Group H Sequences

<table>
<thead>
<tr>
<th>HPV5</th>
<th>HPV8</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV9</td>
<td>HPV12</td>
</tr>
<tr>
<td>HPV14d</td>
<td>HPV15</td>
</tr>
<tr>
<td>HPV17</td>
<td>HPV19</td>
</tr>
<tr>
<td>HPV20</td>
<td>HPV21</td>
</tr>
<tr>
<td>HPV25</td>
<td>HPV47</td>
</tr>
<tr>
<td>HPV49</td>
<td></td>
</tr>
</tbody>
</table>

INTRODUCTION

Group H consists of human papillomaviruses HPV-5, HPV-8, HPV-9, HPV-12, HPV-14d, HPV-15, HPV-17, HPV-19, HPV-20, HPV-21, HPV-25, HPV-47, and HPV-49, a group primarily associated with the multifactorial disease, Epidermodysplasia Verruciformis (EV). Patients with EV tend to have depressed cell mediated immunity [1]. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas [2]. 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 [2, 3]. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding [3, 4, 5].

This group forms two major branches based on phylogenetic analysis, one which can be subdivided into two minor branches. These clusters have been designated as a, b, and c. In addition, HPV-49 forms a remote branch off of the b cluster. HPV-49 is curious in so far as EV associated lesions; the awkwardness of its position is perhaps reflected in the distance of its relationship to the other papillomaviruses in the group.

Ensßer proposed a classification scheme of those sequenced EV types based on the presence or absence of conserved EV-specific and other regulatory regions within the LCR. This categorization is consistent with that obtained through our phylogenetic analysis. In this system, viruses in groups A1 and A2 possess the EV-specific M33 and M29 regulatory regions, however viruses in the B group contained only segments of these motifs. Subgroup A2 differed from A1 by the presence of two of four degenerate E2 binding sites [4]. Other researchers have also devised classification schemes based on other criteria [6]. Dr. H. Pfister classified the EV-associated viruses by the level of cross-hybridization to each other and to those in other groups. In his system, the D1 viruses correspond to both the a and a2 categories proposed here. D2 corresponds to the b cluster, and D3 is composed solely of HPV-24, a virus not presented in this compendium [7].

Cluster a consists of HPV-5, HPV-8, HPV-47 and HPV-12. Both HPV-5 and HPV-8 are associated with macular lesions which frequently progress to malignancy [8, 9, 10]. 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 [10]. 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 [11]. In addition, amplifications of the LCR have been reported in
HPV-5 associated carcinomas [12]. 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 [13]. HPV-47 is primarily associated with benign lesions, however, it has also been detected in cases of malignancy [6]. HPV-12 induces benign macular and flat wart-like lesions [14].

Cluster a consists of HPV-19, HPV-25, HPV-14, HPV-21 and HPV-20. 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 [7, 6, 15]. HPV-14, HPV-20 and HPV-21 induce flat-wartlike lesions; HPV-20 and HPV-14 have been detected in carcinomas [6, 15].

Cluster b consists of HPV-15, HPV-17, and HPV-9. HPV-15 was isolated from a benign flat wart-like lesion [15]. HPV-17 was isolated from benign macules and subsequently from squamous cell carcinomas and the malignant melanoma of an immunosuppressed patient [15, 16]. HPV 9 DNA induces both macular and flat wart-like lesions, however it has also been identified in a keratoacanthoma [14, 17].

HPV-49, a type which clusters with the b group of the EV associated viruses, was isolated from the flat warts of a Polish renal transplant patient. Favrre 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 [18].

HPV-5 and HPV-47 are close enough to each other to be considered “close types”- sequences which 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).


HPV-5 has been associated with macular lesions which frequently progress to malignancy. Yabe et al. (Int J Cancer 43: 1022-8) 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. 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 (Yabe et al. Virolgy 183: 793-8). In addition, amplifications of the LCR have been reported in HPV-5 associated carcinomas (Deau et al. Virology 184: 492-503). HPV-5 and HPV-8 have also been found in significant numbers in squamous cell carcinomas of renal allograft patients. Barr et al. (Lancet 1: 124-9) detected either HPV-5 or HPV-8 in nearly 60% of the cases surveyed in the Scotland area. HPV-5 is considered to be part of the a$_1$ cluster based on phylogenetic analysis. This cluster includes HPV-5, HPV-8, HPV-47, and HPV-12. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

BASE COUNT 2376 a 1547 c 1736 g 2087 t
ORIGIN 354 bp upstream of HindIII site.

E2-bind <- signal ->
61 ttgtgctcag catatctag gatttcgctg tacatatctg tttttgcctg tatcgttttc gtatcctgta
121 atatatccat atatatgatg acatatAATAA ATATATATAT ATATAAgtgt ctaagattgg
signal -> E6 orf start ->
signal ->
181 gttcctttctg aatcagcgcA TTgcttgagg agcgcgaacct cacacagaaac tcaacagaaaaa
E6 cds ->
241 agataaggca gaatttacct tattgatag agacttcagt gcattcctag gcattcctgt
301 gattgatgt ttaatacct cgaatttctg tgtgcaacct ctaatattg tggaaacctg
361 tgttgatcag catcctaacat gatttctag tggaaagat tgtgcttctg tctctgtctg
421 tggctgtcag tgggctgcca ctgcaacctta gattttaac caattttctg agcagaaagt
481 gttgaggaga gatttctagc tattttctg caatttctg tttctgttac atattatag
541 ttttacctgct tattagcctt ctgctattc agaagattg tattgtctg gcagagcctc
601 tcctctctctgc aagttgaggaa agcgcgaacct gggcaacctg agcagaaagt
E7 orf start ->
661 tcaATGatgac TAAagaggct aaggaggcaac atatttacct ggagctcagt gaggtgcacac
E7 cds -> <- E6 end
HPV5

I-H-5

SEP 94

721 cccaagtgct accagtgac cttttttgtg aagaggaatt accaaacgag cagggaaaggg
781 aggaggacgc tgacacgacaa aggatctcttg acaaaagatt gagagatcaca
841 acgtgtgagct ctaacctgtct attttttgct acgcaccaga acetttgatt agaccttttccc
901 aaacgcaccgc GaACggagat ctgcagctcc tgtgccctga cgtgcgcca aactgcagac
E1 orf start ->
961 ATGacggagc cTAAttctaa aggtagtaca tctaaagaag ggtttggtga ttggtgttta
E1 cds -> <- E7 end
1021 ttggagagctc actggtgaatt ttgatagaaat gattttgggac aattatttga gagagataca
1081 gaagccggcctatt actgcggacta agttgagaaa atgatttgggac aattatttga gagagataca
1141 cagagcgagctaat gcaacctgct tccagggtcg ccaggaggac aagcagacac
1201 aacaccaccgc ccgcacccct ccaccacctc caaaccgcag ccACGactcca cgtccagaca
E2 orf start ->
1261 gatcaagaag aggagggcga ggATGgagaa tctcagcgag cgtttcaatg ctctgcaaga
E2 cds ->
1321 tcagctaatg aacatttgggt Aagctgcaga acaaacattg caggcacaaa ttaaacattg
1381 gcaaacctta cgaaaagaac ctgtattact ctactatgcag cgggggaaag gtgttacaag
1441 gcttggatat caacctgtgc ctgtaaaggc agtatcagaa acaaaggcta aagaagccat
1501 gaattgcgct atatgcagct attttaag aaggggaaag gcacatggac cagcatatct
1561 gaatcgcgata acataaattt caaggggaag gcacaggtgc agatagtagt
1621 aagtttaatttta gatactttttt atcagctgtt caacactata aaagaatggtt
tcatacctag tgcacccaaaaa attttctgct ttaaatttggt gatctccaaa tctctggaaa
1681 aaccaattgacctggc agacagtctca taagtaattt acctcaatgt caaagatacct
1741 atattttgctt caacggcctaa atttggaaag aacccggcttg tatattttgt gataaaggtg
1801 tgaatgatcgc ttttgttctt ggcagctgtcag ctgtaaggcc aaggtgaggtt catttagct
1861 atattagttgctt cacgtgagct aagttgagata tggatgtcatc cagctcagct
1921 gcttggatat caacctgtgc ctgtaaaggc agtatcagaa acaaaggcta aagaagccat
2041 gaattgcgct atatgcagct attttaag aaggggaaag gcacatggac cagcatatct
2101 gaatcgcgata acataaattt caaggggaag gcacaggtgc agatagtagt
2161 aagtttaatttta gatactttttt atcagctgtt caacactata aaagaatggtt
tcatacctag tgcacccaaaaa attttctgct ttaaatttggt gatctccaaa tctctggaaa
1681 aaccaattgacctggc agacagtctca taagtaattt acctcaatgt caaagatacct
1741 atattttgctt caacggcctaa atttggaaag aacccggcttg tatattttgt gataaaggtg
1801 tgaatgatcgc ttttgttctt ggcagctgtcag ctgtaaggcc aaggtgaggtt catttagct
1861 atattagttgctt cacgtgagct aagttgagata tggatgtcatc cagctcagct
1921 gcttggatat caacctgtgc ctgtaaaggc agtatcagaa acaaaggcta aagaagccat
2041 gaattgcgct atatgcagct attttaag aaggggaaag gcacatggac cagcatatct
2101 gaatcgcgata acataaattt caaggggaag gcacaggtgc agatagtagt
2161 aagtttaatttta gatactttttt atcagctgtt caacactata aaagaatggtt
tcatacctag tgcacccaaaaa attttctgct ttaaatttggt gatctccaaa tctctggaaa
2221 ctacacctag tcgacccaaaaa attttctgct ttaaatttggt gatctccaaa tctctggaaa
2281 tcatacttgactc aaagatggtc gaagtttctgg ttaaaggttgta gattttgtgtc ataattttgc
2341 tcataactgct acatgagtttc gcaacccctt tgctgatagta ctaaagtcggt ctaattttgc
2401 gtaaagacgcc ctgttaagatt atcaactttt aaatattttg gattttgtgtc ataattttgc
2461 gatcaagaggg cgggtgtttttt ctgctgatagta ctaaagtcggt ctaattttgc
2521 tataaagagc gaagtttctgg ttaaaggttgta gattttgtgtc ataattttgc
2581 aagtttctgg ttaaaggttgta gattttgtgtc ataattttgc
2641 acaaatgagtc ctaaagtcggt ctaattttgc
2701 gatcaagaag aggagggcga ggATGgagaa tctcagcgag cgtttcaatg ctctgcaaga
2761 tcagctaatg aacatttgggt Aagctgcaga acaaacattg caggcacaaa ttaaacattg
E4 orf start ->
2821 gcacaccttc caaaagacgct ctgctgatagta ctaaagtcggt ctaattttgc
2881 gctgtaagtttc gcaacccctt tgctgatagta ctaaagtcggt ctaattttgc
2941 gtaaagacgcc ctgttaagatt atcaactttt aaatattttg gattttgtgtc ataattttgc
3001 tcataactgct acatgagtttc gcaacccctt tgctgatagta ctaaagtcggt ctaattttgc
3061 ccaccccttttt tgtgaaggta tatttggaa tgaagattta gattttgtgtc ataattttgc
3121 gtgaccaatt tttgcttcatg ttttcagcag caaattggca cagcatatct
tgaccaatt tttgcttcatg ttttcagcag caaattggca cagcatatct
3181 gtaaagacgcc ctgttaagatt atcaactttt aaatattttg gattttgtgtc ataattttgc
3241 tgaagtttctgg ttaaaggttgta gattttgtgtc ataattttgc
3301 tgtgttttgtc cctgctcacc gcctgacgcc tccagggtcg ccaggaggac aagcagacac
3361 aacacccacc cccgcacccc ccacaccacct cacaacgacc gcTGActcca cgtgcgcca
E5 orf start ->
3421 gcctcaccaga aatgtcagcaa gatgcacaaag gcacaggtgc ctcgagacacc
E2-bind ->
3481 gcctcaccaga aatgtcagcaa gatgcacaaag gcacaggtgc ctcgagacacc
3541 GTCCCGGCTg cggctccgggt cgcggctccgg gcctcggccacc gcacaccctc ggcggttaagc
3601 caggtccggg ctcagctgcc ttcacaaaggc acaacaccgact ccaccgcacct ggcggttaagc
3651 aggaggtgcc cccacccctc gaagaggggg gaagtaggaag gcctcggccacc gcacaccctc ggcggttaagc
3712 gtccaccgcc aatgtcagcaa gatgcacaaag gcacaggtgc ctcgagacacc
3781 tccacaccag caggagggccc gaggtcaggg aggggtcaagc gcagggaggg gcgtggtggtg
3841 agggcgagcc cagagggagt tctccctcct ctcctcccccc cccacaccag gcgagacacc

