat aaggaaactg tctttgcccc
                                E4 orf start ->
                                NH₂ terminus unknown
3301 tgtcaccagc tccacccctc cagggtcgcc aggaggacaa gcagacacaa acgcctcctc
3361 caagacctcc accaccacca cagccaccgt tgactccacg accaagcagc tcaccacatc
3421 agaacagcca caacaaaccg aaaccaaagg aagaaagtac ggacggaggc cctccagcag
3481 gacaaggaga ccgcaagcga agcaaaggcg atcaaggtcc agacaccgat cctctaggtc
3541 ccgatcgcgg tcccagtccc ggtcccacac cccaaccact cggtctgcca ccacccggtc
3601 taggtccccg tcgctcgcca agactggggt ccagcgggta tcaaccagat cacgatccag
3661 aagcacctct agaaggggag gtagaaggcg gaggtcacgg tcaccatcca cctcctcctc
3721 caccaccacc accaacaaac ggtcacgagt gcgggccgaa accacagggt ccagaggggc
3781 gcgaggggt agaggggcca ggggtgggag cggtgggggg cggcgacgag gacgatcatc
3841 ctcctccacc tcccccgccc acaaacggtc aagagagcac tctgttagga gccgtggcgt
3901 gtctcctgac caagtgggaa agtcacttcg atctgttagt tcaaaacata caggacgact
3961 tggaagatta ctggaagaag ctctcgatcc cccagTGAtc cttgttagag gggaggcaaa
                                              <- E4 end
4021 tacactaaaa tgctttcgca acagagctaa gataaaatac atgggactgt ataggtcatt
4081 tagtacaact tggtcatggg tggcaggaga tggcactgag cgtctaggca ggcccagaat
4141 gctcattagc tttttcgtcct acaatcagag aagggatttt gatgacgtgg tgagataccc
4201 gaaaggagtt gaaaaatcat atggcaacct tgacagtctc TAAccactaa tgctatgctg
                                              <- E2 end
4261 ctttgctact aacaacacTA Acaaattagc ttttttatact tttttacttt tgtacctgca
                L2 orf start ->
4321 ATGgcgcgtg ctaaaagggt caagcgagac tctgtaacac atatatacca gacctgcaaa
L2 cds ->
4381 caagcaggca catgcccccc tgatgttgtg AATAAAgtgg aacaaacaac agttgctgac
                                signal ->
4441 aatattttga aatatggcag tgctggtgtc tttttttggtg gccttggcat tggttcgggc
4501 cgaggtactg ggggtgctac cgggtacgtg ccacttagtg aaggtcctgg tatccgtgtc
4561 ggaggtaccc ccacggttgt aaggccttca ttagttcctg aagcaattgg gccagtcgat
4621 attttgccca ttgatacaat tgatcctgtg gagcctacag catcgtccgt ggttcctctt
4681 actgaatcca ctggacctga tttacttcca ggtgaggtgg aaacaatagc tgaaattcat
4741 cctgttgctg aaggcccatc agttgatacc cctgtggtta ccacaagcac gggttccagt
4801 gctgtttttgg aagttgctcc agagcctata cctccaacac gggttagaat ctcacgtaca
4861 caatatcaca atccttcctt tcaaattatt acagaatcaa cacctgcaca gggtgaaagc
4921 tctcttgcag atcacatttt agtgacatca gggtccgggg acaaagaat aggggctgat
4981 attactgatg aaattgaact tcaagaactt cctagtagat atactttga aaatgaagaa
5041 ccaactccac ctagacgtag cagcacacct ttacaggcca cacgagctgc aggtagacgg
5101 agaggtgttt ctttaactaa tagacgtcta gtacagcagg tacctgtaga aaatcctttg
5161 tttttaactc aaccctctcg attggtgcgt tttgcttttg aaaatcctgc atttgaagaa
5221 gaagttacaa atatatttga acatgatgta gatgcttttg aagaaccacc agatagggat
5281 tttcttgatg tccaacggtt gggtcgtccc caatattcta caactcctgc agggtatgtg
5341 agggttagta gattaggtac tcgtgcaacc attcgcacgc ggtctggtgc acaaataggg
5401 tcacaggtgc acttttatag ggatctcagt tccattaata ctgaagatcc tatagaattg
5461 cagttgttgg ggcaacattc ggggatgct agtatagttc aaggccctgt ggaaagcaca
5521 tttatagacg taaatgtgtc tgaaaatcct ttgtctgaaa gtgtggaagc ctttttctgat
5581 gatttattgc tggatgaagc tgtggaggat tttagtgggt cacaattagt tattggtaat
5641 agaagaagta ccacttctta cactgttccc agatttgaaa ctactaggag tggttcttat
5701 tatgtccagg acagtaaagg atattatgtt gcatacccag aatctcgcaa taatgcagaa
5761 ataatttatc ctacacctga tatACCTGTA GTGGTAAtac acactcATGa caatacaggt
                 E2 binding ->L1 orf start ->
                                first 'ATG' of L1 cds ->
5821 gactttttatt tacatcctag tcttcgatgg cgcaaacgta aaagaaaata tttgTGAttt
                                                  <- L2 end
```

```
5881 tgcattgcAG ATGgcagtgt ggcattcggc taatggtaaa gtataccttc ctccatcgac
                  /\ 3' sj
      probable L1 ->
      cds start
5941 accggttgcc agggtgcaaa gcacggatga atatatacaa aggacaaata tttattatca
6001 tgcattcagc gacagactac tgactgtagg ccatccatac ttcaatgttt acgatattac
6061 tggtaacaaa ttagaagtcc ctaaagtgtc tggtaaccaa cacagggttt ttcgcttgaa
6121 attacctgat cctaatagat ttgcattagc tgacatgtca gtttataacc cagataaaga
6181 acggttggtg tggagctgta ggggccttga aataggtaga ggacaaccgt taggtgttgg
6241 cagtacggga cacccactat ttaataagtt gaaagacact gaaaacagta attcatatat
6301 aaaatcttct aaagatgata gacaggatac atcatttgat cccaaacaaa ttcaaatgtt
6361 tatcgtggga tgcacacctt gtaggtgaa catttgggac aaggctatac cttgtgaaaa
6421 ggagcgccag gacaacagac tatgcccacc aattgaattg aaaactactt atatagaaga
6481 tggcgacatg gcagatatag gttttggaaa tttaaacttc aaaaatctgc aagagagtag
6541 gtcagatgta agtttggata tagttaatga aacctgcaaa tatcctgact tcttaaaaat
6601 gcagaatgat gtttatggag atgcctgttt tttttatgct cgcagagagc aatgttatgc
6661 cagacatttt tttgtccgtg gaggtaaaac gggtgacgac atccctgacg ccagaattga
6721 taatgggact tttaagaatc agttttttcat tcctgggggct gacggccaag atcaaaagac
6781 cataggaaat gccatgtatt acccaactgt tagtggctca ttggtgtcta gtgatgctca
6841 attgtttaac aggcccttct ggctccagcg agcacaaggt cataataatg gcattctgtg
6901 ggctaatcag atgtttatta cagtggtaga caacacacga aatactaact ttagtatttc
6961 aatatataac aataatgggg cactaaagga catcaatgat tacactgcag agcaatttag
7021 agaatatcaa aggcacgtgg aggaatatga aatttcatta atattacagc tatgtaaggt
7081 tcctctgaag gcagaagtat tggctcagat aaatgctatg aattcttctt tattggaaga
7141 ttggcagtta ggttttgtac ctactccaga taaccctatt caagacacct atcgatatat
7201 tgattcatta gccactcgct gtcctgataa aaccccccct aaggaaaaag aggatcccta
7261 caagggggtta aagttctggg atgtggatct tactgaacgg ttgtcactag atttggatca
7321 atactcttta ggtagaaagt ttctatttca agccggctta cagcagacgA CCGTAAGCGG
                                            E2 binding ->
7381 Tacaaaatca gtgtcttatc gagggttcac cagaggaacc aagcgcaagc gaaaacagTA
7441 AtatgACCGT TTTTGGTaca gatttataaa cttttacaca gtattcaagg aatgtttgtt
       <- L1 cds
  E2 binding ->
7501 tactctgact aagtataact ctaccaaggg aaccgACCGC ACCCGGTaca gtcaacgata
                              E2 binding ->
7561 ctgctgccaa tatagactct gtttagtgcc agaacatatc atcttggaag cagatcgacc
7621 gtgttcgttg taacacgctc ggattagaga cattgccaag gaagatttaa tctacaatcg
7681 ctgttggcaa tcgcttttgg ctgagatagc tgaccggtaa cg
//
```

```
LOCUS       HPV37          7421 bp    DNA             VRL        18-JUL-1995
DEFINITION  Human papillomavirus type 37, complete genome.
ACCESSION   U31786
KEYWORDS    .
SOURCE      Human papillomavirus type 37.
REFERENCE   1  (bases 1 to 7421)
  AUTHORS   Delius,H.
  JOURNAL   Unpublished, Sequenced by Hajo Delius, Deutsches
            Deutsches Krebsforschungzentrum, Angewandte Tumorvirologie,
            I.N.F. 506, W-6900 Heidelberg, Germany
REFERENCE   2  (sites)
  AUTHORS   Scheurlen,W., Gissmann,L., Gross,G. and zur Hausen,H.
  TITLE     Molecular cloning of two new HPV types (HPV 37 and HPV 38) from a
            keratoacanthoma and a malignant melanoma
  JOURNAL   International Journal of Cancer 37 (4), 505-510 (1986)
REFERENCE   3  (bases 1 to 7421)
  AUTHORS   Farmer,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     HPV37, as well as HPV9, was found in a keratoacanthoma in a patient
            who also had a basaloma. The HPV37 DNA was present as a circular
            monomeric episome, with approximately 10 copies per diploid cell.
            HPV37 was not detected in the basaloma. Keratoacanthoma is "a
            rapidly growing benign skin tumor which originates from the hair
            follicle and may resolve spontaneously after a period of months.
            It grows invasively and occurs as a single lesion or as multiple
            tumors. Tumors are found preferentially in areas exposed to
            sunlight. This proliferation has been suspected to be of viral
            origin, but there exists as yet no evidence to support this
            concept." [2] HPV37 was not found in 231 other tumor DNAs
            originating from different tissues (including 6 keratoacanthomas and
            35 malignant melanomas, as well as 190 other tumors); thus no
            correlation has been found so far between HPV 37 and any tumors of
            the skin or other tissues. HPV-37 is closely related to HPV 9, 15,
            17, 22, 23, 38 by cross-hybridization as well as phylogenetic
            analysis. These types also have in common a relatively short (7.4
            kb) genome.
BASE COUNT     2345 a   1343 c   1666 g   2067 t
ORIGIN      200 bp upstream from beginning of E6 cds
        1 agccaagaat atttggcaga acattttctt ggaagacaAC CGATAACGGT aagattgtaa
                                      E2 binding ->
       61 tctttcaACC GTAGGCGGTa ctttctgatt ggtttggccg attgtagcta acaacaatct
     E2 binding ->
      121 ttcttcaTAA atacatgtaa ccgcctgcgt taacttacat gatctaaata aatatgatga
      E6 orf start ->
      181 gcaatactta agagaatatA TGgctaggcc taagcctcaa tctgttcaac agcttgcaga
                   E6 cds ->
      241 tactttatgt ataccctttag tagatgtttt actgccttgc agattttgtt atagattctt
      301 agcatatata gaattgatcg catttgatcg aaaaggtctt caactaattt ggaccgaaga
      361 agatttagtg tatgcgtgct gtactagctg tgcctatgct acagcacagt ttgaatttac
      421 cagtttctat gagcactcag ttagtgggag ggagatagaa gagatagagc aaaagccaat
      481 aggagaaata gccatacgct gcaaattttg cttaaagtta ttggatttgt tagagaagtt
      541 ggagacttgc tatactcagc aacaatttca caaggttagg cgcaattgga aaggcttgtg
      601 TAGacattgt gggtcgatag gATGAttggg aaagaagcta caataccaga aatagtgctt
    E7 orf start ->                     <- E6 end
                         E7 cds ->
      661 gagctgcaag agcttgtcca gcccactgct gacctgcatt gttacgaaga gttgagtgaa
```

```
 721 gaagagacag aggaggagcg tcctcacatc ccttacaaga ttgtagctcc gtgctgcttt
 781 tgtggttcta aactacgact gaTAGttgtt gcaacgccta ttggaattag atcacaagaa
                       E1 orf start ->
 841 gagctattac ttggtgaagt gcagctggtt tgtccaaact gtcgggggaa gcttcgccAT
                                                                E1 cds ->
 901 GacTGAcgac acgaaaggta caaaatttga tcctaaagaa ggatgtagtg attggtttgt
         <- E7 end
 961 gctagaagca gaatgctctg acaatagttt agatggtgat ttggaaaagt atttgaaga
1021 agggaatgat actgacattt ctgatttaat agatgatgag gacactgttc agggaaattc
1081 ccgcgaattg ttatgccagc aacaaagtga ggaaagcgag caacaaatac atttgctaaa
1141 acgaaagtat ttcagttcac aagagattct gcagttaagt cctcgtctgc agtctattac
1201 tatttcgcca cagcataagt ctaaaaggag attatttgaa ggagacagcg gactagaact
1261 gtcatttaat gaagctgaag attttactca gcagactttg gaggtgcagg aggtatcggc
1321 atccggctct gagccggcag accagggtgc caagggactg ggcattgtta agaccttct
1381 taaatgtagt aatgttaaag ctatgttgtt agcaaaattt aaagaagcat ttggagttgg
1441 ctttatggaa cttactaggc aatataaaag ttgtaaaaca tgttgcagag attgggttgt
1501 aacgttgtat gcagttcaag atgaactgat agaaagctcc aaacagctgt tgcttcaaca
1561 ctgtgcttat atatggttgc agcatatgcc tccaatgtgt ttatatttat tgtgtttaa
1621 tgtgggtaaa agtagagaaa ctgtttttag actgctaatg aatttattgc aagtagcaga
1681 agtacaaata ttggctgaac ctccaaagct tcggagcaca ttatctgcac tgttttggta
1741 taaaggtagc atgaatccaa atgtctatgc acatggtgaa atcctgagt ggattatgac
1801 acaaaccatg atcaatcacc aatcagcaga agctacacaa tttgatttat ccactatgat
1861 acaatatgca tatgacaatg atttaataaa tgaagatgaa attgcttata attatgccaa
1921 attagcagat acagacgcta atgccagagc tttttttacag cacaatagtc aagccagatt
1981 cgttagagaa tgtgcactaa tggttagata ttacaaacga ggtgaaatga agatatgag
2041 catatctgcc tggatacata ataaaatgtt agttgtggaa ggcgaaggac attggtctga
2101 tattgtaaag tttgtaagat tccaagatat caatttttata aggtttctag atgtctttaa
2161 atcattttg cataacactc ctaaaaagaa ttgtcttta ttttatggtc cacctgatac
2221 aggcaaatca atgtttacta tgtctttaat taaagtgtta aaaggaaaag ttttatcctt
2281 tgcaaattat aaaagtaatt tttggttgca gccgttggca gatactaaaa ttgctttaat
2341 agatgacgtc acgcatgtgt gttgggatta catagatcaa tatttaagaa atggattgga
2401 tggtaatttt gtttgtttag acctaaaaca tagagcgcca tgtcaaatta gtttccacc
2461 attattactg acttccaata tggatattat gaaggaagaa aggtatagat atttacatag
2521 caggtgcat gcttttgcat ttccaaataa gtttcctttt gatagtaaca ataagccaca
2581 atttcgactt actgaccaaa gctggaaatc ttttttttgaa aggctttgga aacagtTAGa
                                                        E2 orf start ->
2641 tctcagtgac caagaagacg agggagacgA TGgacacact cagcgatcgt ttcaatgcac
                              E2 cds ->
2701 tgcaagagaa cctaatggac atttaTGAgt caggtcgaga tgacctagag acccaaatta
                                  <- E1 end
2761 tgcattggca acttctaagg caggagcaga tcctgtttca ttatgccaga aaaaatggag
2821 tcatgcgttt aggatatcaa cctgtacctc ctttagccac cagtgaagct aaagcaaaag
2881 atgcaattgg catggttata ttattagaaa gtttacaaca gtctgcttat ggtaaagagt
2941 cctggacact tacacaaact agtttggaga ccgtgaggag tccacctgca aattgtttta
3001 aaaagggccc tcagaacatt gaagtgatgt ttgacaatga ccctgaaaat ctaatggtgt
3061 atactgcctg gtcatttatt tattatcaga ctgtagatga cacgtggaac aaggttgagg
3121 gacatgttga ctactatggt gcatattatt ttgaaggaga ttTAAaagtc tattatatac
                                         E4 orf start ->
                                         NH2 terminus unknown
3181 aatttgaagg tgatgctgcc aggtttagca aaaactggacg ctgggaagta catgttaaca
3241 aggacactat ctttgctcct gttactagct cttcgccggc agctggagaa gggacagacg
3301 gggcagcctc cgtccacacc gtatccgggt cgccgctcgc acgggggattc tctaccacct
3361 ccgtgtccac cagaaaacgg acaccaccac ggcgatacag aagaaaagca tctagcccta
3421 caaccaccgc cgcccggcaa aaaagacaag gagaagacac cgcaacaagg cgatcaaggt
3481 ccacctcccg ggggaaacaa gcaacctcca ggggagggga ccgacgcaga cggagacgag
3541 aacgctccta ctcccgagac acctccagtt cccccgacag gggaagggga gggagaagta
3601 gagggggggcc cgagacacga tcccaatcaa ggtccctctc acgatcccgg tcgcggtcgc
3661 gatccagagg gtcttcttcc aggggtggcg ttgcgcctga cgcagtggga aagtcagttc
3721 gaacagttgg tagagaccat agtggacgac ttaaaagatt actggacgaa gctagggatc
```