NH₂ terminus unknown

E4 orf start ->
3301 tgtgttttgtc cctgctcacc gcctgacgcc tccagggtcg ccaggaggac aagcagacac
3361 aacacccacc cccgcacccc ccacaccacct cacaacgacc gcTGActcca cgtgcgcca
E5 orf start ->
3421 gcctcaccaga aatgtcagcaa gatgcacaaag gcacaggtgc ctcgagacacc
E2-bind ->
3481 gcctcaccaga aatgtcagcaa gatgcacaaag gcacaggtgc ctcgagacacc
3541 GTCCCGGCTg cggctccgggt cgcggctccgg gcctcggccacc gcacaccctc ggcggttaagc
3601 caggtccggg ctcagctgcc ttcacaaaggc acaacaccgact ccaccgcacct ggcggttaagc
3651 aggaggtgcc cccacccctc gaagaggggg gaagtaggaag gcctcggccacc gcacaccctc ggcggttaagc
3712 gtccaccgcc aatgtcagcaa gatgcacaaag gcacaggtgc ctcgagacacc
3781 tccacaccag caggagggccc gaggtcaggg aggggtcaagc gcagggaggg gcgtggtggtg
3841 agggcgagcc cagagggagt tctccctcct ctcctcccccc cccacaccag gcgagacacc

NH₂ terminus unknown
HPV5

3901 ggggtctgcT AAgctccggt
gcgtctcc  tggtaagaTt  gagggtcac  ttgatcaqt
<- E5 end
3961 tagttcaaag catacaggac
gacttgcaag attactgga  gaagtcgctg  acccccccaqT
4021 AAtcattgc  aaaggggccc
taacactc  gaaaaatgtc  cgcacagag  ctaaattaag
4081 atacagggaa cttgtagtac
tccttagta  taccggtcga  tgggtgcaag  gagaagggc
4141 tgagctccta ggagccccca
gaatccttc  tagttttttt  tctcatatac  aaagagaga
4201 ttttattgaa gcggtgcgat
taccacccag  gtttgaTAAg  gcctatggca  acctggacag
L2 orf start ->
4261 tctttTAcat  ttacattagc
tggctgtcga  aaaccgtcaca  gcggagacct
4321 tatttttttt tacattttgt
tttgctATG  gcgcgtgcaa  aaacggtcaa  gcgagacct
L2 cds ->
4381 gtaactcata  tttaccaaac
gcaggcactt  cctctctctc  tttgagatgaa  gcggtgcgat
signal ->
4441 AAAgtggaac  aaacaacagt
ttgctgacaat  atttttaaaat  atggcagtgc  tggtgtattt
4501 tttggtggcc  ttggtattag
tacaggccga  ggaactgggg  gtgctacagg  gtacgtgcca
4561 cttggggaag  gtcacagtgt
cctgtcgcga  ggaACCCCA  CGGTGTaag  gccttctctg
-> E2 bind
4621 gttctgaaaa  caatccggcc
cgcgtgatat  ttgccatatt  atacagttaa  cccctggtgaa
4681 cctcagacat  catccggtgt
cctctctact  gatccacctt  gatccacctt  gatccacctt
4741 gaagtagaaa  caattgctga  aatccatcct  gtacctgagg  ggccatcagt  ggatacccct
4801 gtagttacca  ctagcagcag
ttcaggtctc  ggcagctcct  ggcagctcct  ggcagctcct
4861 ccaagctggg  aacagagaggg
gttgatata  actgcataaa  tttgagatgaa  gcggtgcgat
4921 agtattcata  cattgcaacac
ttcagcataa  tttgctctct  ttgagatgaa  gcggtgcgat
5000 ggctgatgaa  ggcagctcct  ggcagctcct  ggcagctcct  ggcagctcct  ggcagctcct
5061 ttttgaaacta  caagaaatgg  ttcatactat  acacaagaca  caaagggata  ttatgctgtc
5121 gatgaactga  aacatgcttc
5181 gatgaactga  aacatgcttc
5241 gatgaactga  aacatgcttc
5301 gatgaactga  aacatgcttc
5361 gatgaactga  aacatgcttc
5421 gatgaactga  aacatgcttc
5481 gatgaactga  aacatgcttc
5541 gatgaactga  aacatgcttc
5601 gatgaactga  aacatgcttc
5661 gatgaactga  aacatgcttc
5721 gatgaactga  aacatgcttc
5781 gatgaactga  aacatgcttc
5841 gatgaactga  aacatgcttc
5901 gatgaactga  aacatgcttc
5961 gatgaactga  aacatgcttc
6021 gatgaactga  aacatgcttc
6081 gatgaactga  aacatgcttc
6141 gatgaactga  aacatgcttc
6201 gatgaactga  aacatgcttc
6261 gatgaactga  aacatgcttc
6321 gatgaactga  aacatgcttc
6381 gatgaactga  aacatgcttc
6441 gatgaactga  aacatgcttc
6501 gatgaactga  aacatgcttc
6561 gatgaactga  aacatgcttc
6621 gatgaactga  aacatgcttc
6681 gatgaactga  aacatgcttc
6741 gatgaactga  aacatgcttc
6801 gatgaactga  aacatgcttc
6861 gatgaactga  aacatgcttc
6921 gatgaactga  aacatgcttc
6981 gatgaactga  aacatgcttc
7040 gatgaactga  aacatgcttc

I-6
SEP 94
7081 tctttataac tactactcgt taaggttcct ttaaagagcag aggtattggc acagatcaat
7141 gcaatgaact ttcgattat gaggatttg caggttagcttttccac tcttgataat
7201 ccaattcagg acacctacag atatatggac tctttggtca cagcggtgctc agataagaat
7261 cctcctgaaag aaagagaga ccctttataag ggctttacatt tttgggaggt agatttaact
7321 gaagatggtt cattagattt agatcaatat tccttaggca gaaaattttt atttccaaggct
7381 gggttacac ccacgACCCT TAAGGTGca aagcagtgct tttaaagg gtttaataaga
  -> E2 bind
7441 ggcaaccaac gcaacgtaa aaatTGAaggt ctagccgaaa gttgtaacatt tttaaaact
  <- L1 end
7501 tttaacagt attcaaggaa tggttgntta ctctgactaa gtataagtct tccaagagta
7561 ccgACCACAC CCGGTacact cagtcgaatt gttgccaata tagaatcagta cagtgccaa
  -> E2-bind
7621 acacacgctc ttgacactag acagaccggg gttcgattata acatgctcgg attagggacc
7681 tccccaaaga agatttaact taCAATCGCT TTTGGCAATC GCATTTGGCA cttgtaaaag
  -> overlapping repeat  <-
7741 ACCGTT
  -> E2-bind
HPV5b

LOCUS HPV5b 7779 bp ds-DNA circular VRL 07-AUG-1991
DEFINITION Human papillomavirus type 5b (HPV-5b), complete genome.
ACCESSION D90252
KEYWORDS complete genome.
SOURCE Human papillomavirus type 5b DNA from benign lesions of an EV patient.
REFERENCE 1 (bases 1 to 7779)
AUTHORS Yabe,Y., Sakai,A., Hitsumoto,T., Kato,H. and Ogura,H.
TITLE A subtype of human papillomavirus 5 (HPV-5b) and its subgenomic segment amplified in a carcinoma: Nucleotide sequences and genomic organizations
COMMENT These data kindly submitted in computer readable form by: Yoshiro Yabe Department of Virology, Cancer Institute Okayama University Medical School 2-5-1 Shikata-cho Okayama 700 Japan Phone: 0862-23-7151 x2630 or 2632 Fax: 0862-22-2846

Yabe et al. studied characterized HPV-5 in lesions of differing severity. They cloned and sequenced HPV-5b from benign lesions of a patient with EV. 40% of the genome was amplified in carcinomas, and was present in an episomal state. In the metastatic tumor, only the 40% subgenomic region was present, and integrated into the host genome. 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. In addition, amplifications of the LCR have been reported in HPV-5 associated carcinomas (Deau et al. Virology 184: 492-503) HPV-5 is associated with macular lesions which frequently progress to malignancy. HPV-5 and HPV-8 have also been found in significant numbers in squamous cell carcinomas of renal allograft patients. Barr et al. (Lancet 1: 124-9) detected either HPV-5 or HPV-8 in nearly 60% of the cases surveyed in the Scotland area.

HPV-5 is considered to be part of the α$_{1}$ cluster based on phylogenetic analysis. This cluster includes HPV-5, HPV-8, HPV-47, and HPV-12. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

BASE COUNT 2378 a 1542 c 1752 g 2107 t
ORIGIN 207 bp upstream from beginning of E6 cds
   1 AACGGAagtt agcaagtTTCC TTTTCCCGT TACGAGGTGC GGTattggga ttttgcaatt
       E2 bind <- -> repeat <- -> E2 bind
   61 gtaatgttg ttgccaacct ccataggcac attcaagttt ttgcctgtat cgttttcgta
121 tcctgttaa aatatccaat gtagttagc ataaataaaT ATATAATATA ATAAgtgtcctct E6 orf start -> signal ->
181 aagattgggt tattctgtaa tcaggcaATG gctgaggag ccaaacaacct acagaaactg E6 cds ->
241 acagaaacaat ataaaggcag aatactctctta accatactagc accttagctg gtaaatatatgtaaatttattg
301 attacactct aattctttctg ttaaattttttt gagaaagatatttagtttgttcttáttggtgttt
361 gtagggtgtag cgtgtgtagtag ggtgcggggca gaaacaacct atatatgttagttgtgtagttgttggtggctggctggtgtgctttggtgttgggtgtggtgtttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
HPV5b

I-H-11
SEP 94

5821 gtggtcatta tacacactca tgacaataca ggggactttt atttacatcc cagtcttcgc
5881 aggcgcccaac gttaaagaaa atatttgTGA ttgtcatgct gaggAAGTgc gttggcactcg

L1 orf start →  L1 cds →
← L2 end

5941 gctaatggta aagtatacct tccaccaatcg acaccgggtg ccagagtcca aagcaccgat
6001 gaatacatcct aaagaccaaa tattctactat cattgcatta gtgacagatt gttaactgtct
6061 ggtctatcctt atttcaatgt atataatatt actggtgata aacctgtagct cctaagaattt
t121 tcaggaatac aacagcaggc atttgcctta aaaaacacac atcccccaacag atttgcatta
6181 gctgatatct gtttttcaac cccttgacaaa gaacctattg tgtgggctct tagagggctta
6241 gaaataggta ggggccagcccc atctggtgta gggagcactgt tccacccctta ttttaaat
6301 gttaaagata cagaaaacag taattaacttt acatgcattta ggtacagatt gttaacttga
6361 gccttttcta aagtgatcag actgaatata cttctttgatct ctaaacaacat cccataatgtc
6421 atgttaggat gcacaccttctg catagagagat cattgggata aagctgtaggc tgtgcacaaa
6481 aatgaccacgc aacagctgcc tttctctcct attgaattaa aaaaatacatata tagaagat
6541 ggtgatatgg catagattgag ttttggaatt atgaacctttta aggcaccttcac agataagtaga
6601 tcagagtcta gttttgatata tgtcaataa aatctcagat cttgccaatct ttttaaatgat
c6661 ccaaatgata tcataggcctg tcctcctctt ttttatgcct ctagggagca atgtaagctct
6721 agacatccttt tctggagag ggttaaatct ggtgatgaca ttccagtgct caaatagtagac
6781 aatggtctac cacaatcta atttcatttt cctgagctct atggccaaag ctaaagactc
6841 atcagggatgc cactgtaatt ccaacagcgt cagttgcttc agtgctctct cagttgcataa
6901 ttctttttca aggctgcttta gtctcagat gcaacacacag taattaacttt aatctcagat
7081 gaataaacc cactgtatgta agaatattag aatctattaa aatcttctaaatt taatttactat
7141 ctttaaagga ggggagtatt ggcaccagct aatcaagaaa ctgtctctct ttttaaatgat
t7201 tcgatttttgag aatcttcttt aatcttcctct aacacacacag taattaacttt aatctcagat
7261 gactcctgag tcaacgcttt tccagataaa aatcttcctct aacacacacag taattaacttt aatctcagat
7321 aaaaagcttc atcttttctga aatctttgta aactgtagttg cactgacagct gacagttgatt
7381 tattttctgag cacaactctac gctgttcttac tcgggtttattata aggcgcccaac atatttgTGA
E2 bind →
← L1 end

7441 acaaaaacag tgcttattac aagggtctaag agaggaacaa agcgccaaag ctaaataATGA

7501 ggcctgACCG AAAGTGTTGac atttttataa acttttacac agatatcttaa gagatttttgt
catagttgat antagon E2 bind →
7561 ttaacctctgct cattgattac tttaaagaag ccagagcacc gatacgttag ccctcctcct
catagttgat antagon E2 bind →
7621 gttggtcacc atatagatac agatgaggttc aacccacacc gggttcacct actgacagcag
7681 gtcgtccgctt ataaacatgt cggataggg acccttccaga aagagatttaag atctcaAAZT
repeat E2 bind →
HPV5d

LOCUS HPV5d    7746 bp ds-DNA    VRL 15-JUN-1989
DEFINITION Human papillomavirus type 5d (HPV5d), complete genome which contains naturally occurring deletions in the late region of the virus.
ACCESSION M22961 M18452 M18453 M18454
KEYWORDS complete genome.
SOURCE Human papillomavirus type 5d DNA.
REFERENCE 1 (sites)
AUTHORS Ostrow, R.S., Zachow, K.R. and Faras, A.J.
TITLE Molecular cloning and nucleotide sequence analysis of several naturally occurring HPV-5 deletion mutant genomes
JOURNAL Virology 158, 235-238 (1987)
REFERENCE 2 (bases 1 to 7746)
AUTHORS Ostrow, R.S.
JOURNAL Unpublished (1988) Univ. of Minnesota, Minneapolis MN 55455
COMMENT Ostrow et al. conducted a study to characterize HPV-5 detected in lesions of differing severity. They mapped three naturally occurring HPV-5 deletion mutants with deletion sizes of 1.33 kb, 1.9 kb, and 2.3 kb. All of these deletions were located within the late region of the genome. Other than these deletions, the HPV-5d sequence is virtually identical to the HPV-5 prototypic sequence. The features table in this entry represents the locations of the deletions found in the three mutants. The first form of HPV-5 characterized contained a deletion of 1.3 kb and is represented by the HPV-5d sequence with the naturally occurring deletion A. The second form represents the 1.9 kb deletion, it is composed of the sequence with naturally occurring deletions B and C. The third and final form characterized was a 2.3 kb deletion, the 5d sequence with naturally occurring deletion D. The third form of HPV-5d characterized was derived from a metastatic lesion. For other site information, see entry HPV-5.

NCBI gi: 333086
BASE COUNT 2376 a 1534 c 1749 g 2087 t
ORIGIN 362 bp upstream of HindIII site.