```
    3781 ccccagTAAt tgtgctgcgt ggtgatgcta acaaattaaa atgctatcgc tatagagcta
                   <- E4 end
    3841 agaaaaagca tggaaaccta gttaagtact acagtaccac gtggtcatgg gttgggggca
    3901 gcaccaatga tagaattgga aggtcacgca tgttacttgc atttcaatcc aatacagaaa
    3961 gagagttgtt tttaaaaact atgaaattac caccaggagt tgattggtca ctgggtcatt
    4021 tagatgaatt gTGAaaacag ctttttttata acaaactaac attgctttttg cttttgctac
                       <- E2 end
    4081 TAAcctacta acgttccaAT Ggctcgcgca cgtcgtacca aacgtgcgtc tgtaactgac
L2 orf start ->        L2 cds ->
    4141 atttacaggg gttgcaagca ggccggcact tgccccccccg atgtaattAA TAAAgtggaa
                                                        signal ->
    4201 caaacaacaa ttgcagacaa aattttgaag tatggtggtg ctggtgtttt ttttggtggg
    4261 cttgggatta gcaccggccg aggaacaggt ggtgctacag gatatgtccc tttgggggaa
    4321 ggccctggag tgcgcgtagg aggcgcaccc accattgttc gccctggggt catacctgaa
    4381 ttgattgggc cagcagatgt aatacctatt gacacagtca ctccaattga ccccgcagca
    4441 cccagtattg tcacaattac agacagtagt gctgttgacc ttttacctaa tgaaatagaa
    4501 acaattgcag aagtgcatcc tgtgccaaca gacaatttgg atattgatac tcctgtagtt
    4561 acaggaggcc gggattccag cgctgttttg gaagttgctg atcctagtcc ccctgtgcga
    4621 acaagagttt ccagaacaca atatcataat ccttcttttc aaataataac tgaatctaca
    4681 cctttagcag gagaatctgc tttagctgac catgttattg ttttgaagg cactggagga
    4741 caaaatatag gtggttctcg aaatgcaact atagaaacag ctcaagaaag ttttgaaatg
    4801 caaagttggc cgagtaggta tagtttttgaa atagaagaag gaacacctcc tagatctagc
    4861 acaccagtac aaagagcagt acaatcactc tctagtttaa gacgggcatt gtataatagg
    4921 agattaacag aacaggtagc agtcacggat cctttattct tgagtagacc ctcacaatta
    4981 gtacagtttc agtttgacaa tcctgcattt gaagaagaag taactcaaat atttgagagg
    5041 gatttagagg ctgtagaaga acctccagat agacagtttt tggatgttat tcgcttaggt
    5101 agacctactg ttgctgaaac accacaagcg tatttaagag taagcagatt aggacgtcgt
    5161 gctaccatcc gtactcgtag tggagcacag gtgggggctc aggtacattt ttatagagat
    5221 ttaagtacta tagattctga tgccctagaa atgcaattat taggagaaca ttcaggtgat
    5281 actactatag tacaaggacc tgtagaaagt tcatttgttg atataaatat tgatgaacca
    5341 ggtcccttaa atatagggca acaagagtct actatggcag atgacacaga tttttaattct
    5401 gcagatttat tgttagagga tgctgtagaa gacttctcag gatctcagtt ggttttttgga
    5461 acctcacgcc gcagtacaaa ttctatcaca atacctagat ttgaaactcc aagagatact
    5521 ggattttata tacaagatat tcaaggttac aatgtagcct atcctgagtc acgtgacaca
    5581 acacaagtta tcttgccaca acctgaaaca ccaactgtag ttattagatt tggagaggca
    5641 ggtacagact attatttaca tcctagctta aaaaagaaaa agagaaaacg caaatatttta
    5701 TAAttgtttt tacagATGac tttgtggctg ccagcgacgg gtaaagtata cttgcctcca
                  <- L2 end
L1 orf start ->        L1 cds ->
    5761 acaccaccag tagcccgggt gcaaagcacg gatgattatg tggaaagaac aaatgtattc
    5821 tatcatgcca tgagcgatcg tctcctaact gtaggacacc catattatga tgtaagatct
    5881 agtgatggct aaaaatcga ggttcctaaa gtatctggaa atcaatacag agctttttagg
    5941 gttaggttgc cagatccaaa taaatttgct ttagcagata tgtcagtata taatccagaa
    6001 aaggaaaggt tggtgtgggc ctgtgcgggc ttggagatag gccgagggca accacttgga
    6061 gtaggaacga caggtcaccc tttatttaat aaaattaaggg acactgagaa taatagtaat
    6121 tACCAAGGGG GGTcacggga tgatagacaa aacacatcat ttgatccaaa acaagtacag
E2 binding ->
    6181 atgtttgtgg ttggatgtgt gccatgcatg ggtgaacatt gggataaagc accagtttgt
    6241 gcatcagagg aaaataatca gacaggacag tgtccaccac ttgaattaaa aaacacagtg
    6301 attgaagatg gggacatgtt tgatataggg ttcggaaata ttaacaataa ggttctctct
    6361 actaataaat cagatgttag tttagatata gtaaatgaaa atgcaaaata ccctgatttt
    6421 ttaacaatgg ctaatgatgt ttatgggggat gcatgtttct tttttgctag gagagaacaa
    6481 tgttatgcca gacattattt tgtaagaggg ggaaatgtag gtgatgctat ccccgatggc
    6541 actgttaatc aggaccacaa atattactta cctgccaaat cagaccagca gcagtatctg
    6601 ttaggcaatt ctacctattt tcccactgtt agtggatctt tagtaacatc tgatgctcag
    6661 ctctttaaca ggcctttttg gttacgcaga gctcaaggtc acaacaatgg cattttatgg
    6721 ggtaatcaaa tgtttatcac agttgctgat aatacacgga acacaaactt ttctattagt
    6781 gtgtctactg acaatggcga agttacagaa tataattctc aaacactcag agaataccta
    6841 agacatgttg aagaatacca gctttcaatt attttacaac tttgtaaagt tcctttaaag
```

```
6901 gctgaggttt taactcagat aaatgcaatg aattctggta tattggaaga gtggcaatta
6961 ggatttgtac ctactccaga taattcagta catgACCTTT ATAGGTacat taattcaaag
                             E2 binding ->
7021 gctaccaagt gtcctgatgc agttgttgaa aaagaaaagg aagatccctt tgcaaaatat
7081 acatttttgga atgtagattt aactgaaaaa ttatcattgg atttagatca atatccttta
7141 gggaggaaat tcatctttca gtcgggattg caaagtagac ctagaattgt tcgatcgtct
7201 gtaaaagtgt ctaaaggtac aaagcgtaaa cggtcgTGAC CGTTTTCGGT ttccAATAAA
                                           <- L1 end    signal ->
                             E2 binding ->
7261 caAATAAAcc aataaggtat gtgaagcatt ttttaccatg ttcgtgacta aaccatataa
signal ->
7321 gtcaacgcca acaACCGCAC CCGGTttaat cagatataaa acacctggtg cgattttatc
          E2 binding ->
7381 agagctttg tggaagcacc tgaggcgacc gccagaactg c
//
```

```
LOCUS       HPV38           7400 bp    DNA              VRL        18-JUL-1995
DEFINITION  Human papillomavirus type 38, complete genome.
ACCESSION   U31787
KEYWORDS    .
SOURCE      Human papillomavirus type 38.
REFERENCE   1  (bases 1 to 7400)
  AUTHORS   Delius,H.
  JOURNAL   Unpublished, Sequenced by Hajo Delius, Deutsches
            Deutsches Krebsforschungzentrum, Angewandte Tumorvirologie,
            I.N.F. 506, W-6900 Heidelberg, Germany
REFERENCE   2  (sites)
  AUTHORS   Scheurlen,W., Gissmann,L., Gross,G. and zur Hausen,H.
  TITLE     Molecular cloning of two new HPV types (HPV 37 and HPV 38) from a
            keratoacanthoma and a malignant melanoma
  JOURNAL   International Journal of Cancer 37 (4), 505-510 (1986)
REFERENCE   3  (bases 1 to 7400)
  AUTHORS   Farmer,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     HPV-38, as well as HPV17a, was found found in high copy numbers in
            a superficial spreading malignant melanoma of an immunosuppressed
            patient. The HPV38 DNA was present as a circular monomeric episome,
            with approximately 50-100 copies per diploid cell. Superficial
            spreading malignant melanoma is a tumor derived from
            pigment-producing cells of the skin with rapid invasive growth and
            high incidence of metastasis. Tumors are found preferentially in
            areas exposed to sunlight. HPV38 was not found in DNA from 231
            other tumors originating from different tissues, including 6
            keratoacanthomas and 35 malignant melanomas, as well as 190 other
            tumors. Thus, no correlation has been found so far between HPV 38
            and any tumors of the skin or other tissues.  HPV-38 is closely
            related to HPV 9, 15, 17, 22, 23, 37 by cross-hybridization as well
            as phylogenetic analysis. These types also have in common a
            relatively short (7.4 kb) genome.
BASE COUNT     2305 a   1383 c   1625 g   2087 t
ORIGIN      200 bp upstream from beginning of E6 cds
        1 catctttggc agacgaagtg cACCGATAAC GGTaagactt ttctctttta accgtaggcg
                         E2 binding ->
       61 ttggtttatt attcctggca acaatggtgg ttaacaacca tcacacgtaa tcggtacaag
      121 caACCGCTTG TGGTagtaaa atgaattaaa aaaaaaaaca aggatatatt taaggggcct
 E2 binding ->
      181 gTAAgcttgg gatgtattcA TGgaactacc aaaacctcaa actgtgcagc agctcagtga
 E6 orf start ->       E6 cds ->
      241 taagttaaca gttcctgtag aggatctgtt attaccctgt agattctgca acagtttcct
      301 cacgtacatt gaattacgtg agtttgatta caagaactta cagttaatct ggactcaaga
      361 ggattttgtt tttgcatgtt gtagcagttg tgcttatgct tctgctcaat atgaatgtca
      421 gcagttttat gaattaactg tctttggccg tgaaattgaa caggtggagc aacagacaat
      481 aggccttatt gttataaggt gtcagtattg tttaaagtgt cttgatttga tagaaaaatt
      541 agatatctgt tgctctcatc aagcatttca caggtTAGa ggcaattgga aaggaaggtg
                                       E7 orf start ->
      601 caggcattgc aaagcaatag aATGAttggg aaacaagcta ctcttcgtga tatagttctt
                                       <- E6 end
                         E7 cds ->
      661 gaagagcttg tccagcccat tgacctgcat tgccacgagg agttgcctga tcttccagag
      721 gatattgaag catcagtggt agaggaggag ccagcataca ccccatacaa aatcatagtt
      781 ctttgtgggg gttgtgaagt aaggctaaaa ctatacgtgt gggccaccga cgctgggatt
```

```
 841 cggaatctgc aagattgttt gctgggcgac gTAAggcttc tgtgtcccac ctgtcgagaa
                                   E1 orf start ->
 901 gacattcgca ATGgcggacg aTAAaggtac tgatcctaaa gaaggctgta gtgattttat
          E1 cds ->              <- E7 end
 961 atatttagaa gctgaatgct ctgacattag tgacttagat aatgatttgg aaacattatt
1021 ggaagaaggt gcgggatccg atatttctga cttaataaat gatgaggttg ttgagcaggg
1081 aaattcccgc gaattattat gtcaacaaga gagagaggag agcgaactgc aggttcaata
1141 tctaaaacga aagtgtttca gtccgaaagc tgttcaggag cttagtcctc gtctgcagtc
1201 tatgaatata tcttcagagc ataaatctaa aaggagatta tttgtggagc aagacagtgg
1261 actggagcta tctctaaatg aagctgaaga ttctactcaa gagttggagg taccggcgag
1321 cgctccagcg ccggcagcag agggtgatat agggctgggt actgtaagag atctttttaag
1381 gagcagtaac agcagagcaa cactgttaag caaatttaaa gactcgtttg gggtcagctt
1441 tacagaactg acaagacaat ataaaagcaa taaacgtgt tgccaccatt gggtcttggc
1501 agtgtatgct gctaaggatg acttgataga tgcgtccaaa caattgttac agcagcattg
1561 ttttatata tggcttcaat cattttgtcc catgtcactt tatttatgtt gctttaatgt
1621 tggtaaaagt agagacactg ttgtaagact aatagctaca ttattacagg tgcatgaaaa
1681 tcatatattg tcagagccac caaaaaatag aagtattcca gcagcgttat tttggtataa
1741 aggaagtttg aatagtaatg tgttttgttt tggtgaagct cctgattgga ttctatcaca
1801 aacaatgata cagcatcaaa ctgctgacac tttgcagttt gacttgtctc gaatgattca
1861 atgggcctat gataatgatc atatagacga aagcattata gcttatcaat atgctaaatt
1921 agcagatatt gatagtaatg ctaaagcttt tttagctcat aacagccaag ttaaatatgt
1981 taaagagtgt gctttaatgg taagatatta taaaagagga gagatgaaag aaatgtctat
2041 ttctgcttgg attcatcact gcatatctaa agttgaagga gaaggcaatt ggcagcatat
2101 tgttaggttt attagatacc aaaaatttgaa ttttattatg tttctagata agtttcggac
2161 ctttttaaaa aatctgccaa aaaaaaattg tttattaata tatggtcctc ctgacacagg
2221 aaagtctatg tttgcaatgt cacttattaa actattgaaa ggtagtgtag tatcttttgc
2281 taattcgaaa agtcaatttt ggttacagcc actagctgat gggaaaattg gtttattgga
2341 tgatgcaact gatgtgtgtt ggcagtatat agattctttt cttagaaatg gtttagatgg
2401 taatttagtg tcgttagata taaaacataa agcaccttgt caaatgaaat ttcctccatt
2461 aattattact tccaatatta atttattaaa agaggaacga tacagatttt tacacagtag
2521 agtaacacaa attgattttc caaataagtt tcccttttgac tcagataata agcctttgtt
2581 tgaacttact gatcaaagct gggcatcttt ctttaaaagg ctgtggacac aatTAGagct
                                         E2 orf start ->
2641 cagtgatcaa gaagacgagg gagacaATGg aaactctcag cgcacgtttc actgtactgc
                       E2 cds ->
2701 aagagaagtt aatggacata taTGAatcag gtgtagagga cctggataca caaattcagc
                        <- E1 end
2761 attggcagct tttaagacaa gagcaaatta tttatcacta tgcaaggaga catggtgtta
2821 ctcgattggg ctatcaacct gtaccttctt tggcaagttc agaagccaaa gcaaaagatg
2881 ccatttctat ggtcctttta cttgaaagcc tgaaaaaatc caaatatgca gatgaacaat
2941 ggacattagc tcaaactagc ctggaggctg ttcgcagccc tcctgcagac tgttttaaaa
3001 aaggacctaa aaatattgaa gttgtatttg atggtgaccc tgaaaatctt atgtcatata
3061 ctgtgtggac atatatatat tacctgacag atgaggacat atgggaaaaa gtggaaggcc
3121 atgtggatta tacaggagcc tattattatg agggcaaatt aaaggtgtat tattTAAaat
                                         E4 orf start ->
                                         NH₂ terminus unknown
3181 ttgaaaatga tgctaaacga tatggtgtca caggattatg ggaagtacat gttaataaag
3241 acactgtgtt taccccccgtt accagttcta cgccgccagt tggagactcc accgactccg
3301 catccagggc ggcactcccc gagccttcca cctccgtgtc ccccgaacgg ccaccatccc
3361 aaacagcacg gcgatacggg agaaaagcat ctagcccttc aaccacctcc cgcaggcaaa
3421 ggaaaggaca aagagaaacc acaggcaccc aaaggagaag aaaaagcaga tcaaggtccc
3481 gaagcaccaa caggggaggg agggacaccc ggcgatcctc ctccagagga tcctcagtct
3541 cccccaccag gggaaggaga aggggaggag gggacagcag aaggcggggg ccggtcaccc
3601 gctcgagatc aagatccctc tcacgagcct cttctgcagg gggtggcata tcgcctgaca
3661 aagtgggaac ggcagttcga tcagttggta gacaaagtgg tggaagactt acgcggctac
3721 tggcagacgc tgcagacccc ccagTAAtat tgttacgtgg agatgccaat accttaaaat
                         <- E4 end
3781 gctatcgcta tcgatttaga aaaaaacatg ctggtggctt tcgctttgtt agcacaacat
3841 ggtcatggat aggagatgca tcaaatgatc gcataggg cg ctcacgaatg cttctagctt
```

```
3901 tttattcaga atcacaaaga gaaaagttta tacagactat gaaattacct acaggtgtag
3961 agtggtcatt aggacaattt gatgatttaT AGaataatct ataagagata tttttTATAT
                                    <- E2 end          signal ->
4021 ATtatgtaac cttttttacT AAcaatactg ctttgctact aacactATGg ttcgagcacg
                L2 orf start ->                    L2 cds ->
4081 tagaaccaaa cgtgcatctg ttactgatat atacaggggc tgcaaggctt ctaatacttg
4141 ccctcctgat gtaatcAATA AAgtggaaca atcaacaata gcagataaga ttttaaaata
                signal ->
4201 tggtagtgct gctgtctttt ttggtgggct gggtattagc actggtcgtg gtacaggcgg
4261 tgctacaggc tatgtgcctt tggggcaagg acctggagtg cgagtgggtg gcgcccccac
4321 agtggtccgc cccggggtga tacctgaagt aattggacca accgaactga tacctattga
4381 ctcagtcaca ccaattgacc ctacagcacc ttcaattgtg tcattaactg acagtagtgc
4441 tgttgacctt ttACCTGGAG AGGTtgaaac tattgcagaa gttcatcctg gccctataga
            E2 binding ->
4501 ccctatagaa attgataccc ctgttgtgag tggaggccgc aataccaatg ctatattgga
4561 agtggctgac cctcatccac ccactagagc tactgttagc agaactcaat ataataatcc
4621 tgctttccaa ataatttctg aagtaatccc tacctctgga gagtcttctc ttgcagatca
4681 cgtgttagtg tctgaagggt ctggtggcca gcagatagga ggtaccagaa cagcagaaga
4741 aattgagttg cagcctttgt tatctagata tagttttgaa attgaggagc caacaccacc
4801 gcgaagaact agcacccccт tacaaagagc aagacaacag ttttcatcat tacgcagagc
4861 attatataat agaaggctaa ctgagcaagt gggtgtcact gacccttat ttttcacatc
4921 accttccaaa ttggtgcgtt ccaatttga caatcctgta tttgatgaac aagtaacaca
4981 gatatttgag caggacatag cagactttga ggaaccaccc gatagacagt ttttggatgt
5041 ggttaaatta ggtaggccaa cattaactga gtctgcagag ggatatgtta gagtgagtcg
5101 tttgggaaga cggggaacga tccgaacacg cagtggtaca caaataggat cacaagtaca
5161 tttctatagg gatttaagta caattaatac agaagaaccc ttagaaatgc agttattggg
5221 tgagcattca ggtgatgctt caattgtaca aggtcctgta gaaagcactt tagtggatgt
5281 aaatgtgact gaggttcctg aaggtgttct tacagaaact tctatggatc cagatacttt
5341 taattcagag gatttattac tggatgatgc tatagaagac ttcagcggat ctcagttagt
5401 tgtaggaact ccacgcagat ccactacgtc aatcactgta cctagatttc agactcctca
5461 aaatcctacc atatattatc aggatataca ggggtatcat gtttcatatc ccgaaagcag
5521 agaaagaccc gccattattt atcctacacc cgatattcct acagTAGtta tacATGttgc
                                       first 'ATG' of L1 cds ->
                                       L1 orf start ->
5581 tgattcctct ggagattttt atttacatcc cagtttacga tggcgacggc gcaaacgcaa
5641 atatttaTAA tgtttttcAG ATGacacttt ggcttcctgc atctggtaaa atatacttgc
             <- L2 end     /\ 3' sj
        probable L1 cds start ->
5701 caccaacacc tccagttgcg cgcgttcaaa gcacggatga atatgtggaa cgaacagaca
5761 tctattacca tgcaactagc gatcgcctat taacagtagg ccatccatat tttgatgtca
5821 gatcacagga tggtcaaaaa attgaagttc ctaaggtgtc aggaaatcaa tataggtcat
5881 ttcgggtaac cttttccagat cccaataagt ttgctttggc agacatgtct gtttatgatc
5941 cagataaaata taggctggtg tgggcctgca aaggccttga ataggccga ggacagccat
6001 taggagttgg aactacagga catcctctat ttaataaagt aagagatact gaaaactcca
6061 gtaattatca aaacacatct actgatgaca gacaaaatac ctcttttgat cctaaacagg
6121 tgcaaatgtt ataataggc tgcactcctt gtctaggaga atactgggat aaagcacctg
6181 tatgtgataa tgcaggggac cagacaggcc tatgtcctcc actagaattg aaaaatagtg
6241 taattgaaga cggagacatg tttgatatag gatttggcaa tataaacaac aaaactctat
6301 cctttaatag atctgatgta agtttggata ttgttaatga aacctgcaaa tacccagatt
6361 ttcttaccat gtctaatgat gtttatggtg attcctgctt cttttttgtg cgacgggagc
6421 aatgctatgc cagacattat tttgttcgag gtggtgcagt gggtgacgct attccagatg
6481 gtactgtcaa ccaaaatcat aattattatt acctgcaaa aatggacag ggtcaacgca
6541 ctttagggaa ctctacgtat tttccaacag ttagtggatc cttggtgacg tctgatgctc
6601 agttatttaa tagaccattt tggttacaaa gagcacaagg ccacaataat ggtatttttat
6661 gggcaatca aatgtttgtt acagtcgctg ataatacccg aaatacaaac tttacaatca
6721 gtgtatccac tgaaaacggg ggtgctcaag aatatgattc tgcaaatatt agagaatatt
6781 taagacatgt tgaggaatac caattgtcat ttatattgca attgtgtaag gttcctttaa
6841 atgctgaagt gctgacacag attaatgcta tgaattctgg aatattagaa aattggcaat
6901 taggctttgt acccacccca gacaattctg tacacgatac atatcgttac ataacatcta
```

```
6961 aagcaactaa atgtccagat gcagtgcctg aaacagaaaa agaagatccc tttggtcaat
7021 atacattttg gaatgtggac atgactgaaa aattgtctct agatttggat caatatcctt
7081 tggggcgcaa atttttattc caagcaggtt tacaaacagc acgaacacgt gctgtcaaac
7141 ggccgttagt aagaaaatct tccaaatctg taaaacgcaa aaggacccag TAACCGTTTT
                                                               <- L1 end
                                                      E2 binding ->
7201 CGGTcgcccA ATAAAattta ttaactaatg tggtatgtga agcatttttt gaccttcttt
        signal ->
7261 gtgactaaac cgaacaagtc aacaccagta ACCGCGCCCG GTtaatcaga ttataaattc
                            E2 binding ->
7321 ctgaagggca gatttcaatc agtgcagata tcatctagca cctgcagcaa ccgccaagac
7381 tttgccagga cttggcagaa
//
```

ICPX1

```
LOCUS       HPVICPX1               276 bp   DNA             VRL        13-JAN-1995
DEFINITION  Human papillomavirus unidentified type (ICPX1) L1 gene, partial
            cds.
ACCESSION   L38914
KEYWORDS    L1 gene.
SOURCE      Human papillomavirus unidentified type (ICPX1) (individual_isolate
            Kremsdorf et al.) DNA.
REFERENCE   1  (bases 1 to 276)
  AUTHORS   Berkhout,R.J., Tieben,L.M., Smits,H.L., Bavinck,J.N., Vermeer,B.J.
            and ter Schegget,J.
  TITLE     Nested PCR approach for detection and typing of epidermodysplasia
            verruciformis-associated human papillomavirus types in cutaneous
            cancers from renal transplant recipients
  JOURNAL   J. Clin. Microbiol. 33 (3), 690-695 (1995)
  MEDLINE   95270695
COMMENT     HPVICPX1 was isolated from an immunocompetent patient [1],
            and is likely to constitute a new HPV type.  The most closely
            related known type appears to be HPV-8.
            NCBI gi: 623453
BASE COUNT       91 a     51 c     54 g     80 t
ORIGIN
        1 atttcggttt atagtgaagg tggacaaata aaagatatca gggactacac atctacacag
  L1 cds ->
       61 ttcagggaat atttaaggca tgtggaggaa tatgaaatat ctgtcatatt gcagttatgt
      121 aaaatacctt taaaagcaga agtcttggct cagataaatg ccatgaaccc cttattattg
      181 gaggactggc aattaggatt tgtccctaca cctgacaatc caattcatga tacctacaga
      241 tttattgact ctttggctac ccgatgccct gacaaa
                                          L1 cds ->
//
```

```
LOCUS       HPVRTRX1              267 bp    DNA              VRL       13-JAN-1995
DEFINITION  Human papillomavirus unidentified type (RTRX1) L1 gene, partial
            cds.
ACCESSION   L38918
KEYWORDS    L1 gene.
SOURCE      Human papillomavirus unidentified type (RTRX1) DNA.
REFERENCE   1  (bases 1 to 267)
  AUTHORS   Berkhout,R.J., Tieben,L.M., Smits,H.L., Bavinck,J.N., Vermeer,B.J.
            and ter Schegget,J.
  TITLE     Nested PCR approach for detection and typing of epidermodysplasia
            verruciformis-associated human papillomavirus types in cutaneous
            cancers from renal transplant recipients
  JOURNAL   J. Clin. Microbiol. 33 (3), 690-695 (1995)
  MEDLINE   95270695
COMMENT     HPVRTRX1 was isolated from renal transplant patients [1],
            and is likely to constitute a new HPV type.  It was isolated
            during a search for HPV DNA in 5 out of 53 squamous cell
            carcinomas from 2 of 26 renal transplant patients.  The most
            closely related known HPV types appear to be HPVs 22, 23, and
            38.  The partial L1 sequences for HPVVS42 and HPVVS73 do not
            overlap with that of HPVRTRX1; from their similarity to
            HPVs 22, 23 and 38, it is possible that HPVRTRX1 and one of the
            other two isolates are from the same novel HPV type.
            NCBI gi: 623461
BASE COUNT       98 a     41 c      52 g      76 t
ORIGIN
        1 atcagtgtta ctagtgaaga tttaagtaca gcaaaatatg atgctaaaaa tatcagggaa
  L1 cds ->
       61 tatatgagac atgtagaaga atatcaatta tcatttatat tacagttatg tagggtaccc
      121 ttagaggctg aggtgctaac ccagattaat gctatgaatt caggtatatt agaaaactgg
      181 caactagggt ttgttccaac accagataat gcagtgcatg acacatatcg ttaccttagt
      241 tcaaaagcta caaaatgtcc agatgca
                                  L1 cds ->
//
```

RTRX2

```
LOCUS       HPVRTRX2                276 bp    DNA             VRL       13-JAN-1995
DEFINITION  Human papillomavirus unidentified type (RTRX2) L1 gene, partial
            cds.
ACCESSION   L38919
KEYWORDS    L1 gene.
SOURCE      Human papillomavirus unidentified type (RTRX2) DNA.
REFERENCE   1  (bases 1 to 276)
  AUTHORS   Berkhout,R.J., Tieben,L.M., Smits,H.L., Bavinck,J.N., Vermeer,B.J.
            and ter Schegget,J.
  TITLE     Nested PCR approach for detection and typing of epidermodysplasia
            verruciformis-associated human papillomavirus types in cutaneous
            cancers from renal transplant recipients
  JOURNAL   J. Clin. Microbiol. 33 (3), 690-695 (1995)
  MEDLINE   95270695
COMMENT     HPVRTRX2 was isolated from renal transplant patients [1],
            and is likely to constitute a new HPV type.  It was isolated
            during a search for HPV DNA in 2 out of 53 squamous cell
            carcinomas from 2 of 26 renal transplant patients.  The most
            closely related known HPV types appear to be HPVs 22, 23, and
            38.  Although clearly more closely related to EV-associated types
            than to other HPV types, it is not particularly closely related
            to any known types, but it does appear to be more closely related
            to the novel sequences for HPVRTRX4 and HPVRTRX5.
            NCBI gi: 623463
BASE COUNT       99 a     50 c     53 g     74 t
ORIGIN
        1 atagctgttt atcaggaaca gaagaaggta aaagaaatag agagttacga ttctaccaag
  L1 cds ->
       61 tttaatgaat tccaaaggca tgttgaagaa tatgaagtat cacttatttt acagctttgt
      121 aaaattccac taaaggctga ggtgctagcc aaaattaatg caatgaactc tgacattctg
      181 gagaattggc agttagggtt cgtacctaca ccagataatc ctattcacga cacatacaga
      241 tatttagact ctttggccac acgctgccca gaaaaa
                                         L1 cds ->
//
```

```
LOCUS       HPVRTRX3              267 bp    DNA              VRL        13-JAN-1995
DEFINITION  Human papillomavirus unidentified type (RTRX3) L1 gene, partial
            cds.
ACCESSION   L38920
KEYWORDS    L1 gene.
SOURCE      Human papillomavirus unidentified type (RTRX3) DNA.
REFERENCE   1  (bases 1 to 267)
  AUTHORS   Berkhout,R.J., Tieben,L.M., Smits,H.L., Bavinck,J.N., Vermeer,B.J.
            and ter Schegget,J.
  TITLE     Nested PCR approach for detection and typing of epidermodysplasia
            verruciformis-associated human papillomavirus types in cutaneous
            cancers from renal transplant recipients
  JOURNAL   J. Clin. Microbiol. 33 (3), 690-695 (1995)
  MEDLINE   95270695
COMMENT     HPVRTRX3 was isolated from a renal transplant patient [1],
            and is likely to constitute a new HPV type.  It was isolated
            once during a search for HPV DNA in 53 squamous cell carcinomas
            from 26 renal transplant patients.  The most closely related
            known type appears to be HPV-9.  The partial L1 sequence for
            HPVVS92 does not overlap with that of HPVRTRX3; from their
            similarity to HPV-9, it is possible that HPVRTRX3 and HPVVS92
            are from the same novel HPV type.
            NCBI gi: 623465
BASE COUNT      90 a     45 c     50 g     82 t
ORIGIN
        1 attagtgtcg ccacagatgc aggtgtaact actgaatatc aggctaatac aatcagagaa
  L1 cds ->
       61 tatttaagac atgttgagga atatcaacta tctctcattt tacaactctg taaagtacct
      121 ttaaaggctg aggtattaac tcagatcaat gcaatgaatt ctggtatttt agaggactgg
      181 cagttaggct ttgtgcctac agcagataat tcagtacatg atatttacag atatattaat
      241 tctagagcca caaaatgtcc agatgct
                              L1 cds ->
//
```

RTRX4