1 aacgttaagt tgaatcttcc tgttacagag tgcgggtatt ggatttcaca attataatgg
61 ttgttgccaa ctaccatagg catattcaag tttttgccct tatcgttttc gtaatcctgt
121 ataatatcca atatatgtat acataaaaat atatatatat atataagttg ctaaatttggt
181 gattgattgt ttaatacctt gcaatttctg tggcaacttt ctaaattatt tggaagcttg
241 tgcggtatgc tgtggcgcca ctgcaactta tgaatttaac caattttatg agcagacagt
301 gttaggaaga gatattgaat tagcttcagg actttcaata tttgattatg atatcaggtg
361 tccacgctgac cagaggttcg cttttttttg aagggaattt caaacaagc caggaaacgg
421 cggaggaagcc agatctcttg aacctcagag cggcgagcaat ttcgagttct cagagtttta
481 aagcagctact gaccggagat ctgcagctcc tgtgccctga ctgtcgcgga aactgcaaac
541 ttgcgtcggc cagagctggg cgggttgcagga gatattgctg tcttccctgt ctcggtcgg
601 tgcgtctcgc tcgagtcggc ggctgagctt ctcggtcggc cgggttgcagga gatattgctg
661 tgcgtctcgc tcgagtcggc ggctgagctt ctcggtcggc cgggttgcagga gatattgctg
721 atgtttttaag ccatctccag atgttttttg aagggattc ccatctccag atgttttttg
781 aacgagacgg tgcaccaacgg aagtctctct ctaaattttg gattctcagc gcgtttcggg
841 actgctgcgt ctaatcctcc attttttttc acgcacaccc atttgttgat aatcagttttc
901 aacgagacgg tgcaccaacgg aagtctctct ctaaattttg gattctcagc gcgtttcggg
961 atgctgcgt ctaatcctcc attttttttc acgcacaccc atttgttgat aatcagttttc
1021 aacgagacgg tgcaccaacgg aagtctctct ctaaattttg gattctcagc gcgtttcggg
1081 gattctcagc atgttttttg aagggattc ccatctccag atgttttttg
1141 aacgagacgg tgcaccaacgg aagtctctct ctaaattttg gattctcagc gcgtttcggg
1201 aacgagacgg tgcaccaacgg aagtctctct ctaaattttg gattctcagc gcgtttcggg
1261 cccacgctg acgttaacag aacctctcttt cgcagctggc acgcctggct gcaagcctctg
1321 ttaaacaagt aagctctgaag ttgagataat tttatcttttt cgcctcgag cttgctacag
1381 caccctgctg gacagcgcgc tgcagcgtg gatcttgctt tttgttgctt tttgttgctt
1441 aaggaaatcc aacgtgcctg attttttttg gattctcagc gcgtttcggg
1501 gattctcagc atgttttttg aagggattc ccatctccag atgttttttg
1561 aacgagacgg tgcaccaacgg aagtctctct ctaaattttg gattctcagc gcgtttcggg
1621 cccacgctg acgttaacag aacctctcttt cgcagctggc acgcctggct gcaagcctctg
1681 aacgagacgg tgcaccaacgg aagtctctct ctaaattttg gattctcagc gcgtttcggg
1741 atatttttgg ccggggcacg aaacgtgcctg attttttttg gattctcagc gcgtttcggg

I-H-12
SEP 94
1801 tgtatgggat cgggggcgtt tagccatgga ccatatcctg attggattgc ccaacaaact
1861 atattaggtc acaaaagtgc tgaggcaagt acttttgatt tttcagcaat ggtccaatgg
1921 gcatttgata atcacttatt agacgaagca gatatagcat accagtatgc aaggcttgct
1981 cccgaagacg cgaatgcagt agcttggctt gcacataaca accaggccaa atttgtgaga
2041 gaatgtgcat atatggtacg attttataag aagggacaaa tgagagacat gagtatatct
2101 gaatggatat acactaaaat caatgaagta gaaggggaag ggcactggtc agatatagta
2161 aagtttatta gataccaaaa tataaacttt attgtattcc taactgcatt aaaagaattc
2221 ctacactcag tgccaaaaaa aaattgcatt ttaatttatg gtcctccaaa ttctggaaag
2281 tcatcatttg caatcgcatc aataaggtag ttgaagggtg tagtgtgtgc attttaatat
2341 tctaaagtgc aatgtttgtct gccacccct ccaagttgca agatatgctt atttgatgat
2401 gtaacgaccc ctggtggtat atacattgat atacattaata gaatgtcctt ggatggatcc
2461 tataattgtaa atatacgacc ccaacgccaat tgaatattc tcccataatta
2521 tattaacctca acattaattg gcctggtgaa atataccatcc tcatgcagaatt
2581 aaaggttttg aacgtgggaa atatacttga ttagaattat cagcagttgg cggaggttgg
2641 tacgcaattt cagcagttgg cggaggttgg cggaggttgg cggaggttgg cggaggttgg
2701 tacgcaattt cagcagttgg cggaggttgg cggaggttgg cggaggttgg cggaggttgg
2761 tacgcaattt cagcagttgg cggaggttgg cggaggttgg cggaggttgg cggaggttgg
2821 ccacacacta ccacacacta ccacacacta ccacacacta ccacacacta ccacacacta
2881 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
2941 ctaactgacc aaagctggaa atcttttttt acaaggcttt ggacacaatt agacctgagt
3001 gatcaagaag aggagggcga ggatggagaa tctcagcgag cgtttcaatg ctctgcaaga
3061 tcagctaatg aacatttatg aagctgcaga acaaacattg caggcacaaa ttaaacattg
3121 gcaaacctta cgaaaagaag ctgtattact ctactatgct agggagaaag gtgttacaag
3181 gcttggatat caacctgtgc ctgtaaaggc agtatcagaa acaaaggcta aagaagccat
3241 gcatttgata atcacttatt agacgaagca gatatagcat accagtatgc aaggcttgct
3301 cccgaagacg cgaatgcagt agcttggctt gcacataaca accaggccaa atttgtgaga
3361 gaatgtgcat atatggtacg attttataag aagggacaaa tgagagacat gagtatatct
3421 gaatggatat acactaaaat caatgaagta gaaggggaag ggcactggtc agatatagta
3481 aagtttatta gataccaaaa tataaacttt attgtattcc taactgcatt aaaagaattc
3541 gtaacgaccc ctggtggtat atacattgat atacattaata gaatgtcctt ggatggatcc
3601 tacgcaattt cagcagttgg cggaggttgg cggaggttgg cggaggttgg cggaggttgg
3661 tacgcaattt cagcagttgg cggaggttgg cggaggttgg cggaggttgg cggaggttgg
3721 ccacacacta ccacacacta ccacacacta ccacacacta ccacacacta ccacacacta
3781 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
3841 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
3901 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
3961 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
4021 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
4081 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
4141 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
4201 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
4261 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
4321 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
4381 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
4441 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
4501 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
4561 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca
4621 gctacatttg ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca ccaaatgaca

deletion B |->

deletion D |->

deletion A |->

deletion E |->

HPV5d

5521 atagtccagg gacctgttga aagcacaattt atagatatgg atatctctga aatattcatta
5581 tctggaagca ttagaagata tacattcatat tttttagttg atgtgaagctt ggaagatttc
5641 agtgggcttc actgcttgaa gagagaacct gaacctgtttc ctcgcagtgtggc
t5701 tttgaaacta caagaagact tttttcctctt tttttggttt cttggtgtgcc ttttcagttgc
5761 atacagagctacctgttctt ggtgacatag atctccccacag ggtggccatgttt tcctcgcagtt
t5821 atcagaaaaa aagacaagac cagtactgtt gccagcaagtc cagtttcttgc ttagattcgt
t5881 ttaggaaactt cttttagttc ctctgctttg ccttgcgttcc ccaagagcttt tataaggtgct
5941 ggttacgcag attcctttcc caagtttctt ggtgagttcc ggattttcag gccagtcgac
6001 atcagctcgg atggtaaactc tatttctttc tttgccttttt cctttagttt cagtttcttgc
t6061 aataggtttt ctctgctttt gccagcaagtc aagacttttt ttttcaagtt ccctttttcgc
t6121 aagccatacg aacttttttttt accggagactt ttttctttgt cagtttcttgc aagacttttt
6181 atgtccttgtt acaacccctga aaagaactgt tttggttgggc cctttagttt cctttagttt

| deletion A | -> |
| deletion B | <- |

6241 ggttagggccc agcccattaggg tttggtgagtactgcttc accatctttt atcatttacc tttttagttt ggttacgcag atacagagct
6301 ttagacagca acagcagcagac acataagact ttcctttttt tttttagttt ggttacgcag atacagagct
6361 ttgtagaattt catttttagtt tttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
6421 tggcagaaa cttgttagtt cttgttagtt catttttagtt tttttagttt ggttacgcag atacagagct
6481 gactactctct ctttttagtt tttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
6541 atctctctct catttttagtt tttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
6601 gtagtagttt catttttagtt tttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
6661 gatatttttt catttttagtt tttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
6721 gtagttagtt tttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
6781 gtagttagtt gtagttagtt gtagttagtt gtagttagtt gtagttagtt gtagttagtt gtagttagtt gtagttagtt
6841 ggtcgtcgtg gcttcagacc cgtttccgaa ctcacgcgag cttttttttt tttttagttt ggttacgcag atacagagct
6901 caagctcagc aataactgaa aataactgaa aataactgaa aataactgaa aataactgaa aataactgaa aataactgaa

| deletion C | -> |
| deletion D | <- |

6961 acaaggtacc atatactttttct catttttagtt tttttagttt ggttacgcag atacagagct
7021 gagtagttt tttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
7081 ttcttttagtt tttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
7141 gcagcttttt catttttagtt tttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
7201 cctattttgtt tttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
7261 ccttttttt cttttttttt tttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct

7321 gagaagaaag ttttagttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
7381 gccagccagc cttttttttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
7441 ggagttaattt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
7501 ctaatcagtt caatcagtt caatcagtt caatcagtt caatcagtt caatcagtt caatcagtt caatcagtt
7561 cccacaatgc ggtcgtttagtt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
7621 aaccccagc cttttttttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
7681 ttgacagact gttttttttt ggttacgcag atacagagct tttttagttt ggttacgcag atacagagct
7741 acaggctt

//
HPV-8 is associated with macular lesions which frequently progress to malignancy. HPV-8 has also been found in significant numbers in squamous cell carcinomas of renal allograft patients. Barr et al. (Lancet 1: 124-9) detected either HPV-5 or HPV-8 in nearly 60% of the cases surveyed in the Scotland area. HPV-8 is considered to be part of the α6.15 cluster based on phylogenetic analysis. This cluster includes HPV-5, HPV-12 and HPV-47, in addition to HPV-8. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

Fuchs et al. note that HPV-8 is unique among the papillomaviruses in several respects. These characteristics include: the small noncoding region with little potential to form complex secondary structures, a cluster of promoter elements in the 3' half of the E1 ORF, and the homology between the E4 ORF and the Epstein-Barr virus nuclear antigen 2 protein.