```
LOCUS       HPVRTRX4              276 bp   DNA              VRL       13-JAN-1995
DEFINITION  Human papillomavirus unidentified type (RTRX4) L1 gene, partial
            cds.
ACCESSION   L38921
KEYWORDS    L1 gene.
SOURCE      Human papillomavirus unidentified type (RTRX4) DNA.
REFERENCE   1  (bases 1 to 276)
  AUTHORS   Berkhout,R.J., Tieben,L.M., Smits,H.L., Bavinck,J.N., Vermeer,B.J.
            and ter Schegget,J.
  TITLE     Nested PCR approach for detection and typing of epidermodysplasia
            verruciformis-associated human papillomavirus types in cutaneous
            cancers from renal transplant recipients
  JOURNAL   J. Clin. Microbiol. 33 (3), 690-695 (1995)
  MEDLINE   95270695
COMMENT     HPVRTRX4 was isolated from a renal transplant patient [1],
            and is likely to constitute a new HPV type.  It was isolated
            once during a search for HPV DNA of 53 squamous cell carcinomas
            from 26 renal transplant patients.  The most closely
            related known HPV types appear to be HPVs 22, 23, and 38.
            Although clearly more closely related to EV-associated
            types than to other HPV types, it is not particularly
            closely related to any known types, but it does appear to
            be more closely related to the novel sequences for HPVRTRX2 and
            HPVRTRX5.
            NCBI gi: 623467
BASE COUNT       89 a     51 c     48 g     88 t
ORIGIN
        1 attgctgttt acaatgaagc aggtaaaata caggatattt cttcatatga ttctaccaaa
  L1 cds ->
       61 tttagagaat ttcaaaggca cgtagaggaa tatgaagttt ctttaatttt acaactatgt
      121 aagattcctc taaaggctga ggtgctggcc caaatcaatg ctatgaatcc atctatttta
      181 gaagattggc aattgggctt tgtacctaca ccagacaatc ctattcatga tcgttatcgc
      241 tacatagatt ctttagccac tcgatgtccc gaaaaa
                                          L1 cds ->
//
```

```
LOCUS       HPVRTRX5              276 bp   DNA              VRL      13-JAN-1995
DEFINITION  Human papillomavirus unidentified type (RTRX5) L1 gene, partial
            cds.
ACCESSION   L38922
KEYWORDS    L1 gene.
SOURCE      Human papillomavirus unidentified type (RTRX5) DNA.
REFERENCE   1  (bases 1 to 276)
  AUTHORS   Berkhout,R.J., Tieben,L.M., Smits,H.L., Bavinck,J.N., Vermeer,B.J.
            and ter Schegget,J.
  TITLE     Nested PCR approach for detection and typing of epidermodysplasia
            verruciformis-associated human papillomavirus types in cutaneous
            cancers from renal transplant recipients
  JOURNAL   J. Clin. Microbiol. 33 (3), 690-695 (1995)
  MEDLINE   95270695
COMMENT     HPVRTRX5 was isolated from renal transplant patients [1],
            and is likely to constitute a new HPV type.  It was isolated
            during a search for HPV DNA in 2 out of 53 squamous cell
            carcinomas from 2 of 26 renal transplant patients.  The most
            closely related known HPV types appear to be HPVs 22, 23, and
            38.  Although clearly more closely related to EV-associated types
            than to other HPV types, it is not particularly closely related
            to any known types, but it does appear to be more closely related
            to the novel sequences for HPVRTRX2 and HPVRTRX4.
            NCBI gi: 623469
BASE COUNT       91 a     42 c      49 g      94 t
ORIGIN
        1 atagcagttt ataatgattc tggtgaaatt aaagacattg cttcttatga ttccactaaa
  L1 cds ->
       61 tttcgagagt ttcaaagaca tgtggaagaa tatgagattt ctttaatttt acagttatgc
      121 aaaattcctt taaaatcaga ggtattagct caaattaatg ctatgaatcc tacaatactt
      181 gaggattggc aattaggttt tgtgccaact cctgataatc caatacagga tgcttacaga
      241 tatttggatt ctttggctac acggtgccca gataaa
                                        L1 cds ->
//
```

RTRX6

```
LOCUS       HPVRTRX6              267 bp    DNA              VRL        13-JAN-1995
DEFINITION  Human papillomavirus unidentified type (RTRX6) L1 gene, partial
            cds.
ACCESSION   L38923
KEYWORDS    L1 gene.
SOURCE      Human papillomavirus unidentified type (RTRX6) DNA.
REFERENCE   1  (bases 1 to 267)
  AUTHORS   Berkhout,R.J., Tieben,L.M., Smits,H.L., Bavinck,J.N., Vermeer,B.J.
            and ter Schegget,J.
  TITLE     Nested PCR approach for detection and typing of epidermodysplasia
            verruciformis-associated human papillomavirus types in cutaneous
            cancers from renal transplant recipients
  JOURNAL   J. Clin. Microbiol. 33 (3), 690-695 (1995)
  MEDLINE   95270695
COMMENT     HPVRTRX6 was isolated from a renal transplant patient [1],
            and is likely to constitute a new HPV type.  It was isolated
            once during a search for HPV DNA in 53 squamous cell carcinomas
            from 26 renal transplant patients.  Although clearly more
            closely related to EV-associated types than to other HPV types,
            HPVRTRX6 constitutes a relatively isolated sequence.
            NCBI gi: 623471
BASE COUNT       89 a     46 c     52 g     80 t
ORIGIN
        1 atttgtgtac cttcagatgc aggtgctgta actgagtatg attctagcaa atttagagaa
  L1 cds ->
       61 tttttaaggc acgtggaaga gtatcaaata tctgtaatat tacaactgtg taaagtatca
      121 ctgcaacctg atgtgctagc ccagatcaat gcaatgaatt caggtatatt agaagattgg
      181 cagttaggat ttgtaccaac tcctgacaat gcagtacatg acacctatag atttataaat
      241 tcctcagcca ctaaatgtcc agataag
                              L1 cds ->
//
```

```
LOCUS       HPVTogawa          395 bp ds-DNA            VRL         10-DEC-1994
DEFINITION  Human papillomavirus L1 gene fragment.
ACCESSION   L38388
SOURCE      Human papillomavirus DNA.
REFERENCE   1  (bases 1 to 395)
  AUTHORS   Togawa,K. and Rustgi,A.K.
  TITLE     A novel human papillomavirus sequence based on L1 general
            primers
  JOURNAL   Virus Res. 36, 293-297 (1995)
COMMENT     Togawa et al (Gastroenterology 107:128-36, 1994) identified
      an HPV strain with a novel restriction digest pattern in
      14% of squamous cell carcinomas.  Sequence of a fragment
      of the L1 gene indicates that this sequence is related to
      HPV-types associated with EV.  The most closely related
      known HPV types appear to be HPVs 22, 23 and 38.
      The partial L1 sequences for HPVVS42 and HPVVS73 do not
      overlap with that of the Togawa isolate; from their similarity
      to HPVs 22, 23 and 38, it is possible that the Togawa strain
      and one of the other two isolates are from the same HPV type.
          NCBI gi: 598358
BASE COUNT     137 a     60 c     76 g    122 t
ORIGIN
        1 gatactacac gcagtactaa ttttagtatc agtgttaaaa atgaggatag cttagcaaat
       61 tataatgcta gaaatattag agaatacatg agacatgttg aggagtatca gttgtctttt
      121 atattacaat tgtgcagaat acctttaaag gctgaggttt taacacaaat caatgcaatg
      181 aactctgata ttttagagaa ttggcaattg ggctttgtac ctacaccaga taatgcagta
      241 cacgatacat ataggtattt agcctcaaaa gccactaagt gtcctgatgc agtacctgaa
      301 acccaaaaag aagatccttt tggaaagtat tcattctgga atgttgatat gacagaaaaa
      361 ttgtcgctag atttagatca gtttcctttt ggacg
//
```

HPV75L1

```
LOCUS       HPV75L1      661 bp    DNA            VRL       15-NOV-1994
DEFINITION  Human Papillomavirus type 75 ds-DNA, part of L1 ORF.
ACCESSION   X79942
KEYWORDS    capsid protein; L1 gene.
SOURCE      Human papillomavirus.
REFERENCE   1  (bases 1 to 661)
  AUTHORS   Shamanin,V., Glover,M., Rausch,C., Proby,C., Leigh,I.M., zur
            Hausen,H. and de Villiers,E.M.
  TITLE     Specific types of human papillomavirus found in benign
            proliferations and carcinomas of the skin in immunosuppressed
            patients
  JOURNAL   Cancer Res. 54 (17), 4610-4613 (1994)
  MEDLINE   94340583
REFERENCE   2  (bases 1 to 661)
  AUTHORS   Shamanin,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-1994) to the EMBL/GenBank/DDBJ databases. V.A.
            Shamanin, Deutsches Krebsforschungszentrum, Im Neuenheimre Feld
            242, 69120 Heidelberg, FRG
COMMENT     HPV-75 (VS40) was isolated from a dysplastic art biopsy of a renal
            allograft patient, and is likely to constitute a new type.  It
            was found in a search for HPV DNA in 118 skin lesions taken from
            46 renal allograft patients [1].  HPV-49 appears to be the most
            closely related known type.
            NCBI gi: 562314
BASE COUNT      195 a    118 c     155 g     193 t
ORIGIN
        1 tctagggggc aacccttggg ggtaggttct acaggccatc ctttgttcaa taaagtaaag
  L1 cds ->
       61 gatactgaaa attcaaataa ttatataaca atgtctaaag atgataggca ggacacctcg
      121 tttgacccta agcaggttca aatgtttatt attggctgtg caccttgtat aggggagcac
      181 tgggatgctg ccaaaccctg tgacgctgac aaaggagacg gtaaatgtcc acctttagaa
      241 ttagtaaata cagttattga ggatggggat atggtggata taggtttttgg taacataaat
      301 aataaaacct tgtcagcaaa taaatcagat gtcagtttgg atatagttaa taacatttgt
      361 aagtatccag acttccttaa aatggccaat gacatatatg gggactcctg ttttttttat
      421 gccaggcggg aacaatgtta tgctagacac ttttttgtta ggggaggcaa tgtaggcgat
      481 cgaattccta atgctgcagt gggtcaggac aataatttta tgttacctgc agccgctggg
      541 caggctcaaa acactttggg caactctatt tatgttccca cggtcagtgg ttctttggtg
      601 tccacagatg ctcaattatt taacaggcca ttttggctgc aacgagcaca aggtcacaac
      661 a
    L1 cds ->
//
```

```
LOCUS       HPV76L1     668 bp    DNA              VRL        15-NOV-1994
DEFINITION  Human Papillomavirus type 76 DNA, part of L1 ORF.
ACCESSION   X79948
KEYWORDS    capsid protein; L1 gene.
SOURCE      Human papillomavirus.
REFERENCE   1  (bases 1 to 668)
  AUTHORS   Shamanin,V., Glover,M., Rausch,C., Proby,C., Leigh,I.M., zur
            Hausen,H. and de Villiers,E.M.
  TITLE     Specific types of human papillomavirus found in benign
            proliferations and carcinomas of the skin in immunosuppressed
            patients
  JOURNAL   Cancer Res. 54 (17), 4610-4613 (1994)
  MEDLINE   94340583
REFERENCE   2  (bases 1 to 668)
  AUTHORS   Shamanin,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-1994) to the EMBL/GenBank/DDBJ databases. V.A.
            Shamanin, Deutsches Krebsforschungszentrum, Im Neuenheimre Feld
            242, 69120 Heidelberg, FRG
COMMENT     HPV-76 (CR148) was isolated from a skin wart biopsy of a renal
            allograft patient, and is likely to constitute a new type.
            It was found in a search for HPV DNA in 118 skin lesions
            taken from 46 renal allograft patients [1].  HPV-49
            appears to be the most closely related known type.
            NCBI gi: 562308
BASE COUNT      203 a    110 c    158 g    197 t
ORIGIN
        1 tcaagaggac acccattagg agtagggtct acaggtcatc ccctatttaa taaagtgaag
  L1 cds ->
       61 gatacggaaa atgctaataa ttatatagta acatctaagg atgataggca ggacacctca
      121 tttgatccta aacaggttca aatgtttatt attggctgcg caccgtgcat aggtgagcac
      181 tgggatgcag ccaagccctg tgatgctgac agaggggtag gcaaatgtcc acctttggaa
      241 ctggtaaata ctgtaataga agatggagat atggtggata taggtttttgg aaaatataaat
      301 aataaaaccc tgtcagcaaa taagtcagat gtcagtttag atatagttaa taatatttgt
      361 aagtatccag acttttttaaa aatggccaat gacatatatg gagactcctg ttttttttat
      421 gctagacggg agcaatgtta tgctagacat tttttttgtta gaggaggtaa tgtaggagat
      481 gctattcctg atgctgcagt gggtcaggac aataactttg tgttgcctgc agctgttgga
      541 caggcccaaa acactttggg tagctctatt tacgtgccta ccgttagtgg ttctttggta
      601 tccacagatg cacaattatt taataggccc ttttggctac aacgagcaca gggtcataat
      661 aacggtat
          L1 cds ->
//
```

```
LOCUS       HPVVS20L1      677 bp    DNA              VRL         15-NOV-1994
DEFINITION  Human Papillomavirus vs20-4 DNA, part of L1 ORF.
ACCESSION   X79941
KEYWORDS    capsid protein; L1 gene.
SOURCE      Human papillomavirus.
REFERENCE   1  (bases 1 to 677)
  AUTHORS   Shamanin,V., Glover,M., Rausch,C., Proby,C., Leigh,I.M., zur
            Hausen,H. and de Villiers,E.M.
  TITLE     Specific types of human papillomavirus found in benign
            proliferations and carcinomas of the skin in immunosuppressed
            patients
  JOURNAL   Cancer Res. 54 (17), 4610-4613 (1994)
  MEDLINE   94340583
REFERENCE   2  (bases 1 to 677)
  AUTHORS   Shamanin,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-1994) to the EMBL/GenBank/DDBJ databases. V.A.
            Shamanin, Deutsches Krebsforschungszentrum, Im Neuenheimre Feld
            242, 69120 Heidelberg, FRG
COMMENT     HPVVS20 was isolated from a skin wart biopsy of a renal allograft
       patient.  It was found in a search for HPV DNA in 118 skin
       lesions taken from 46 renal allograft patients [1].  HPV-24
       appears to be the most closely related known type.
            NCBI gi: 562312
BASE COUNT      225 a    115 c    145 g    192 t
ORIGIN
        1 ggaagtggac aaccattagg cataggcagc agtggtcacc ctctgtttaa caaggtaaat
  L1 cds ->
       61 gatacagaaa atggcaacac atataaaggg acaactaaag atgatagaca aaacatttca
      121 tttgatccta aacaattaca gatgtttata attggctgta caccatgtat tggtgaacat
      181 tgggataagg ctcctgcatg tgttaatgat attcaacaag gtagttgccc accaatagaa
      241 ttagttaaca catacataca gggtggagat atggctgata taggatatgg caatctaaat
      301 tttaaagctt tacagcaaaa tagatcagat gttagcttgg atattgtaga tgaaatatgc
      361 aaatatcctg actttttacg aatgcaaaat gatgtatatg gcgatgcctg ttttttttat
      421 gctcgacggg agcaatgtta tgccaggcac ttttttgtgc gtggtggcaa acctggtgat
      481 gatatacctg gtgcccaaat tgatgcaggg tcacataaaa atgaatatta catacaggca
      541 gcttcagacc aatcacaaaa tagtttgggg aattctatgt atttcccaac tatcagtggc
      601 tcattagttt caagtgatgc tcaattattt aataggccct tctggctaca gcgagcacaa
      661 ggccaaaaca acgggat
                   L1 cds ->
  //
```

```
LOCUS       HPVVS42L1      665 bp    DNA            VRL        15-NOV-1994
DEFINITION  Human Papillomavirus vs42-1 DNA, part of L1 ORF.
ACCESSION   X79943
KEYWORDS    capsid protein; L1 gene.
SOURCE      Human papillomavirus.
REFERENCE   1  (bases 1 to 665)
  AUTHORS   Shamanin,V., Glover,M., Rausch,C., Proby,C., Leigh,I.M., zur
            Hausen,H. and de Villiers,E.M.
  TITLE     Specific types of human papillomavirus found in benign
            proliferations and carcinomas of the skin in immunosuppressed
            patients
  JOURNAL   Cancer Res. 54 (17), 4610-4613 (1994)
  MEDLINE   94340583
REFERENCE   2  (bases 1 to 665)
  AUTHORS   Shamanin,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-1994) to the EMBL/GenBank/DDBJ databases. V.A.
            Shamanin, Deutsches Krebsforschungszentrum, Im Neuenheimre Feld
            242, 69120 Heidelberg, FRG
COMMENT     HPVVS42 was isolated from a verrucous biopsy of a renal allograft
    patient.  It was found in a search for HPV DNA in 118 skin
    lesions taken from 46 renal allograft patients [1].  HPV-22
    HPV-23 and HPV-38 appear to be the most closely related known
    types.  The partial L1 sequences for the Togawa sequence
    and HPVRTRX1 do not overlap with that of HPVVS42; from their
    similarity to HPVs 22, 23 and 38, it is possible that HPVVS42
    and one of the other two isolates are from the same HPV type.
            NCBI gi: 562316
BASE COUNT      206 a    104 c    146 g    209 t
ORIGIN
        1 ggtaggggc aaccatttgg ggtaggcact acaggtcatc cattatttaa caaattacgt
  L1 cds ->
       61 gatgcagaaa attccagcga acgtcaggga gatactgctg cagatgacag aatgaatata
      121 tcttttgatc ctaagcaggt acaaatgttc ataataggtt gcacaccgtg tttaggtgaa
      181 tattgggatc aagcgcctgt atgtaaagat gcaggtaacc aaatgggctt atgtcctcct
      241 cttgaactaa agaatagtgt catagaagat ggagatatgt ttgatatagg ctttggtaac
      301 attaataata agacactgtc attcaataga tcagatgtta gtttagatat tgtaaatgaa
      361 atatgcaaat atccagattt tttaacaatg tccaatgatg tttatggtga ctcctgtttt
      421 ttttgtgctc gaagagagca atgttatgct agacattatt ttgtacgagg cggtgttgtt
      481 ggagattcta taccagacgg tgcagtccag cagagtaaca aatattattt agcttcagct
      541 caaaataata gcttggaaaa ttctacctat ttcccaactg taagtggttc tttagtgact
      601 tctgatgctc agctatttaa cagacccttt tggttaaagc gtgctcaagg gcataataat
      661 ggaat
        L1 cds ->
//
```

```
LOCUS       HPVS73L1      612 bp    DNA              VRL        21-NOV-1994
DEFINITION  Human Papillomavirus vs73-1 DNA, part of L1 ORF.
ACCESSION   X79944
KEYWORDS    capsid protein; L1 gene.
SOURCE      Human papillomavirus.
REFERENCE   1  (bases 1 to 612)
  AUTHORS   Shamanin,V., Glover,M., Rausch,C., Proby,C., Leigh,I.M., zur
            Hausen,H. and de Villiers,E.M.
  TITLE     Specific types of human papillomavirus found in benign
            proliferations and carcinomas of the skin in immunosuppressed
            patients
  JOURNAL   Cancer Res. 54 (17), 4610-4613 (1994)
  MEDLINE   94340583
REFERENCE   2  (bases 1 to 612)
  AUTHORS   Shamanin,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-1994) to the EMBL/GenBank/DDBJ databases. V.A.
            Shamanin, Deutsches Krebsforschungszentrum, Im Neuenheimre Feld
            242, 69120 Heidelberg, FRG
COMMENT     HPVVS73 was isolated from a skin wart biopsy of a renal allograft
       patient.  It was found in a search for HPV DNA in 118 skin
       lesions taken from 46 renal allograft patients [1].  HPV-22
       HPV-23 and HPV-38 appear to be the most closely related known
       types.  The partial L1 sequences for the Togawa sequence
       and HPVRTRX1 do not overlap with that of HPVVS73; from their
       similarity to HPVs 22, 23 and 38, it is possible that HPVVS73
       and one of the other two isolates are from the same type.
            NCBI gi: 562307
BASE COUNT      193 a     98 c     131 g     190 t
ORIGIN
        1 taaacttaga gattcagaaa attctgcaga acgtctggaa ggaacaagtg atgataggag
  L1 cds ->
       61 gaatatatca tttgatccta agcaagtgca aatgtttgtg ataggctgca ccccctgttt
      121 agggaggtat tgggatacag ctccagtatg taaagatgca ggaagtcaat taggcctttg
      181 ccctccatta gaattaaaaa acagtgttat agaagatggc gatatgtttg atataggatt
      241 tggcaatatt aacaacaaaa cattaagttt taataagtca gatgttagtg tggacattgt
      301 taatgaaatt tgtaaatatc ctgatttttt aactatgtcc aatgatgttt atggagactc
      361 ttgcttttttc tttgctcgca gagagcgatg ttatgcaagg cattattttg tacgcggagg
      421 ggcagtgggt gatttaatac cagatgctac agttaatcag gaccataaat attacttacc
      481 agcaaatcca cctgccacat tggaaaactc tacatacttt ccgactgcta gtggctcctt
      541 agtgacatct gatgcacaat tatttaatag gcccttttgg ttaaaacgtg cacaaggtca
      601 taataatggt at
            L1 cds ->
//
```

```
LOCUS       HPVVS75L1      686 bp    DNA              VRL        15-NOV-1994
DEFINITION  Human Papillomavirus vs75-3 DNA, part of L1 ORF.
ACCESSION   X79945
KEYWORDS    capsid protein; L1 gene.
SOURCE      Human papillomavirus.
REFERENCE   1  (bases 1 to 686)
  AUTHORS   Shamanin,V., Glover,M., Rausch,C., Proby,C., Leigh,I.M., zur
            Hausen,H. and de Villiers,E.M.
  TITLE     Specific types of human papillomavirus found in benign
            proliferations and carcinomas of the skin in immunosuppressed
            patients
  JOURNAL   Cancer Res. 54 (17), 4610-4613 (1994)
  MEDLINE   94340583
REFERENCE   2  (bases 1 to 686)
  AUTHORS   Shamanin,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-1994) to the EMBL/GenBank/DDBJ databases. V.A.
            Shamanin, Deutsches Krebsforschungszentrum, Im Neuenheimre Feld
            242, 69120 Heidelberg, FRG
COMMENT     HPVSV75 was isolated from a skin wart biopsy of a renal allograft
      patient.  It was found in a search for HPV DNA in 118 skin
      lesions taken from 46 renal allograft patients [1].  HPV-24
      appears to be the most closely related known type.
            NCBI gi: 562318
BASE COUNT      208 a    122 c    164 g    192 t
ORIGIN
        1 gggagaggac agccattagg cgttggtacc agtggacatc cactgtttaa caaagttaat
  L1 cds ->
       61 gatgccgaaa atcccttagc ttacagggca caggcctttt ctactgatga taggcaaaac
      121 acatcctttg atcctaaaca aatacaaatg tttataatag gttgtgcacc ctgtattgga
      181 gagcattggg atgtaggtga acgttgtgca ggagccaata atgaaaatgg tcgatgcccc
      241 cctattaaat tggtaaattc agtcatccaa gatggagata tggcagatat tggttatgga
      301 aacctaaatt tccgtacctt acaggaaaac agatctgatg taagtttaga tatagtgaat
      361 gaaacctgta aatatccaga cttttttaaag atgcagaatg atatatatgg cgattcttgc
      421 ttttttctttg ctcgccggga gcaatgttat gcaagacatt ttttgttcg tggggggtaag
      481 gcgggggatg acattcctgg tgcgcaaatc gatgcaggta catataaaaa tgattttttac
      541 atacctggag cgtcaggtca gacacaaaag aatataggta actcgatgta tttcccaaca
      601 gtaagtggct cattggtgtc tagtgatgct caattgttta ataggccctt ctggctccaa
      661 cgggcgcagg ggcaaaacaa cggaat
                           L1 cds ->
//
```

```
LOCUS       HPVVS92L1     674 bp    DNA            VRL       15-NOV-1994
DEFINITION  Human Papillomavirus vs92-1 DNA, part of L1 ORF.
ACCESSION   X79949
KEYWORDS    capsid protein; L1 gene.
SOURCE      Human papillomavirus.
REFERENCE   1  (bases 1 to 674)
  AUTHORS   Shamanin,V., Glover,M., Rausch,C., Proby,C., Leigh,I.M., zur
            Hausen,H. and de Villiers,E.M.
  TITLE     Specific types of human papillomavirus found in benign
            proliferations and carcinomas of the skin in immunosuppressed
            patients
  JOURNAL   Cancer Res. 54 (17), 4610-4613 (1994)
  MEDLINE   94340583
REFERENCE   2  (bases 1 to 674)
  AUTHORS   Shamanin,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-1994) to the EMBL/GenBank/DDBJ databases. V.A.
            Shamanin, Deutsches Krebsforschungszentrum, Im Neuenheimre Feld
            242, 69120 Heidelberg, FRG
COMMENT     HPVVS92 was isolated from a skin wart biopsy of a renal allograft
     patient.  It was found in a search for HPV DNA in 118 skin
     lesions taken from 46 renal allograft patients [1].  HPV-9
     appears to be the most closely related known type.
            The partial L1 sequences for HPVRTRX3 does not overlap with that
     of HPVVS92; from their similarity to HPV-9, it is possible
     that HPVVS73 and one of the other two isolates are from the same
     HPV type.
            NCBI gi: 562320
BASE COUNT      223 a    110 c    141 g    200 t
ORIGIN
        1 tcaaggggac agccattggg tgtaggaaca tcaggtcatc ctttatttaa caaagtcagg
  L1 cds ->
       61 gatactgaaa actcaggtaa ctatcaagca gtttctcagg atgacagaca aaatacatct
      121 tttgatccta aacaagtgca aatgtttgtc attggctgtg tgccgtgtat gggtgaacat
      181 tgggacaaag ctaaggtttg tgaatcagaa gcaaataatc aacaaggctt atgtccaccc
      241 atagagttaa aaaattcagt aattgaagat ggagatatgt ttgatatagg ctttggaaat
      301 attaataaca aagcactatc ttataacaag tcagatgtta gtttagatat agttaatgaa
      361 gtgtgcaaat atccagactt tttaaccatg gctaatgatg tgtatggaga tgcttgtttt
      421 ttctttgcta gacgagaaca atgttatgcc agacattatt ttgttagggg aggcaatgtt
      481 ggcgatgcaa tccctgatgg agcagtacaa caggatcaca actattattt acctgcacaa
      541 aatgcacagc aacaacacac cttgggaaat tctatatatt atccaactgt tagtgggtct
      601 cttgtaacat ctgatgctca gttatttaat agaccatttt ggttacaacg tgctcaagga
      661 caaaacaacg gtat
                L1 cds ->
//
```

```
LOCUS       HPVVS102L1    674 bp    DNA           VRL       15-NOV-1994
DEFINITION  Human Papillomavirus vs102-4 DNA, part of L1 ORF.
ACCESSION   X79946
KEYWORDS    capsid protein; L1 gene.
SOURCE      Human papillomavirus.
REFERENCE   1  (bases 1 to 674)
  AUTHORS   Shamanin,V., Glover,M., Rausch,C., Proby,C., Leigh,I.M., zur
            Hausen,H. and de Villiers,E.M.
  TITLE     Specific types of human papillomavirus found in benign
            proliferations and carcinomas of the skin in immunosuppressed
            patients
  JOURNAL   Cancer Res. 54 (17), 4610-4613 (1994)
  MEDLINE   94340583
REFERENCE   2  (bases 1 to 674)
  AUTHORS   Shamanin,V.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-1994) to the EMBL/GenBank/DDBJ databases. V.A.
            Shamanin, Deutsches Krebsforschungszentrum, Im Neuenheimre Feld
            242, 69120 Heidelberg, FRG
COMMENT     HPVVS102 was isolated from a dysplastic wart biopsy of a renal
         allograft patient.  It was found in a search for HPV DNA in 118
         skin lesions taken from 46 renal allograft patients [1].  HPV-15
         appears to be the most closely related known type.
            NCBI gi: 562310
BASE COUNT      223 a    102 c     148 g     201 t
ORIGIN
        1 ggaagaggtc tccatttggg tgtaggtaca gcaggccatc cactattcaa taaagttaga
  L1 cds ->
       61 gatacagaaa ataatagtgg ctatcaagat acgtctacgg atgacagaca aaatacatca
      121 tttgatccaa aacaagttca aatgtttgta gtaggatgtg ctccttgttt gggagaacat
      181 tgggataaag ctcctgtctg tgactcagat aaaaataacc aggctggaaa atgccctcca
      241 ttagaactga gaaacacagt aatagaagat ggagatatga ttgatatagg ctttggcaat
      301 ataaacaaca aggttttatc agttactaag tcagatgtta gtctggatat agttaatgaa
      361 acttgtaagt atccagattt tttaactatg gccaatgatg tatatggtga ctcttgtttt
      421 ttctttgcaa ggagagaaca gtgttatgct agacattatt atgttagggg aggtgtagta
      481 ggtgatgcta ttcctgatga agctgtgaat caagataaaa actttgtgtt acctgcacaa
      541 ggcactcagc aacaaaagga tatagctagt tctatatatt ttccaactgt tagtggttcc
      601 ttagtaactt ctgatgctca attatttaac agaccatttt ggttacgcag agcacaaggg
      661 caaaataacg ggat
                 L1 cds ->
//
```

Group B2 Sequences

HPV4 **HPV48**
**HPV50** **HPV60**
HPV65

INTRODUCTION

Group B2 is equivalent to part of the old group G. It consists of the human papillomaviruses HPV-4, HPV-48, HPV-50, HPV-60, and HPV-65, a group primarily associated with the benign cutaneous lesions, commonly seen in the general population. HPV-50 is the only member of this group associated with EV lesions. A number of apparent new types belonging to this group have recently been isolated [1], but no published information is currently available.

HPV-4, in conjunction with HPV-1 (group E1) and HPV-2, (group A4) are the major etiological agents of benign cutaneous papillomas in the general population. HPV-1 is primarily associated with deep palmo-plantar warts, while HPV-4 has been correlated with common warts and keratotic flat lesions on the hands and feet [2,3]. HPV-4 is also frequently present in hand warts of meat handlers [4]. HPV-65 has been linked to pigmented common warts and keratotic flat lesions [3].

HPV-4 and HPV-65 produce homogeneous intracytoplasmic inclusion bodies in most infected epidermal cells. The inclusion bodies primarily contain E4 proteins that can be used to histologically identify these viruses.

The primary target tissue of the Group B2 viruses is the epithelium, however rare mucosal infection has been reported for HPV-4, which has been identified in isolated cases of both normal and malignant oral lesions [5].
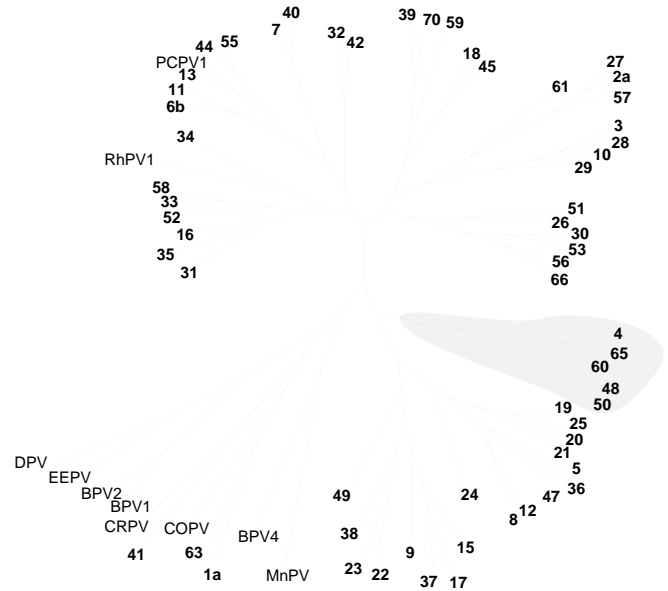
What's new?

The complete genomes of HPV-48, HPV-50, and HPV-60 have been released since last year's publication and are given on the following pages. The sequences of other members of this group were published in *Human Papillomaviruses 1994* pp. I-G-8, and I-G-2.

References

[1] Shamanin,V.A., and de Villiers,E.M. Unpublished data presented at the 14th International Papillomavirus Conference, Quebec City, July 23–28, 1995. Abstract title: HPV in non-melanoma skin cancers.

[2] Danos,O., Katinka,M., and Yaniv,M. Human papillomavirus 1a complete DNA sequence: a novel type of genome organization among Papovaviridae. *EMBO* **1**: 231–236 (1982)

[3] Egawa, K., Delius,H., Matsukura,T., Kawashima,M., and de Villiers,E.M. Two novel types of human papillomavirus, HPV 63 and HPV 65: comparisons of their clinical and histological features and DNA sequences to other HPV types. *Virology* **194**: 789–99 (1993)

[4] Melchers,W., de Mare,S., Kuitert,E., Galama,J., Walboomers,J., van den Brule,A.J. Human papillomavirus and cutaneous warts in meat handlers. *J Clin Microbiol* **31**: 2547–9 (1993)

[5] Yeudall,W.A., and Campo,M.S. Human papillomavirus DNA in biopsies of oral tissues. *J Gen Virol* **72**: 173–6 (1991)