Stubenrauch et al. (J. Virol. 66: 3485-93) have deduced the transcript structure of HPV8. They identified a late promoter at P₇₅₃₅$ which gives rise to mRNAs consisting of three exons: an LCR leader, a short segment from the early region, and the L1 gene. There is no classical TATA box recognizable in the late promoter, as is the case with the BPV1 late promoter. Both the HPV8 and BPV1 late promoters show some sequence similarity to the SV40 major late promoter. They have also identified another promoter at P₇₁₇₅$. They have deduced that this promoter is most likely responsible for the early genes. All splice sites annotated in the text have been experimentally determined.
I-H-16

SEP 94

HPV8

301 tcagtacgcg gcaacttttg tggtaacctt tagaaacttg tggatgtgac
361 aaaaagacag tgtgcctaat ttggaaaaat tacgtgtta ctgcgcctgt tcgtgttctg
421 tcgtaccaat ccgaactatt gaaatttatg gattgcgctgt gctgactcag ttgctgcaggt
481 gtatatagaa tgtgcctaat ttggaaaaat tacgtgtta ctgcgcctgt tcgtgttctg
541 ttcgactatt ttaaatgatc aagaaaaatt aagatgtga gaaactgtgc tacggtgtgg ttccttcctcat

E7 orf start ->

601 aaagttagag gaggctggaa agaagtggcg cgcctgtta ctgcgcctgt tcgtgttctg
661 TAAagaggtc actgtgcaag attttgtgtt gaagttaagt gagatacaac ctgaactttg ttgctgcaggt
721 cacgttgac gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg

E7 cds ->

781 AAgacatggaa gaaatgtctt tcaatatagt tgtgcctaat ttggaaaaat tacgtgtta ctgcgcctgt tcgtgttctg
841 caagctcagt ctttttgcga aagcatccttc agaacttttct gtcgtgcctgt tcgtgttctg
901 gtcagagac tcaagacttc tgtgcctaat ttggaaaaat tacgtgtta ctgcgcctgt tcgtgttctg

E6 orf start ->

961 aTAAaggtag tgcgtgcctg ggagagcctt tccctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1021 aaggggtaaa gtcggtggtg tggctgaact ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1081 actttttcat gtttccttat ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1141 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1201 tcgtgcagtg ttcctttcat ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1261 ccagggcagc cctttttcat ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1321 caagactttt ctgcgcctgt tcgtgcagtg ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1381 tgcagagac tcaagacttc tgtgcctaat ttggaaaaat tacgtgtta ctgcgcctgt tcgtgttctg
1441 catttttttt ttcgagctgt gcttcctttt gcgtgcctgt tcgtgttctg

E1 orf start ->

1501 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1561 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1621 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1681 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1741 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1801 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1861 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1921 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
1981 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
2041 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
2101 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
2161 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
2221 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
2281 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
2341 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
2401 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
2461 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg

E2 orf start ->

2521 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
2581 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
2641 aagaaatgtta cagccccata cattttatgt ggtgctgaac ttcctttcat
ggcggtgtg gcgtttgctg ggcctgcttt gcgtgcattt gcgtgcctgt tcgtgttctg
2701 aacATGgaga atctcagcga gccttcctgtc ttaccctagt aaataggtaa agtacagcag

E2 cds ->

2761 GAagctgcg aaaaaaactt ctagcagcag ggcacactct gcgaaaaaga
2821 ggcagactct gcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt
tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt
tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt
tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt
tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt
tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt tgcgtgtgcagt

NH2 terminus unknown

AGcctcactcc ccaccaccgc ggcacacttc ctagcagcag ggcacactct gcgaaaaaga
3361 ccaccacact ccgctgcgtt ccgactaccgc ctagcagcag ggcacactct gcgaaaaaga

/ 3' sJ
6541 cctgattttt tgaaatgca gaatgatgtg tgttgtgtt tttatgcacga
6601 agggacaagt gtcatgtcag acatatttttt gtgagaggg gtaagacagg tgatgacata
6661 cctgtcgctc aaaaatgata tgtcataagag aaaatcaaat attcatatttc tttatatggacaa
6721 gataatcatttttcttgatgaagtttataaagtttaaatgtaaatggctgattttt ggtggtgatg
ttgcttcacttt
6781 gttttcaagt tgtcataattt tttaaacagg cccttcctgc gtcagctgc ccaaggtcctt
6841 aataatggca tttctctgggc tattcataat gttggtgaactt tttgtagccaa cagcccccaac
6901 acacctttta gtttctttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
LOCUS HPV9 7434 bp ss-DNA VRL 04-OCT-1993
DEFINITION Human papillomavirus type 9 (HPV-9), complete genome.
ACCESSION X74464
SOURCE Human papillomavirus type 9 DNA.
REFERENCE 1 (bases 1 to 7434)
AUTHORS Delius,H. and Hofmann,B.
TITLE Primer-directed sequencing of human papillomavirus types
REFERENCE 2 (bases 1 to 7434)
AUTHORS Delius,H.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG
COMMENT HPV-9 was isolated by Kremsdorf et al. in 1982 (J Virol 43: 436-47) and subsequently sequenced by Dr. H. Delius. It has been associated with both macular and flat wart-like lesions of EV, a multifactorial disease, and has also been identified in a keratoacanthoma. HPV-9 is considered to be part of the b cluster based on phylogenetic analysis. This cluster includes HPV-15, HPV-17 and HPV-9. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. These types are rarely detected in lesions afflicted 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

BASE COUNT 2363 a 1393 c 1654 g 2024 t
ORIGIN 220 bp upstream from beginning of E6 cds
1 cgccagcgaac cggcacaattt caactgcaaag gttcgtttgc agaccgtcct ggttcaaaa
61 cgACCGATAA CGGTaagtct tggcacgtag gtggttattt gatcgttggg atgattgtgg
-> E2 bind
121 ttaacacaac aacatataca cattttcata tgACCGCCTT CGTTaataag cttatataga
-> E2 bind
181 catataata TAAGttgcCA TGTatattaa agacgcagatt atggacaggc caaacaacctag
E6 cds ->
E6 orf start ->
241 aacagttaag gaactagcag acactctttg gattctttta atagattttg tgcacccctg
301 taaattttgc atagatatttc tatcattatt gtagcttttt aatagttgtc acaagtgttt
361 aacagtattt gtagcacagg aggattttg gtatggacttc tggtagagct gtgccttgct
421 gtggcagc gtaattttat tcaattttgtg tagctctgtg tagctgtg gagaatagct
481 aacttgatag gtagacgaat tggaaatat tggatattag tttacgctact gtttaagtt
541 atattagcattat gtagctttag cttataagtt gacagtatt TAAGttgctag
E7 orf start ->
HPV9

601 aaacagctgg aaaggattg tggtgcggta gaATGAtttg gaagaagact
E7 cds ->
<-- E6 end
661 actatACCAG AGGTGCTTct aagaacctgg aagagttgcc accaccaccgc tgaACCTGAT
<-- E2 bind
E7 cds ->
721 TGGTacgaag aatttgcaga aagaacctga cggagggagc agttgctcgc acctcataaag
tacaataatc
781 atctgactctct ctgctgctgc ggctgacaa aacccactgc tggacaaatc ctgatgctg<-- E4 end
841 gagaattgcag cgcagcagac atttttggctc ctggtatatat ccgagagctc ccgagagctc
901 cagctgcagc ttcgcACATg gTGACAAata atgcctgcagc ttcgcACATg<-- E7 end
961 tagtgcttgcttat ccctagatgta aagacagagt cctctgatttc tttacaag
1021 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1081 tgtacggctag agttgacatttg cttcaagtact ctgacaattat ccctcagtct ccctcagtct
1141 atgcctgact ctgacatattt cccagctcag cctcagatcct ccctcagatcct
1201 gttcctgact ctgacatattt cccagctcag cctcagatcct ccctcagatcct
1261 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1321 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1381 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1441 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1501 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1561 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1621 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1681 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1741 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1801 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1861 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1921 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
1981 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2041 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2101 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2161 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2221 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2281 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2341 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2401 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2461 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2521 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2581 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2641 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2701 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2761 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2821 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2881 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
2941 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
3001 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
3061 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
3121 tagtgcttgct tatccaaagctc cctcagtcag actataaatc ccctcagatcct ccctcagatcct
3181 tagtgcttgct tatccaaagCT Gcagccacagt atggcagact tgggcttttgg gaagtgcatg
E4 cds ->
3241 tagtgcttgct tatccaaagCT Gcagccacagt atggcagact tgggcttttgg gaagtgcatg
3301 tagtgcttgct tatccaaagCT Gcagccacagt atggcagact tgggcttttgg gaagtgcatg
3361 tagtgcttgct tatccaaagCT Gcagccacagt atggcagact tgggcttttgg gaagtgcatg
3421 tagtgcttgct tatccaaagCT Gcagccacagt atggcagact tgggcttttgg gaagtgcatg
3481 tagtgcttgct tatccaaagCT Gcagccacagt atggcagact tgggcttttgg gaagtgcatg
3541 tagtgcttgct tatccaaagCT Gcagccacagt atggcagact tgggcttttgg gaagtgcatg
3601 tagtgcttgct tatccaaagCT Gcagccacagt atggcagact tgggcttttgg gaagtgcatg
3661 tagtgcttgct tatccaaagCT Gcagccacagt atggcagact tgggcttttgg gaagtgcatg
3721 tagtgcttgct tatccaaagCT Gcagccacagt atggcagact tgggcttttgg gaagtgcatg
3781 tagtgcttgct tatccaaagCT Gcagccacagt atggcagact tgggcttttgg gaagtgcatg