```
LOCUS       HPV48          7100 bp    DNA             VRL        18-JUL-1995
DEFINITION  Human papillomavirus type 48, complete genome.
ACCESSION   U31789
KEYWORDS    .
SOURCE      Human papillomavirus type 48.
REFERENCE   1  (bases 1 to 7100)
  AUTHORS   Delius,H.
  JOURNAL   Unpublished, Sequenced by Hajo Delius, Deutsches
            Deutsches Krebsforschungzentrum, Angewandte Tumorvirologie,
            I.N.F. 506, W-6900 Heidelberg, Germany
REFERENCE   2  (sites)
  AUTHORS   Muller,M., Kelly,G., Fiedler,M. and Gissmann,L.
  TITLE     Human papillomavirus type 48
  JOURNAL   Journal of Virology 63 (11), 4907-4908 (1989)
REFERENCE   3  (bases 1 to 7100)
  AUTHORS   Farmer,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     HPV-48 DNA was isolated from a squamous cell carcinoma of the hand
            of an immunosuppressed 36-year old female patient; the DNA was
            present in episomal form. HPV 48 was not found in any of the 21
            similar tumors subsequently tested. Weak hybridization was observed
            under stringent conditions with HPV types both of the EV-associated
            groups and the anogenital-associated groups. HPV 48 is
            phylogenetically associated with HPV 4, 50, 60 and 65, a set of
            primarily cutaneous types.
BASE COUNT     2286 a   1163 c   1447 g   2204 t
ORIGIN      102 bp upstream from beginning of E6 cds
         1 tctgacgcat aataattgtt ggcaacaatc gtgcataata acttaggaac cgagacagta
        61 tcaTATAAAT actgctgaac agtagatttc ttcagaagct gATGgagcca caatttccta
     E6 orf start ->                                 E6 cds ->
     signal ->
       121 ctgatttgga ttcatactgc aaatatttta acataagctt ttttgactta gttcttaagt
       181 gtatcttttg taaattttct gtatctattg ttgatcttgc ttcttttcat aacaaacgac
       241 tgtctgtaat ttggagagat aatacaccat ttgcatgttg tactaaatgt ttaagattaa
       301 ctgcattata tgaaaaagat aatttttttg tctgtactgc aaaaagtcat ttgttaactg
       361 gtttagttaa aaaggaattg tcagatatta atattaggtg ccaacattgt tattcatttt
       421 tagattattt agaaaagttg taccatttat ataatgatgt tgattttctg tTAAttcgcg
                                                              E7 orf start ->
       481 ggacttggag aggtgtttgt agaaattgta tcagccATGa ggggagaTAA ggctactatt
                                           E7 cds ->             <- E6 end
       541 cctgacatag agttagaaga gcttgttttg cctgctaact taatttctga cgagtcattg
       601 tcaccagatg ctactgcaga ggaggagttt gtccgtacc ggatagactc caagtgtcac
       661 aattgtggct gtcgcataag ggtgacagtc gctgcaactg aatttggaat acgttgtttt
       721 gagcaactgt tacTGAaaga gctttgttta ttttgtcctg cgtgctccag acaacttccc
              E1 orf start ->
       781 cgcaATGgca gatcaTAAag gtactgataa tattgatcat aatgatgttt tagatgggtc
        E1 cds ->            <- E7 end
       841 ttggtgttta ataactgaag ctgaatgtga ggacgatacg ttagtggatt gtttgagga
       901 gagcacaaat gactctgtag tttcgaatct gttagacgat tccgaaagta taattcaggg
       961 gaattctgag gaaagtgacc gctgtattca ggagctaaaa agaaagttaa atgttactcc
      1021 tgaaaaacag atctcagagc taagtccgcg gttatcggca gtgcatataa cgcctgagag
      1081 agcatctaag cggagactgt ttaatgatag tggagttgtt gaagatgaag ctgaaagtaa
      1141 tactattcag gtagattcac tgttggtgca aaaagacgcg ggaaatcaaa atggcgctga
      1201 gtgtgaactt aattctattt tgcgtagtaa taatatcaga gctactgtac tttgcaaatt
      1261 taaagataaa tttggggttt catttaatga gcttacaaga agttttaaaa gtgataagac
```

```
1321 atgtacACCA AATTGGGTaa taacagcaat aggtattaga gaagatttac gagatgcttg
E2 binding ->
1381 taaagtttta ttacagcaac atgttgagtt tttagaaatg atttgcaatg attttttctgt
1441 gctactattg gtagaattta aggtaactaa aaatagagaa acggttttaa agttaatgtg
1501 tagtatgtta aatgctaaag aagaacaaat tttgtgtgaa ccgccaaaat tgaaaagtac
1561 agctgcagct ttatatttct ataagaagat tataactgat acatgtttta aatatggcac
1621 tttgccttcc tgggttagta gattaactat agtagaacat cagttagctt cagcagacac
1681 attttcatta tctgaaatgg tacaatgggc ttatgacaat gattttactg aagaagcatc
1741 tgttgcttat aattatgcat gttatgccac agaaaataca aatgcagcag cctttttagc
1801 cagtaatatg caagttaagt atgttaaaga ttgtgtagca atggttagaa tgtataaaag
1861 acaagaaatg aaatcaatga caatgtcaga gtggatttcc aaatgttgta agaagaaac
1921 tattggtgaa gagtggaaag aaattgtaca gttttttaaaa tatcagggag taaacttttt
1981 agaattttta atagctttaa aacaattttt taaatgcact cctaaaaaga tgtgtatagt
2041 aatatatggt ccaccagata ctggaaagtc aatgtttttgt tttaaattag tacaatttt
2101 gaaaggtcaa gttgtatctt atattaacaa atcgagtcag ttttggttaa tgccattaca
2161 agatgctaaa attggtttgt tagatgatgc cacacataat tgttggattt atttagacac
2221 atatttgaga aatgcatttg atggtaacac gttttgttta gatattaaac ataaaaacct
2281 gcaacagaca aaacttccac ctatgataat aactactaat gttaatgtaa caactgatga
2341 atctctattt tatttgagaa gtagacttac gtgtttcaat tttccgaata agttaccaat
2401 gtctgacaaa gatgaacctt gtttaccat ttctgacaaa agctggacct gttttttttag
2461 aaagttttgg aatcaatTAG aacttcaaga agATGcagcc agagacccag gagagcctga
            E2 orf start ->    E2 cds ->
2521 gcacccgttt tgctgcacag caagaaattc agttgacttt gatTGAaaaa gaatcttatg
                                                    <- E1 end
2581 atcttaaaga tcatttagca tactggaaag ctgtaagatt agaaaatgtt attgcctatt
2641 atgctagaaa agaacatatt actaaattag gtctgcagcc attaccaaca ttagctgtaa
2701 ctgaatataa agcaaaagaa gcaattaata tacagctgct tattcaaagt ttattgaaat
2761 cagagtttgc tttggaaaga tggactcttg cagaaactag tgcagaaact ataaacagtt
2821 cccccagaaa ttgtttcaaa aaagtacctt tcattgtaaa tgtatggttt gataatgatg
2881 aaagaaattc ttttccttat acctgttggg attttatata ttatcaagat gaccaaaaca
2941 aatggcacaa gactgaaggg ctggttgatc ataatggatg ttattatgta gatctaaatg
3001 gtgattttgt atactttact ttatttcaac ctgatgctgT AAatatgga aaaactggac
                                  E4 orf start ->
                                  NH2 terminus unknown
3061 tatggacagt tagatttaaa aacaaaacta tttctgcctc tgttactagc tcctcgagaa
3121 atacaaatcc ctcttctgaa agcagggtcg ggctctcgac ctccagcagc tcggagagcc
3181 ctcgaaggag accgagcatc tcagaaaact ccaacaccga tcgcccacc tcctcgacat
3241 ccagactacg agagcgacga cgacgagaac cgagagaatc tggaaccacc gacaccacac
3301 ccagaagacg aggaacaaag aggaaattgg ggtccgactc tgcaccaact ccttcggaag
3361 tgggatcaag atctacaacg cttgcaagac acggttactc acgacttgga cgattacagg
3421 aagaagctcg ggatccgcca tTAGtgttat ttacaggtca acaaaataat ttgaaatgtt
                            <- E4 end
3481 ggagaaatcg ctgtactaca aaatatgcta gtttatttt atgcttcagc tctgtttgga
3541 agtggcttgg tcctaattct gatggggggtg ctgcaaagt gttagttgca tttaaaagtg
3601 atgcccaaag acaagtgttc ttaaacacag ttcatattcc taaaggaact actataactc
3661 tgggaagact TGAcagttta TAATGtcctt acgtagaaga aaaagagcaa gtcctactga
        L2 orf start ->        <- E2 end
                        L2 cds ->
3721 tctttataaa acatgcttgc aaggggggga ctgcattcct gatgttaaaa ataaatttga
3781 aaattctact attgcagatt ggttATTAAA aatatttgga agtttggtat attttggaaa
                        signal ->
3841 tttgggaata ggatctggaa aagggtctgg gggatcattt ggatatagac cattaggatc
3901 cgcaggaagt ggaagaccag ccacagactt accagtgact agacctaatg ttgtgataga
3961 acctataggt cctcaaagta tagtacccat tgatcctgga gcgtcatcta tagtccctct
4021 tgttgaggga gggcctgata tatcttttat tgcaccagat gcaggtccag gtataggagg
4081 tgaggatatt gaactattca cctttagaga tccagcaact gatgtaggtg tgtagtggg
4141 aggtcctact actatttcta cagaagaaag tgagacagct ataatagatg ctttaccaag
4201 tgccacaact cccaaacagt gtttttatga ttcttataca caaactatct tgcaaacaca
4261 ggtaaaccca tttttaaata atgctattag tgatactaat gtgtttgtag atccattatt
```

```
4321 tgcaggagag acaattgggg acaacatatt tgaagagata cccttgcaga atttaaattt
4381 cagttttccg cgggaaagta cacctgttaa acctgggagg ggtttacgaa caccagctca
4441 aagatcttac agtagattta tggaacagta cccaatccaa gctccggaat ttcttagtca
4501 gccttctcga ttggtgcagt ttgagtttga aaatcccgcc tttgatccag atattagtat
4561 acagttccag cgtgatgtaa atagcctaga ggctgcgcca aatccagctt ttgctgacat
4621 tgcttattta agtaggccgc acatgtctgc tacatcagaa ggattagtca gagtcagcag
4681 aattggatct cgagcagtat tacaaaccag gagtggatta actataggcc ctaaagtaca
4741 ttattatatg gatttgtctg caatatcaac agaagctata gaattgcaaa cctttgcaga
4801 ttctggacat gttcacacaa tagttgatga tttcttatca gttactgctt tagatgatcc
4861 agcaaatata gctgatataa attatacaga agatgattta ttagatcctt tacttgaaaa
4921 ctttaataat tcacatatta cagtacaagg ggttgatgag gaaggagaaa cagttgctct
4981 tccaattcct tcaattacta attcttccaa aactttttgtt acagacattg cagaaaatgg
5041 tctgtttgca aatgatacag atagtctttt aaccccagca agcactattg tacctgctat
5101 taattggttc cctttatttg atagttactc agactttgct ttagatccct tttttattcc
5161 acgtaagaaa cgacgctTAG atatccttTA Atgttttca gATGgctctg tggtctgcag
                L1 orf start ->        <- L2 end
                                             L1 cds ->
5221 tccctggcaa ggtttaccta cctcctagtg caccagtagc tcgtgtgttg cgcactgatg
5281 aatacgtgca agaaacggat gtatatttt atacaagcac tgaaagactt ttaatagttg
5341 gtaatcctta ttttgatgtt gaaaacagag acactataac agtacctaaa gtttctgcta
5401 atcaataccg agtatttaga tgtaaacttc cagatcctaa caagtttgct ctagttgata
5461 aaaatttata taattcagat aaggaacgtt tagtttggaa acttgttggt ttggaagtag
5521 gcagaggggg tcctcttggc gtaggatcta caggtcatcc tttactaaat aagatagggg
5581 acacagaaaa tcctagcttt tatttaggag aacaaacaaa ggatgagaga caaaatgtttt
5641 ccatggatcc taagcagagt cagatactaa tagttggttg tgctccagcc actggagaat
5701 actgggattt agctaaacca tgcaatgatt tggaaaacgg ggctgcacca ccaattcaac
5761 tagttaatac tgttattcag gatggtgata tgggagatat aggctttggg gctgctaatt
5821 ttcctaagtt gatgcaggat cgtgctggcg tccctctaga gttaatagat tctattagta
5881 tatggccaga tttttttaaag atgaccaaag atatttatgg agactctgtc ttcttttttg
5941 gtaaacggga caatgttat gctcgtcatc tttttgcaag agctggccaa atggggggagc
6001 ctataccaac agaaaatgga gtatattata taactcctga ttctgccgat caaaacaaca
6061 gatcttctca tttaggatct tctgtgtatt ttacaacacc aagtggatcc ttgaatacta
6121 gcgattctca gttatttaat aggccttatt ggcttagacg agcccaggga acaaataatg
6181 gcatttgttg gggaaatgag ctgtttatta cagtttttga taatactcat aatgtgaatt
6241 ttactattag tgttaagaat gataaaactg cattaactga aaactacata gataatggtt
6301 acaaatataa taatgcagat tttaaacaat atcttcgaca tacagaagaa tatgaaaatag
6361 agttggtttt tcagttatgc aaagtgaact tgactgcaga tgttctggca catttacacg
6421 ttatgaatcc cagaatatta gaggagtggc aattagcttt tgttccacca gctccaacgg
6481 gcattgaaga tacttacagg tatatcaaat ccatggctac taaatgccct actgctgaac
6541 cagaagaaga cacagatcca tacaaagctt acagtttctg gactttagat atgacagagc
6601 gcttctcgtc tgatttaagt cagtttcct tgggtcgaaa attttatat caaactggtt
6661 tgttaaatgg taaacgagct agaacagact atacagctgc aggatctagt accagatcta
6721 caaagcgtag gagagtaaga TAGtaaagat gtagaatgtt tacattatta atactgtgaa
                     <- L1 end
6781 tcaaacaaac tgtgatcAAT AAAcactgtg aatattttac cgtggtgtca tttatacgcc
                  signal ->
6841 tccattgggt gggttctgac ctataaactc ctgggtcagt catttggtca gcagacgctg
6901 aggagcagca gacgattcag acgtccattt gtgagacgaa gACCGTTTTC GGTaagtttc
                                        E2 binding ->
6961 atagccgcca ggtattgttt gtaccgatac agttacacct taaatgtac tgccaacttt
7021 tacaagttca gacagagatc aacaagatat cttaaaacag tACCTAAAGC GGTactgtAC
                                        E2 binding ->      E2 binding ->
7081 CGTTTGTGGT atgttgatgc
```

//

```
LOCUS       HPV50          7184 bp    DNA              VRL        17-JUL-1999
DEFINITION  Human papillomavirus type 50, complete genome.
ACCESSION   U31790
KEYWORDS    .
SOURCE      Human papillomavirus type 50.
REFERENCE   1  (bases 1 to 7184)
  AUTHORS   Delius,H.
  JOURNAL   Unpublished, Sequenced by Hajo Delius, Deutsches
            Deutsches Krebsforschungzentrum, Angewandte Tumorvirologie,
            I.N.F. 506, W-6900 Heidelberg, Germany
REFERENCE   2  (sites)
  AUTHORS   Favre,M., Obalek,S., Jablonska,S. and Orth,G.
  TITLE     Human papillomavirus (HPV) type 50, a type associated with
            epidermodysplasia verruciformis (EV) and only weakly related to
            other EV-specific HPVs
  JOURNAL   Journal of Virology 63 (11), 4910 (1989)
REFERENCE   3  (bases 1 to 7184)
  AUTHORS   Farmer,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     HPV50 DNA was isolated from actinic keratosis taken from a
            epidermodysplasia verruciformis (EV) patient, who had previously
            been found to be infected with HPVs 5, 8, 20, 23, and 36. HPV50 was
            not found in cutaneuous lesions of patients with warts (56 cases),
            keratoacanthoma (12 cases), Bowen's disease (20 cases), actinic
            keratosis (25 cases), basalioma (24 cases), and squamous cell
            carcinoma (22 cases). It was however found in specimens of skin
            macules of 2 out of 34 additional patients suffering from EV.
            Phylogenetically, HPV50 is a member of a group of primarily
            cutaneous types including HPVs 4, 48, 60 and 65. HPV-4 and HPV-65
            are associated with Hg-ICB.  The members of this group other than
            HPV-50 are not specifically associated with EV.
BASE COUNT    2333 a   1205 c   1441 g   2205 t
ORIGIN      102 bp upstream from beginning of E6 cds
        1 tctatctaat aatagttgcc aacaattatt agacatagag acgtgacaAC CGACTACGGT
                                                              E2 binding ->
       61 gcaTATAAAA gagctgatca gcacagattT GAaggagact gATGGagcct cagagagcta
   signal ->                    E6 orf start -> E6 cds ->
      121 aaaatttggt ggactattgc aagcaacaac agattagttt ttttgaatta gagctacagt
      181 gcttgttttg taaatttgtt attactttgc ctgacttagc tagctttcat tgtaaaaagt
      241 tagctcttgt ttatagagat ggtattgctt ttgctgcgtg cgctaaatgc ttaagattgt
      301 ctgctatctt tgaaaatgag agatattata cttgctctat aaaagcttac ttgctatctg
      361 atttgattgg tagacctttg agtgaaatcg caattcgctg tgaaaattgt atgtgtttgt
      421 tggattatat tgaaaagtac gattgtattt gtcacggtgg ttattttcat tTAGttagag
                                                             E7 orf start ->
      481 gtaattggag gggctgttgt agaaattgtt acgaacATGa ggggTGAtaa accaactatt
                                       E7 cds ->        <- E6 end
      541 cctgatattg ttttagaaga gcttgttttg cctgtgaatt tgcttagtga cgaatctatt
      601 gaaactgatg acaTAGcaga atctgaaagg tcaccattca agattgattc tacgtgtaag
             E1 orf start ->
      661 cattgtcatt gtcgcgttag actgtgttgt gtcgccaccg acgcagctat ccgtgtgttt
      721 gagtgccttc ttcaatcaga attttctttt ctgtgtctta agtgctccaa agaattgcta
      781 cgaaATGgca gaactTAAag gtacagataa taattctaat gtagaaatta ttaatgagtt
   E1 cds ->              <- E7 end
      841 gtttgataat gaagctgtct gtgatgatga ttcttttcag gaattgtttg atgaaagtac
      901 ggatgaatct actatctcta atttaataga tgactctgaa aacgttgtgc aggggaattc
      961 ccatgcatta ctaaatgcac agctttcaga ggaatatgac aaagacctgg ttacagtaaa
     1021 acgaaagttt tatgccaccc cggaaaagct cgcctctgat ttgagtccga gactttcagc
     1081 ggtgcatata acaccagaaa gacaatccaa aaggcgttta tttagagaca gtgggattga
```

```
1141 agatgaagct gaaaattcta ttgttcaggt acaggatgat agttcgaatg tggcgggaaa
1201 caaaaatggc gctgatgcgg agttgtattc gttgcttcat agcaataatc gtagagctgc
1261 tttgctttgt aagtttaaag aaaaatatgg cattccattt aatgaaatta ctagaacttt
1321 taaaagtaat aagagttgta cacaaaattg gattattgta gttttgcct gtgcagaaga
1381 tttaattgaa gcttctaaaa ccacgatgca aaatcatgtt tcctatttac aaatgattac
1441 ttctgatttt tcagctttgt atattatttg ctttaaggct gctaagagca gagagactgt
1501 tgttaaattg attaatagct tgctaaatac aaaagaggaa caagtactat gtgatcctcc
1561 taaaataaaa agtatggctg cagcattata ttgctataaa aaggtaatag cagatacctg
1621 ttataaatat ggagactttc ctgactggat agccacacat actgttataa atcatcagct
1681 agcaactgca gacagcttta aatttagtga tatggtacag tgggcatacg ataatgatat
1741 gctggatgaa gctgcaattg cttacaatta tgcgtgctat gcaagtgaaa atgaaaatgc
1801 agctgcattt ttacaaacta atagtcaatt aaagtatgtt aaggaatgtt gtgcaatggt
1861 caggttatac aaaaagcaag aaatgaggaa tatgactatg ccagaatgga ttaaatcatg
1921 tttcacaaat aattataatt ctgatgattg gaaagttatt gtaagatatt aaaatatca
1981 aaatattaac tttttagaat ttttgttagc tttaaagtta cttttaaaag catacctaa
2041 aaaaatgtgc ttagttattt atggaccacc agatacaggc aagtcatact tttgttatca
2101 atttattcag tttatgagag gaaaagttgt atccttcatg aacaaaaata gtcatttttg
2161 gttaatgcct ttattagatt caaaaatagg attttttagat gatgcaactc aatgttgttg
2221 gatgtatttg gacacacaca tgagaaatgc atttgatggt aatgcagttt ctgttgatgt
2281 gaaacataaa aatttacaac aaatagtttt accacctatg ttaataacaa caaattgtga
2341 tgtatgcaga atccaacgt ttatgtactt aaggagtaga ttaacatgtt ttaactttcc
2401 aaataagtta ccattgtatg aaaatggaga accaaaattt aaatttactg ataattgctg
2461 gacgtctttt tttagcaagt tttggaagca ctTAGattta cctgaagATG acccagatgg
                         E2 orf start ->    E2 cds ->
2521 agacacagga aactctgagc gcacgttttc ttgcacaaca agacattcaa ttgaatctga
2581 tTGAaaaaga ttcaaaaaat ttaaaagacc atatagacta ttgggaatct atgagaaaag
         <- E1 end
2641 aacaagtgct tgcattttat gcaaaaaaag aaaatatgag caggcttgga ttacaacctc
2701 ttcctcctgc taaggtttct gagcaaaaag caaaagatgc aattagaatt caattacttt
2761 tgcaatcctt gtataagtct gattttgggt ctgaaccatg gactttatca gaatgtagtt
2821 tagaaatgtt aaatgcacct cctagaaatt gttttaaaaa acaaccgttt actgtaactg
2881 tgcaatttga caatgatcct aaaaatgtgt atccttatat atgttatgaa tatatatatt
2941 atcaagatga tagagacaaa tggcataaag ttaaaggtct agtggatcac aatggccttt
3001 attttaaaga agTGActggt gattcagtat actttaaact ctttcaacct gatgccactg
         E4 orf start ->
         NH2 terminus unknown
3061 tatatggcaa atctggacaa tggactgtta tatttaaaaa caaaactatt cattcctctg
3121 tcactagctc ctcaaggagt gcgtttgggc ctgctgacga acagcccggg ccttccacat
3181 cgtacgacaa gagccaacag gaaagatctg gaagcggtca accaaaagcc ctacaggaca
3241 ccgaaccacc cacctcgaca tcaacagtac gacttcgacg aggacgacga gaaagagaac
3301 accattccta cagacacaga aagtcacaat cagaattggg agccgactct cgcccaactc
3361 ctgaagaagt gggaagaaga tctcacacag ttgcagcgca tggtctgtca agacttagac
3421 gattacaaga ggaagctcgg gatccaccag tccTAAttat tacaggtcaa cagaataatc
                                    <- E4 end
3481 ttaaatgttg gaggtatagg ttctcacaaa aatatgcaga tctctatgag tgttgtagct
3541 ctgcatggaa gtggttgggc ccaaaatcag agggatacag gggtgatgct aaattgctta
3601 ttgctttta aaaatccTGAa caaaggctat cttttttaaa cactgttgga ttacccaaaa
             L2 orf start ->
3661 atactacgta ctctatggga catttggact ctctaTAATG cttcgcagac gcaaacgggc
                                            <- E2 end
                                        L2 cds ->
3721 aagcccctacg gacttatata gatcatgctt acaagggggg gattgcatac ctgatgttca
3781 aAATAAAttt gaaggcaaca caatcgcaga ttggttattg aaaatatttg gaggtttagt
signal ->
3841 atactttggc aatttgggaa ttggaacagg aagaggaact gggggtactt ttggctatag
3901 acctttttgg gcccctgggt ctggaaggcc aactcaggaa ctacctatcg caagaccaaa
3961 tgttgttata gatccattag ggccagcgcc tattgtacca gtagatccct ctgcagcttc
4021 aatagttcca ttagtagaag gcgcccctga tgttggattt gcagctccag atgctggacc
4081 tgcagcaggt ggaactgata ttgaattata tactataaca aattccacaa cagatgttgg
4141 tgctgttggt ggtgggccca ctgttacctc taacgaagaa tttgaagtag ctgttatag a
4201 tgcacaacca attgcaccct atccaaagca attattatat gactctacta tagcagcaac
```

```
4261 ttttgaaaca cagattaatc ccttcataaa tccagacata aataatgtta atgtattagt
4321 tgaccccagc tttgcaggag acactgtagg tgactacttt tatgaagaga tacctttgga
4381 acgtttagat attcaaacct ttgacatttt agagccacct actgaaagca cgcccACCCA
                                                       E2 binding ->
4441 ATTAGGTaat aggtttgttt ctagggcaag agatctgtac agtcgatttg ttgcacagca
4501 gccaatatct gagccagatt tttgagtca accttcccgc cttgttcagt ttgaatatag
4561 aaatcccgcc tttgacccag atgtgagcct atactttgag cgggatttgg aaggcctcag
4621 ggcagcgcca ttgcaggaat ttgctgatgt agtttatttta ggcaggccaa gagtatccag
4681 cACCTCAGAA GGTacaataa gagtaagtag attaggtacc agagcagctt taacaacacg
E2 binding ->
4741 tagtggtctt tctgtgggac ctcaggtaca cttttatatg gatttaagtg atatacctcc
4801 agaggattct atagagttgc atacattaaa cgttacacca caaacaagta caattgtaga
4861 tgatatatta gcaacaacta catttgatga tcctgcaaac tccttattta cacagtttaa
4921 tgaagatgta ttaacagacg atgttgaaca taattttaca gagtctcatt tagtttatacc
4981 agctactgat gaagaaaatg acacagcaat aaatattatt aatttaagaa atattcctct
5041 tactgtgggc atgaattcag gtgacatatc tactacatta tcagattata atattttaga
5101 tgcatccttg atagtaaaat caaatgtttc tgaacagccc cttttgtttt tagattattc
5161 agattatgat ttacatccag gtcttcttcc aaaaagaaga cgcatagatt attttTAAtt
                                                        <- L2 end
                                            L1 orf start ->
5221 tatttttacag ATGgctcatt ggtcctcaac ctctggaaag ttataccttc caccaagtac
            L1 cds ->
5281 tcctgttgcc agagtttttaa gcaccgatga atatgtgaaa gaaactgatg tttactttca
5341 tgcacgcagt gagcgtttac taattgtggg acacccatat tatgatatag aagatggtgg
5401 agatatcaaa gttcctaaag tatcagcaaa tcaatacaga gttttttcgtt gtgaattacc
5461 agatcctaac aagtttgctc taattgatac tacattgtat aactcagata cagaacgttt
5521 ggtttggaaa ctagtaggta tagaagtcgg gagaggtggg cctttaggtg ttggatctac
5581 tggtcatccc ttgtttaata aggttggaga cacagaaaat cctagctttt acttaggtcc
5641 acaagaaaaa gatgaaagac aaaacttgtc tatagatcct aaacaaacac agttattgat
5701 tgtaggctgc aaaccagctg ttggtgaata ttgggattta gcagaacctt gtgacaaaaa
5761 cagcttaaat aatggcaagt gtcctccaat acaactagta aacagttata ttcaggatgg
5821 tgatatggga gatataggat ttggtaatgc aaactttcct aaactacaac aggacagagc
5881 aggggtccct ttagatatag tggattcaat cagcttatgg cctgatctat aaaaatgac
5941 taaagatgtc tatggggacc atgtgttttt ctatgcaaag caagagcaat tatatgccag
6001 acatttgttt acccatgcag gaccaattgg ggaacctata cctaatgtct caggagttta
6061 taattatgca gtaaacccta atcaaccaga gcaaaatcgt agaACCAATA TTGGTcccta
                                               E2 binding ->
6121 cttatatttt actacaccaa gtggatcttt aaatacaagc agttcacagt tatttaatag
6181 accatattgg attcgcagag ctcagggtac gaacaattgt atttgttggg gtaatgaggt
6241 atttgttaca gtatttgata acaccagaaa cattaacttt aatattagtg ttaagaaaga
6301 tgtcaatcct ttggatcctt aaaatgtagc aagttcttat atgtactcaa aggatgattt
6361 taatcagtac agccgtcata ctgaagaata tgaattagaa tttatatttc agttatgtaa
6421 agtaggattg gatgctgata tattagctca tcttaatgta atggaccca gaattttgga
6481 aaattggcaa ttagcttatg ttccaccagc tcccagtggt ataggcgata cttacagata
6541 cttaaaatct gatgctacaa aatgtcctgc aaaaagacagt agcgccgaag tagtagacata
6601 ttataaagaa tatactttttt ggaatgtcaa ccttacagaa aagttttcgt ctgaacttga
6661 tcagtatgcc ctaggacgaa aattttttgtt tcagactgga ttattaaaga gaagagttag
6721 aactgattac actgttgcta cagtttccaa accaaacaag agaaaacgta ccagaTAAgc
                                                        <- L1 end
6781 tgctgacatg ctttcttggc tttactgtga atattgtact gtgaattcta tactgtgaat
6841 gctttagatc tgagtcttga acatttcaca aaatgAATAA AcaAATAAAg ttaatgtaca
                                  signal ->signal ->
6901 ggtgtgtgtt gtctttgagt cagagttatt tactcccgcc ctttgtagta aaatatgtt
6961 gcacctctgt gatcaatcat gttgaccagc tggacgtgct gtacaccggt cactttaaga
7021 gagacacgAC CGTTTTCGGT aagtttgtat gcctttaggg aatataaagc gACCGAAGTC
  E2 binding ->                                         E2 binding ->
7081 ggtacttgct aacaaagata aaatctcacg agttagaaca aacctctgag tcacctcagt
7141 gcagctggca gaagaaagta tACCGATATC GGTtgctatg tttt
                  E2 binding ->
//
```