HPV9
I-H-20
SEP 94
HPV9

I-H-21

SEP 94

3841 tgc ttaagtg ctatcgc ttt cgagaacgca aaaaaaag aggccttagt aaatattata
3901 gtactacgtg gtctgggtta gggaagaaca gttgtgatag agttggaaga ggccggaatga
3961 tttaacgttt tcacacatat gcacacagc aacaatcatc tggactaag aatactaccac
4021 ctacagtaga tggcttctta ggaatgttg atgatcgtTA Agccttacta acgtcaacgc
4081 tggcattgct aTAAacccat actaactaac aaaaaccatac taactaacAT Ggttcgtgca
4141 aacgtaacta aacgtgacct tttacagat atatacagag gctgcaaaaa gctgtgtaaca
4201 tgtcaccacag atgtaataAA TAAAggtgag cacacaacta ttgctgataa aatattttgcaat
4261 tattgaagtg ctggtgtgtg tttcgggggc ttgggaataa gtacaggccg tggcactggt
4321 ggtcccttcct gctatgtagt ttatatcagc ttgatcgtac tggacaccc
4381 actatacttc gcctgggttt gatactagaa ataatttgcc caaactcctat cttttcctta
4441 gagacatgca gacaagttgc ccccaacgc cccagtattt ctcacaggcc tgcacgacct
4501 tgtgaccttt tacctgggag aataagacta atgtctgtaa agccagctgg atccatatttt
4561 aatgtctgtaa tagataactcc agtgttaaca gaggttagaa gaggtctgctc tgcactatttta
4621 gaggtggctg accaaaggcc ttcctatgca accctgtttg ccagacacta ataccataat
4681 ccaagtcttc aaaaatcttc tggagtctca cttatgtagt gttgctgtaag gaagactcatc
4741 catattagag tttttgaaga atctgggggc cagcagctag ctggtcttta ggaatgtagt
4801 acagatctct tgggaacact aagattacca gaaatctctta tggatgtagat ttttttagtg
4861 gatgaagaga caccacccct gcgtgctaca cccgttccaa gacagctaca atcattactt
4921 agttcagctg gacctcattaa tcacagagat cttggcctaa aagtgctggta gacagatcca
4981 ttttaaattg tggggttttt ttcctatagc tgggtgtctc ttcacctcag cccacagagg
5041 tgtgaggtca cacaattat tgtgaagagat ctgtagtcc cgaggagcc gccagagatg
5101 catttttag atgtcataag cctgtacgtg ccttttatata caagacccac cccaggtttt
5161 tgtgacgtgtg actgcagact gccaagagca acaacagcagg caccaggttg gccagagtgg
5221 ggcgcacagc ttcatttcta cccgcctact acaaccaga caacagagca ccagctgctg
5281 ataaccatatt tggtgggaaac ctcagaggtc cttctatgcg cggctgtgag gatgttagat
5341 tcaattttgttt cagatgtaac tggtagttcg aggtagggaa ccgagaaACC
5401 CTTTGTTAGg aagagtgtgga tttttatcct gagaactcatc tgtagtagtg ggttgtagaa
5461 gatttttagtg ggtctgactg gatgtgctag gacacccgca gcacaaacttc ataagaactag
5521 ccaagcttttg aaatcccaag gcacagtacg tttatatcct gaggataata gcagactaca
5581 gttgctcttc cccgctgctg caaaccacca gataatattt cccaccacgc tgcaccccc
5641 acagattgact tccagctatca TGAatcagc gggagttttt attttgcac cggctctccaa
5701 aggaaaac gcaaacgcacg atattttaAA ttttttttttt gcagAGTGca tttgtgcttca
5761 cagcaggtgg taaagttatt ttgcagccagc ccaaccaggc gcgagaggtt ccaagccagg
5821 atgaattagc ggaanacaca atatttttt tattctgcaatt tgaattgttg tgtgcactaca
5881 tggaggctact atactgtat gcgtcgctag ggcagcgaca aagattggaat gtcctttaag
5941 tgtctgtgtag ctagtctcg cccctttagaa ttagctttcc gccatatccaa agttttgctt
6001 tgcagcagat tggcttctt atacctttata aggaaagctt ctgtggggcct tgtagaggtta

E2 bind ->

L1 orf start ->

E2 end

L2 orf start ->

L2 cds ->

L1 cds ->
6061 ttgaatagg cagaggacaA CCTTTAGGGG TtngaacacT aggtacccca tttttaata
-> E2 bind
6121 aggttagaga cacaaggaac ttagaacat atcaagggac aaacaaggtt gagagggaa
6181 acacatccttt tgcacccaa caggtaccca tggctcattat aggatgcttt ccatgcctttg
6241 gagaacacgt ggttaaaggg aaaggtgttg aaaaaggttg ctaataatca aaatggttctat
6301 gctctctccttt agaataaaga aacagcttaa ttgaggtggt ggacatgtttt gatattgggat
6361 ttggaataa caacaataag gaaagtctct ttaataagtc ttgtaagac ccagttattg
6421 tttgtaaacac cttgtaaatc ccagacctttt taacaaaaag aattaggttt taaggagttg
6481 cattgttcttt ttttgaagaa agagaacaat gtattgccag gcattatatt ttagagggag
6541 gttcagttgg tgtgatgttt cctgatggtt cagtaaaccac ggattcaata tttttttttcc
6601 cagcaaaaaaq tgtagcaacaa ccaacgaacaa tagcttactt caccaactttt cctacagttta
6661 gttggctcatt agtaacttctga gattctcaat tgggaatattag gcgatttttg gctcaaagag
6721 caacaagtca caacaaatgc aatgatatgg gtaatcagat aatgtcagttt cttgctcaca gttgcagaca
6781 atacagttaa caaccaatttt accattaagtg tgtgatcaagtt gcagtttcaac acagaagat
6841 aatggtcacaatatatt gacatgtgga agaatatcag atttcatttaa
6901 tcttacagtgt gttagaagtg cctttagatgt ctagagtttt atcccaagtaa aatgcaatga
6961 actcaggtatat tttagagggatt ggtgaataggt gtttatgccc aaactctgtga aatgctgttc
7021 atgatcctt caagataatt gttcagatt ccacaaacag cccagagagt gttgcaagttta
7081 cagagaaaagagtctctttt gccaattact catcttggaag agttgacactc atctgaagat
7141 tatactgtgaa tttgatacctg attccttttag gtagaataatt tccttttcgaac gttgggtggtc
7201 aacacagaaacgtctctatattt aaaaatcttg ttaaataatc taaaaatgtt aaagagaagggc
7261 gaaccTTACC GATATCGGGTt tccaaattaa tttaaagttt ccaatgttga atgtagttac
<- L1 end
-> E2 bind
7321 ttttttaacc atcttcggtga ctaacccgta caagtcacaC cagacgcgACC GCACCGGGTt
-> E2 bind
7381 tatctgatta taagttgacat ctggtgtga atgagacgctagattataaacaactc acagagccttt agagagtttt cttttttttt gttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
HPV12

LOCUS HPV12 7673 bp ds-DNA VRL 04-OCT-1993
DEFINITION Human papillomavirus type 12 (HPV-12), complete genome.
ACCESSION X74466
SOURCE Human papillomavirus type 12 DNA.
REFERENCE 1 (bases 1 to 7673)
AUTHORS Delius,H. and Hofmann,B.
TITLE Primer-directed sequencing of human papillomavirus types
REFERENCE 2 (bases 1 to 7673)
AUTHORS Delius,H.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG
COMMENT HPV-12 was isolated by Kremsdorf et al. in 1983 (J Virol 48: 340-51) and subsequently sequenced by Dr. H. Delius. It has been associated with both the benign macular and flat wart-like lesions of EV, a multifactorial disease. HPV-12 is considered to be part of the a$_8$15 cluster based on phylogenetic analysis. This cluster includes HPV-5, HPV-8 and HPV-47, in addition to HPV-12. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

BASE COUNT 2375 a 1531 c 1714 g 2053 t
ORIGIN 199 bp upstream from beginning of E6 cds
1 tttgtACCAGG TGGCGTGaca tttcccaata gcacattata ctagattgtt gttgccaact
-> E2 bind
61 accatcatca gtctcaagtg tttgcctgta tcgctccccgt atcatactaa ttctgtatat
121 aataataata aaatatatat atatatatat atatatatat tataaaatgc tcgctccttt
181 tgcaaatgTA tgagccaaac TGcaccacag gcgcagttcg cagacacag cacagcagac
e6 orf start ->E6 cds ->
241 gcctagccg cccacaacta ttaagaggtt agctgacctt ttagatatac cttaagttga
t301 cttgctgaa ccttgcacat tttgcgagaa cttttcagat tttctggaag tgtgtgattt
361 tgcaaaaaag cagcttacac taatttgga gaagttcagtt gctctgaggg gttgcagcag
421 ttgtgccgc gcctagtccca ttaagaggtt ttagatatac tctagttggc
t481 tagagataa gagctgctta tggagagtc ttagaggtag gcggagctgg gcggagctgg
gtctgtgataa ctttttagata caattggaata gtaattgtgg gcggagctgg tcgagcaga
e7 orf start ->
661 taATGatcgg TAAagaggtc accgtgcaag atttttcatt gtagcttgct gcgttgcaag
e7 cds -> <- E6 end
721 ctaggttgtt accagtaacct gcgcagctgtg tcttgggttg tcttgggttg ggcattacaag
781 agaggtggc aggtgcggtg tcgtagcgaa cttgctgacag gcgtggagtt ggtgccatgc
841 cctgagggc cctggggtg ttagaggtg tgcctctgga ttgagcaggt gcgtgtagcag
901 aggacgcgct GATcgtgctg ctgagtcgac tggcagcagc ggtgcgcagc gccgcacagt
e1 orf start ->
961 ATGcggggg cTAAagaggt caccgtgcaag atttttcatt gtagcttgct gcgttgcaag
e1 cds -> <- E7 end
1021 ctaggttgtt accagtaacct gcgcagctgtg tcttgggttg tcttgggttg ggcattacaag
1081 gagaggtggc aggtgcggtg tcgtagcgaa cttgctgacag gcgtggagtt ggtgccatgc
1141 cctgagggc cctggggtg ttagaggtg tgcctctgga ttgagcaggt gcgtgtagcag
1201 agtccaaaaag ctaggtgcaag atttttcatt gtagcttgct gcgttgcaag