```
LOCUS       HPV60         7313 bp    DNA             VRL       18-JUL-1995
DEFINITION  Human papillomavirus type 60, complete genome.
ACCESSION   U31792
KEYWORDS    .
SOURCE      Human papillomavirus type 60.
REFERENCE   1  (bases 1 to 7313)
  AUTHORS   Delius,H.
  JOURNAL   Unpublished, Sequenced by Hajo Delius, Deutsches
            Deutsches Krebsforschungzentrum, Angewandte Tumorvirologie,
            I.N.F. 506, W-6900 Heidelberg, Germany
REFERENCE   2  (sites)
  AUTHORS   Matsukura,T., Iwasaki,T. and Kawashima,M.
  TITLE     Molecular cloning of a novel human papillomavirus (type 60) from a
            plantar cyst with characteristic pathological changes
  JOURNAL   Virology 190 (1), 561-564 (1992)
  MEDLINE   92410648
REFERENCE   3  (bases 1 to 7313)
  AUTHORS   Farmer,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-1995) Andrew D. Farmer, HIV Sequence Database,
            Los Alamos National Labroatory, T-10, Mail Stop K710, Los Alamos,
            NM 87501, USA
COMMENT     HPV60 was originally isolated from a keratinous plantar cyst
            composed of the wall of the squamous cell layer and the horny inner
            substance in the lower dermis displaying cytoplasmic amorphous
            eosinophilic inclusions and vacuolated structure with or without
            pyknotic nuclei. These morphological changes are not characteristic
            of skin lesions associated with other HPV types (1, 2, 3, 4, and
            7). HPV 60 was found in three out of three other plantar cysts
            with similar pathological changes; a fourth cyst without the
            associated pathological changes as negative for HPV60. These
            results strongly suggest that HPV 60 has unique biological
            properties which can induce a plantar cyst. The HPV60 genome
            appeared to be in an episomal form. HPV 60 showed no similarity with
            other known prototypes (HPVs 1-59) by Southern blot analysis under
            stringent conditions. Phylogenetically, it is associated with HPV
            4, 48, 50, and 65, a group of primarily cutaneous types.
BASE COUNT     2431 a   1241 c   1462 g   2179 t
ORIGIN        102 bp upstream from beginning of E6 cds
        1 tataggatat acactgattg ttggcaacta tcattaaatc ataaaaaaaa tatgACCAGA
                                                               E2 binding ->
       61 AGAGGTATAT ATAAatagct gagacgttcg taaactttag gATGcagatg gaagaagaca
          E6 orf start ->                           E6 cds ->
        signal ->
      121 ggtttccaac aacagtggct gattattgct ctgaatttga tattcctttta aaagatctta
      181 agttaaaatg tgtatttttgt agatttttact taactgaaca gcagttggct gcatttttata
      241 ttaagaattt aaagttagtt tggaaaaacc gttattgctt tgcttgttgt actccgtgtc
      301 ttagacttac tgctaaattt gaagctgaaa actattttca gtgtatgtgt aaaggagaag
      361 tattagaagt tctaactcgt attcctttga gttctctttc tgtacgctgt tttgattgcc
      421 ttacattgtt atcttttgca gaaaaaatag attgtataat tagtggtcaa aactttttatc
      481 ttgtaagggg tcgctggaga tcatattgta gaaattgcat TGAgaaATGA ttggtaatca
                                                  E7 orf start ->   <- E6 end
                                                    E7 cds ->
      541 gcctaatgtt aataaccttg atgtaaattt ggaggagttg gttttacctg tcagtctttt
      601 agctgatgag gagttgtcac ccgatggtga tcctgaggag gaggagcact atccctatac
      661 aatagacacc tgttgcaaac cttgtggagc aggtgttaga tttactatta ttgctactcc
      721 ttctgctgta ataactcttc gtcaactatt gcttcaagaa gtgttcctga cctgttTGAg
                                                              E1 orf start ->
      781 gtgctccaga tcccttttc gacATGGcag atccTAAtaa aggtattaat tctcttgaat
                           E1 cds ->              <- E7 end
      841 taaatgaggg gcatagcgaa tggtatgttg tgacagaagc tgagtgtatt aatagtttgg
      901 atacaatgga agagctattt gaagaaagta cagatggatc tattgtgtcc aatctgatag
      961 acgattccga ggaactggag gagggaaatt ccctggcact ctacaatgaa cagttaacag
     1021 aggattgcaa tagagctatt ttagcgctaa aacgaaagtt aactaaaaca cctttgaaaa
     1081 gccaggacag aacggttgct gacctaagtc caaggttgga agctgtcaca atttcacctc
```

```
1141 agagacaaag caaaaggaga ttatttgagg acagtggatt aggagaagat gaagctacaa
1201 attctattga aaaaaaggta gtttcgaact ctttagagag taatgaaagt gggaccttgg
1261 ttgtggaaac ggacagtata tttcgcagta ctaatagaaa agccacttta ttggctaaat
1321 ttaaagaata ctttggtgta gcatatggag atttaacgcg accgtttaaa agtgatagat
1381 catgttgtga aaactgggtt ataagtgtgt gtgccgctgc agaagaagta atagaagctt
1441 caaaaacagt tatgcagcaa cattgtgatt ttttacaggt tatttcatat gggtttttatg
1501 ctttatattt ggtaaaattt aaaacagcaa aaagtagaga tactataatg aaattatttt
1561 ctttaacatt gaatgtacaa gaacaacaat taatgtgtga tccaccaaaa tctagaagta
1621 ccccaacagc tctttatttt tatagaaggt catttggaaa tgcatcgttt atttatggac
1681 cgtttccaga ttggttagca aaattaacta tgttagacca tgaatccgcc gcgagctcag
1741 aacagtttga acttgctcaa atgattcaat ttgcttatga caataacttg acaactgaat
1801 ctgaaatagc ttataaatat gcattgttag ctgattctga tgctaatgca gctgcatttt
1861 taaaaagtaa tcaacaggta aaatatgtac gagactgcta tgcaatgtta agatattata
1921 aaagacaaga aatgaaagat atgtcaattt ctgagtggat atggaaatgt tgcgatgatt
1981 gtaatcaaga gggcaactgg aaattaatag cacaattttt acgttatcaa gaagttaatt
2041 tcatttcatt tttatgtgca ttaaagacat tgtttaaagg tattccaaaa agaaattgtt
2101 tagtgttttg gggacctcca gatacaggaa agtcttatat ttgttcttca cttacacggt
2161 ttatgcaggg aaaagtggta tctttcatga atagacacag tcaattttgg ctacaacctt
2221 tacaagattg taagcttgga ttcttagatg atgctacatt tcaatgttgg caatacatgg
2281 atgtaaatat gcgaaatgca ttagatggta atcatatttc attagattta aaacataaag
2341 ctcctttaca aataaaatta cctcctttat taataactac caatgttgat gttgaaaatg
2401 aagctagctt aatgtattta aagagtagac tagtattttt taagtttcct aacaaattac
2461 ctttaaaaga aatgatgaa gtactttatg aaattactga tgcatcatgg aaatgttttt
2521 ttatcaaatt tgcaagccat ttagagcTGA cagccagagg agacgagcag cATGaatcag
                       E2 orf start ->              E2 cds ->
2581 gccgatctga ccgagcgttc cgatgcactg caggaacaaa tactgaatct ataTGAgcaa
                                                               <- E1 end
2641 gattcaaagg acattcaggc ccaaatacaa tattgggatt taaatagaaa actatatgtc
2701 acatactatt atgcaagaaa agaaggctat tcgcatttgg gtcttcaacc cttaccagct
2761 ttgcaagtat cagagtacaa agccaagcag gcaatcgaaa tgggattgtt attaaccagt
2821 ttaagcaaat cccaatatgc ttcagaacta tggggtttaa cagatacaag tgcagaattg
2881 ttattaaccc ctccaagaaa tacatttaaa aagaagggat atactgTAAa tgtgtggttt
                                         E4 orf start ->
                                         NH2 terminus unknown
2941 gataataatg aaaataatac atttcccctat acaaattggg aatatattta ctatcaggat
3001 gatattgaac aatggcacag gaccagagga gaggtggact ataatggact ttatttcaca
3061 gaaaataatg gaaatagagc atattttctc ctatttgata gtgatgcaca acatattca
3121 caaactggga catggacagt gcattataaa aaccaaatta tttctgctcc tgttaccagc
3181 tcctcaaaac aatcctccga cgactacact tccaaagccg ggcagcaacc ccacttcttc
3241 gcctcatcct catcgccgac tactaccgac ggaggacaaa cctcacaaga gggagtctct
3301 agtctctacca cgtcgccgag tgctgttcga ttacgacgcc gaagatccaa cgagcaacaa
3361 agagaactat cctccagaga gtcgccccgt accaaaagac gcagagtacc cgacgaagtc
3421 gaccgacaga gtgccgtggg gtctgccca actgctgaag aggtgggaag cagacataga
3481 tctcttccta gaagcggtat atcaagactt gcaagacttc aaggagaagc tcgggatcct
3541 ccaatcctgc TAAttaaagg tctagcaaac tcccttaaat gttggagata tcgattgaaa
               <- E4 end
3601 aagtatacac gttattttaa atgtatgagt acagtattta gatgggtaga catagacgta
3661 ccagaatctt ctagacataa attgttagtt gtttttaaTG Ataccacaca acgagatgtt
                                   L2 orf start ->
3721 tttatgaaat tagtaacttt gccaagaggt tgtacatata catttggaac attaaactct
3781 ttgTAAtatt tgtaaaacaa aacaaaATGt atgctagagt aaagcgtgtt aaaagagact
        <- E2 end       L2 cds ->
3841 ctgttgaaaa cttatataaa caatgtcaac ttggtgctga ctgccctcct gatgtacgaA
                                                               signal ->
3901 ATAAAgtaga aggaaccaca cttgcagatc gattattaca aatatttgga agtatcttat
3961 acctggggaa tttgggtata gggacaggta aaggatcggg gggtgctaca ggctatacac
4021 ctcttggtac tgcaagagtg ccagcgtcta cacctggaac agttataaaa cccacacgAC
                                                               E2 binding ->
4081 CGTTTTCGGT tcctttagat ccaataggtt ctggaattcc atcgcaacct gtaggaggtc
4141 ggttacctgt ggacattata gatgccagtg cttcctctat tatacccttg caggaggtcc
4201 tgccagagac caccattata gtaggcggtg acagcggtcc tggattgggt gcaagtgaaa
4261 ttgatatagt ttcagaacca agACCAGATG TGGTcggtgt tgatactcaa ccaacagtat
               E2 binding ->
4321 atacctccat tgataataca gttgctacat tagacattac accagctaca ccaccagtga
4381 agaaaataat attagaccct ataagttcag gttctgaagg tgctgcagca ataacatttt
```

HPV60

```
4441 cagatataag tgctgcagat ttaaatgtat ttgttgatcc tcagggagca ggagatagaa
4501 taagttttgg agaagaaata gagttaggcc ctattaatca gcctgctcaa tttgaaatag
4561 aagaacctcc acgtaccagt acacctgggg agggatttca acgtgttaca acacgggcaa
4621 gggaattata taatagattt gtgcagcaac agccaacaca aaatattgac ttcttaggac
4681 gtccttcccg cgctgtacaa tttgaatttg aaaatcccgc ctttttaat gatgaggtga
4741 ctatgcagtt tgagcaagat ttgcaagagg tggctgcagc gccagaccag gattttgcag
4801 atgtcaggga attaggacga gctcggtttt cagaaacatc tgcaggaaca atacgagtca
4861 gcagattagg aacaaaagga acaatgaaaa cacgaagtgg actaactatt ggtcagaaag
4921 ttcactttta ttttgatatt agtgatattc cagctgcaga aaccatacaa ctaagaacat
4981 taggcgagtc ttcccacgat ttttctgcag tagataatat tactgaaagt acatatatta
5041 atttaactga aactacaaat gaaggcctaa tacctgacaa tatactagag gatgaattta
5101 cagaaaactt taataatgca cagctgatat ttgcaacaat agatgaagga gaatctatga
5161 tcatgcccac aattcctcca ggtgtagctt aaaagttatt tattccagaa atagctgcaa
5221 gcgtgttaaa tgtggttcat ccttcttctg agtggactat tttgattcct aatgttccag
5281 atgaaattat tcagcctgct atggcagTAG ATGtttatga tgacttttat ttacatcctc
                     L1 orf start ->
                first 'ATG' of L1 cds ->
5341 atcttcttag gcgacgcaaa cgaaaacggt tggatttttt tTAAtgtttt acAGATGgct
                                                      <- L2 end   /\ 3' sj
                                         probable L1 cds start ->
5401 ctttggttgc agacagctgg acaattgtat cttccaccaa gcaagcctgt tgctcgtgta
5461 ctcagtacgg atgaatatgt acagccaaca aatttagttt ttcatacagg aactgacaga
5521 atgctaattg taggacaccc ttatttttgat attatagaca gtggttcaaa taatattact
5581 gtacctaaat gtagtggaaa tcagtttaga gttatgcgac tactatttcc agatccaaac
5641 aaatttgcaa tgatagacag agctgtcttt aacccagaaa gagagagatt agtctggaga
5701 ttggaaggcc ttgaaattgg tagaggtggg cctttaggaa ttggaacttc tggccatcca
5761 ttatttaata aatatggaga tacagaaaat cctgccgcat atcctttaaa gcagaataat
5821 ggtgatgaca atagaatgga tgtttcaatg gatccaaaac aaatgcaatt gtttattgta
5881 ggatgtaaac cagccactgg tgaacactgg atattgctaa accttgtga tcctgctcca
5941 gcaaaaggat cctgtccccc aattaaatta acacagtcta taattcagga cggtgaaatg
6001 tgcgatacag gctttggaaa tgcaaatttt attaccttgc aagaagataa atcaggtgtt
6061 cctctagata ttaccaatga aatttgtaaa tatcctgatt tactaaaaat gacaaaagat
6121 atctatggag atgcagtatt tttctttggt aaacgtgaac aaatttattc caggcattat
6181 tttgtaagag gagggataga tggagatagt ttgccagatt ctggatatta cttagcacca
6241 caaactgata aacctcaaaa taaccttggt ggttatagtt acttccctac accaagtggt
6301 tctgtagctt ctagtgataa tcaacttttt aatagacctt attggcttca cagagcccaa
6361 ggtgcaaata atggtatttg ttggggtaac caattattca ttactatagt agataatact
6421 cgaaatacta atttatctat ttcagtttac aaacaagatg ctgctattga taacagatat
6481 aaatataaac aggaagattt tagacagtat ttacgccata ctgaagaata tgaagtagag
6541 ttaattctga gactctgtaa agtgcctttg aatccagata ttctagctca tttaaatgta
6601 atggacaaaa atattttaga ggattggcaa ctttcatttg tacctccacc accacaaggt
6661 attgaagatg cttacagata tataatgtct caggcaacaa tgtgtccaac agatgttcct
6721 aatactgaaa gggaagaccc ttataaacag tatacatttt ggacaataga tttgcaagaa
6781 cgtttttcta acgaattatc ccagttttct cttggaaaaa gatatttata tcagtatggc
6841 ttacttaacg gacgaaaacg gtctgctagt agttttgtaa cgaagaaatc aaaaactgtg
6901 aaacgtaaaa gaactaaaTA Aatacatata tatattatat gattaatgaa ggatatatta
                     <- L1 end
6961 attaataatg ctgctgttgt aaaatgaagg attAATAAAt tgcacatcag caagtgtctg
                                 signal ->
7021 actcacgggg tctacatttt tgtctaccgc gccttcataa ttatcttggc attgcatatc
7081 gagactataa gaaacaatca cttggcacac tctcggattt aacaagcatt gctggacact
7141 taccaACCGT AACTGGTtaa aaaagagcgc caaaggggaa acaggcaacc gtttttcgttc
     E2 binding ->
7201 tcctgaagga acgttttcca ggtaagtaca tgttcataca aagttttgac tcactctctt
7261 tgttagaaac catttccgtt ctgtggatat cgccagACC AAAAACGGTt gcc
                                 E2 binding ->
//
```