I-H-23
SEP 94
HPV12

1261 cagaaatcga acacgaaggct ctttgcatag cagacacgct gactcgagct atctcttaac
1321 aatgaaatgg aagatggttc tcttgaggtg gaggtaccgg ctatagactc tcggccggtga
1381 gatgagggag gatcaggggc catagatatt gattatctgt cattgctgcg tagtagcaat
1441 attaaatatt actctatgaa acgacatttt cattttctcg ccttttctcc cttggtagcg
1501 actcgcagct ttaaaagtta caaaaactgt tgtgaatggt gggttttagc tgtgtatgca
1561 gtatcagatg tgtacagatg acctacagct tactacagatg aagagagtgt
1621 tgggctcttg ggaatttggt attcatttca ttccttcaggg ccttttttgccc cttttttgcc
1681 cgcggtactg tgcataagtt aatgacatca atgctaaatg tggcagaaca ctggagagtt
1741 cgcagcagag gcagccacat caacctagct gcccacatca gcagcagcag gcagcgagat
1801 ggtgtaaaat ggtgtaaaat ggggagttga aacctgcttc ggtgtaaaat ggggagttga
1861 ggccatcaaa atgctgaagc aagcacattt gatttttcag ccatggtcca atgggccttc
1921 gataataatt acttagaaga accagatata gcttatcaat atgccaagct tgcaccagaa
1981 gatagcaatg cagtagcatg gctagcacat aatcaacaag ctaaatttgt aagagagtgt
2041 gcagcaatgg tacggtttta taaaaaagga caaatgaaag aaatgagtat gtcagagtgg
2101 atacacacaa aaattaatga agtagaaggt gaagggcatt ggtcagatat agtaaaattt
2161 ttaagatatc aagatgtaaa ctttattacc tttttggcag catttaaaaa ctttttgcat
2221 gctgtaacaa aaccaacattg tattcttata tatggcCCCC caaactcttg aaggacatca
2281 tttgcaatgt cactgtaaaa agttttttaa gttagagttt tgcgtatctg gatattttgt
2341 agtcaacttt cgttgcagga gaagctgttt tgctttatta tgctaggcaa aagggtatta caaggcttgg
2401 cacccgcctg ctcctgcaaa tactatctca ctctctctct ctctctctct ctctctctct ctctctctct
2461 ggttcatgatg atctatttga aagctcaaag cagctgttgc aacagcattg tgactatatc
2521 attaaatatt actctatgaa acgacatttt cattttctcg ccttttctcc cttggtagcg
2581 cgcggtactg tgcataagtt aatgacatca atgctaaatg tggcagaaca ctggagagtt
2641 ggtgtaaaat ggtgtaaaat ggggagttga aacctgcttc ggtgtaaaat ggggagttga
2701 gaagaggagg gccaaccATGg aagatctcag cgagcgtttc aatgttctgc aagatcagct
2761 aataaatatt actctatgaa acgacatttt cattttctcg ccttttctcc cttggtagcg

E2 orf start -> E2 bind

E4 orf start -> E4 end

E4 cds ->

3241 tggcgaacctg ttttattttt ctttttttca ctttttttca ctttttttca ctttttttca
3301 gtttttatta ttgctATGgc gcgtgctaag cgggtcaagc gagactctgt tactcatata
3361 tccctactttt aatattcagc aacagctttc ctcttatttt ctttttttca ctttttttca
3421 cagcagcagc cccccctctt atctcctctt catctctctt ctcttatttt ctttttttca
3481 ggaaccgag cagaagctgt gcagtggttc ggggggggg ggggggggg ggggggggg
3541 ctcctactttt aatattcagc aacagctttc ctcttatttt ctttttttca ctttttttca
3601 gtttttatta ttgctATGgc gcgtgctaag cgggtcaagc gagactctgt tactcatata
3661 tccctactttt aatattcagc aacagctttc ctcttatttt ctttttttca ctttttttca
3721 cagcagcagc cccccctctt atctcctctt catctctctt ctcttatttt ctttttttca
3781 tggcgaacctg ttttattttt ctttttttca ctttttttca ctttttttca ctttttttca
3841 cagcagcagc cccccctctt atctcctctt catctctctt ctcttatttt ctttttttca
3901 aaacataatc gcacgcaagc gcaacagctgt gcagtggttc ggggggggg ggggggggg
3961 ttgccaaggg ggggcaacac ccggtctttt cggggtcggc ggggggggg ggggggggg
4021 aagctgtttt aagctgtttt aagctgtttt aagctgtttt aagctgtttt aagctgtttt
4081 ctttcttatttt ctttattttt ctttttttca ctttttttca ctttttttca ctttttttca
4141 tgtgttatttt gcaggtcact tgcagccact ccaacggttc cccctttttt ctttttttca
4201 aataaatatt actctatgaa acgacatttt cattttctcg ccttttctcc cttggtagcg

E4 orf start -> E4 bind

E4 end

L2 orf start -> L2 cds ->

4261 gtttttatta ttgctATGgc gcgtgctaag cgggtcaagc gagactctgt tactcatata
4321 tccctactttt aatattcagc aacagctttc ctcttatttt ctttttttca ctttttttca
4381 aacacagtgt cttgacaatat tgtgaatgtg ggtgcttttt tggagggcct
4441 ggtattggtA CCGTTCGGC Tactgaggat gtcactggat acagacctct acctgaaagg
   -> E2 bind
4501 cccggtatcc gtgtgtaagg gacttcccaag gtgtgtaagg cttacttgt ctctgtaacct
4561 gtgggtccag cagatatact acaactacag acatctcagct ctggtcagcc actgcttccc
4621 tcgtaggctc cttacaagca atctcagcct cttcaggtc sgactcagct ctcctgccct
4681 atactgctggt ttcagctgtt ctcagcagtt cttcagcagtt atactgctggt ttcagctgtt
4741 agagacggtg ccagacgcct ctcagctgtt ctcagcagtt cttcaggtc sgactcagct
4801 agagacggtg ccagacgcct ctcagctgtt ctcagcagtt cttcaggtc sgactcagct
4861 gtggctgagt gcctagactt ggtgcagaat catgtgctgt gcgtgtaagg cagatatact
4921 actataaaga ggctacgacc gcatacgcgt ggtgcagaat catgtgctgt gcgtgtaagg
4981 tttgagcagg caagagccag gcctcttgcga ggtgcagaat catgtgctgt gcgtgtaagg
5041 accactgccg gcctagactt gcgtgtaagg cagatatact acaactacag acatctcagct
5101 gtggctgagt gcctagactt ggtgcagaat catgtgctgt gcgtgtaagg cagatatact
   -> E2 bind
5161 cccggtatcc gtgtgtaagg gacttcccaag gtgtgtaagg cttacttgt ctctgtaacct
5221 cctgtatttg aggaggatat aacaaatatt tttgaacagg acttagaaac atttgaggag
5281 cctgtatttg aggaggatat aacaaatatt tttgaacagg acttagaaac atttgaggag
5341 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5401 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5461 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5521 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5581 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5641 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5701 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5761 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5821 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
   -> L1 orf start ->
5881 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
5941 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
5991 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6051 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6111 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6161 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6221 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6281 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6341 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6401 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6461 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6521 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6581 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6641 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6701 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6761 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6821 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6881 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
6941 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7001 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7061 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7121 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7181 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7241 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7301 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
   -> E2 bind
7361 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
   -> L1 cds <-
7421 ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
5221 cctgtatttg aggaggatat aacaaatatt tttgaacagg acttagaaac atttgaggag
5281 cctgtatttg aggaggatat aacaaatatt tttgaacagg acttagaaac atttgaggag
5341 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5401 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5461 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5521 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5581 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5641 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5701 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5761 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
5821 gcacccatcg ccctacgtct ccagctagtt cctcagctgtt ctcagcagtt cttcaggtc
   -> E2 bind
7881 gaaaagccaa aacatTAAatg tACCGAATT GGTacaatta cctcaacttt tgcacgtatat
<- L2 end
7941 tccagagatg tttgcttact ctgactaagt ctaactactc caaggaaca gACCGGTCA carracccct ctcctgtcgtg atcttc gacttgcac ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
<- L1 end
7701 GGTGACATAAA ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7761 GGTGACATAAA ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7821 GGTGACATAAA ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7881 GGTGACATAAA ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
<- L2 end
7941 tccagagatg tttgcttact ctgactaagt ctaactactc caaggaaca gACCGGTCA carracccct ctcctgtcgtg atcttc gacttgcac ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
<- L1 end
7701 GGTGACATAAA ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7761 GGTGACATAAA ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7821 GGTGACATAAA ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
7881 GGTGACATAAA ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
<- L2 end
7941 tccagagatg tttgcttact ctgactaagt ctaactactc caaggaaca gACCGGTCA carracccct ctcctgtcgtg atcttc gacttgcac ggtgtaaattg tGGATTTGCGt tgtggcaagc ggcccatggt aaggtctatc
<- L1 end
HPV-14d

**LOCUS**  HPV14d  7439 bp ds-DNA  **VRL**  04-OCT-1993

**DEFINITION**  Human papillomavirus type 14D (HPV-14D), complete genome.

**ACCESSION**  X74467

**SOURCE**  Human papillomavirus type 14D DNA.

**REFERENCE**  1  (bases 1 to 7439)

**AUTHORS**  Delius, H. and Hofmann, B.

**TITLE**  Primer-directed sequencing of human papillomavirus types


**REFERENCE**  2  (bases 1 to 7439)

**AUTHORS**  Delius, H.

**TITLE**  Direct Submission

**JOURNAL**  Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG

**COMMENT**  Sequence has a deletion of about 274bp (by comparison to closely related HPV25) at the HindIII cloning site (pos 622 to 627). The plasmid clone sequenced contains at this position a 368bp fragment from HPV15.

HPV-14 was isolated by Kremsdorf et al. in 1984 (J Virol 52: 1013-8) and subsequently sequenced by Dr. H. Delius. It has been associated with the flat wart-like lesions of EV, a multifactorial disease, and infrequently with malignancies. HPV-14 is considered to be part of the α2 cluster based on phylogenetic analysis. This cluster includes HPV-19, HPV-25, HPV-20 and HPV-21, in addition to HPV-14. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

Due to an apparent deletion of approximately 300 bp, the E6 and E7 ORFs of HPV-14d seem to be disrupted. Based on its homology with other HPV types, the ORF appearing in the sequence from bp 82 to 744 is similar at its 5' end to the beginning of E6, and at its 3' end to the end of E7. On this basis, we have chosen to assign bp 196-627 to E6 and bp 628-744 to E7.

**BASE COUNT**  2337 a  1432 c  1612 g  2058 t

**ORIGIN**  195 bp upstream from beginning of E6/E7 fused cds

1 AACGGTaagt tatctgcAC CGGGTGCGGT cactgtatta ctcactatgt ggttgttgtt

E2 bind (end)  ->  E2 bind

61 gccaactacc attgtctgaTA Gctgtttttt gctgttaacg ttatcgacac atacatatct

E6 orf start  ->

121 atgtatatat atatatatat atatatatat atatatatat atatatatac cagaaaaac
181 agagaatgca gactcATGgc gacaactgac tcttaacag agaggtcct
E6 cds ->
241 tctcctaaga gtaactattg tgatagcaca gaagaacaaat ctcttttat agagaatgca
301 ttaacctgca aatatttgg taaaatatgg actataattg aatggtgtg ataattaattg
gacagtgcaga tgaaggtcct
gataagagtt taacctgcaca atatttttttt tactttttaga gataattttt
gacaacagac caacggtgaa gttataataa tttttgtaga gatactgttga gacagacagaa
541 atagagagtt taacctttgaa atagagttta atagagttttag atagagttttag aatattttttt
gacaacagac caacggtgaa gttataataa tttttgtaga gatactgttga gacagacagaa
941 creating a fused E6/E7 orf putting a deletion of approx. 280 bp /
E1 orf ->
181 agagaatgca gactcATGgc gacaactgac tcttaacag agaggtcct
E6 cds ->
241 tctcctaaga gtaactattg tgatagcaca gaagaacaaat ctcttttat agagaatgca
301 ttaacctgca aatatttgg taaaatatgg actataattg aatggtgtg ataattaattg
gacagtgcaga tgaaggtcct
gataagagtt taacctttgaa atagagttta atagagttttag atagagttttag aatattttttt
gacaacagac caacggtgaa gttataataa tttttgtaga gatactgttga gacagacagaa
541 atagagagtt taacctttgaa atagagttta atagagttttag atagagttttag aatattttttt
gacaacagac caacggtgaa gttataataa tttttgtaga gatactgttga gacagacagaa
941 creating a fused E6/E7 orf putting a deletion of approx. 280 bp /
E1 orf ->
181 agagaatgca gactcATGgc gacaactgac tcttaacag agaggtcct
E6 cds ->
241 tctcctaaga gtaactattg tgatagcaca gaagaacaaat ctcttttat agagaatgca
301 ttaacctgca aatatttgg taaaatatgg actataattg aatggtgtg ataattaattg
gacagtgcaga tgaaggtcct
gataagagtt taacctttgaa atagagttta atagagttttag atagagttttag aatattttttt
gacaacagac caacggtgaa gttataataa tttttgtaga gatactgttga gacagacagaa
541 atagagagtt taacctttgaa atagagttta atagagttttag atagagttttag aatattttttt
gacaacagac caacggtgaa gttataataa tttttgtaga gatactgttga gacagacagaa
941 creating a fused E6/E7 orf putting a deletion of approx. 280 bp /
E1 orf ->
181 agagaatgca gactcATGgc gacaactgac tcttaacag agaggtcct
E6 cds ->
241 tctcctaaga gtaactattg tgatagcaca gaagaacaaat ctcttttat agagaatgca
301 ttaacctgca aatatttgg taaaatatgg actataattg aatggtgtg ataattaattg
gacagtgcaga tgaaggtcct
gataagagtt taacctttgaa atagagttta atagagttttag atagagttttag aatattttttt
gacaacagac caacggtgaa gttataataa tttttgtaga gatactgttga gacagacagaa
541 atagagagtt taacctttgaa atagagttta atagagttttag atagagttttag aatattttttt
gacaacagac caacggtgaa gttataataa tttttgtaga gatactgttga gacagacagaa
941 creating a fused E6/E7 orf putting a deletion of approx. 280 bp /
6841 ttctaatatt ttagaggagt ggcaattagg atttgtacct gcaccagaca atcttatcca
6901 tgatacatac agatatattg agtctgcagc gactaggtgt cctgataaaa atctctctaa
6961 agaaagagaa gatctctata aaaactttta cttttggaat gtgagaattc cagagagact
7021 atctttagac ctagatacag attctctttg aggaaatatt ttatatttccg caggtttgca
7081 gcaattcagCC GTTAACGGTa caaaaacagt ttgccactagg ggtccatatc aaggtatattaa
-> E2 bind
7141 acaaaaacgc aagatTAGa cattatcagat ttcgtgtcaca taaaagtcaac tttacacag
<- L1 end
7201 tattcaagga atgttttattc acctctgacta agcaatatag agccgcgccg gatacataaa
7261 gcttggccaaat gaggtgaatt gtttgcaga agagtcgaga gccaactcag gtttgccgca
7321 gatcaqatac agqqcgcqaccc gcqgttggacgc aagctacatc gttctgaacac gcaaaagact
7381 caaggaatgtc aaggtgtgcc agtctattttt gtccgaattt ggcaagttt gaaACCGTT
-> E2 bind
(start)
HPV15

**LOCUS**
HPV15  7412 bp ds-DNA

**DEFINITION**
Human papillomavirus type 15 (HPV-15), complete genome.

**ACCESSION**
X74468

**SOURCE**
Human papillomavirus type 15 DNA.

**REFERENCE**
1 (bases 1 to 7412)

**AUTHORS**
Delius, H. and Hofmann, B.

**TITLE**
Primer-directed sequencing of human papillomavirus types

**JOURNAL**

**REFERENCE**
2 (bases 1 to 7412)

**AUTHORS**
Delius, H.

**TITLE**
Direct Submission

**JOURNAL**
Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG

**COMMENT**
HPV-15 was isolated by Kremsdorf et al. in 1984 (J Virol 52: 1013-8) and subsequently sequenced by Dr. H. Delius. It has been associated with the benign flat wart-like lesions of EV, a multifactorial disease. HPV-15 is considered to be part of the b cluster based on phylogenetic analysis. This cluster includes HPV-15, HPV-17 and HPV-9. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

**BASE COUNT**
2374 a 1321 c 1611 g 2106 t

**ORIGIN**
199 bp upstream from beginning of E6 cds

1 caagtaacct ggccagacac tttcttgga gacagccACCG ATAACGGTaa gattatatct
  -> E2 bind

61 ttgAGCGT GT cttgatggt ttggctgata gtatgtaaca acaatcactc
  -> E2 bind

121 tttaaaatata ttagtaACCG CCTGCGTTac tttatttaat ctacatacaa tatgTGAggt
  -> E6 orf start ->

181 aacaacttta gagagctatA TGgataggcc aaagctttt tctgtgcagc agcttgcaga
E6 cds ->

241 cactcttggt atccctttag tagatataatt attgcccttg atagttttgc agagattttt
301 acatatata gaattagtcg taaaggtctg cagtaaatgt ggaatgagga
361 agatttttgt tttgcctgtt gttctagttg tgcatttgct acagcgcagt ttgaattttc
421 taacttttat gaacagtcgg tgtgtagttg ggaaatagag atagttgagc agaagctgt
481 tagagattag gacaattttg gcacattttc cgcattttg cgcattttg aagaggattg
541 agatattttg tacaagaggg aggaattttc caagtttgg cagaattgc cggacattttg
tattgggtt tttgcctgtt gttctagttg tgcatttgct acagcgcagt ttgaattttc
601 ttagagatttgc gcacattttg gcacattttg cgcattttg cgcattttg aagaggattg
tattgggtt tttgcctgtt gttctagttg tgcatttgct acagcgcagt ttgaattttc
E7 orf start -> E7 cds ->

E7 end

661 gacgctgcaag agctgttcca gcctcactgac ttgctattg tgtgatagtgc acgaagattttt
tatttttcat gcctctgttg ggtgtgggttg aatccttttt ggtgttgagtt ggtgttgagtt
tatttttcat gcctctgttg ggtgtgggttg aatccttttt ggtgttgagtt ggtgttgagtt
tatttttcat gcctctgttg ggtgtgggttg aatccttttt ggtgttgagtt ggtgttgagtt
tatttttcat gcctctgttg ggtgtgggttg aatccttttt ggtgttgagtt ggtgttgagtt
901 TGATgacaaa ggtagatattg atctataaagag atgttttttt ggtgtttttt ggtgtttttt
tatttttcat gcctctgttg ggtgtgggttg aatccttttt ggtgttgagtt ggtgttgagtt
tatttttcat gcctctgttg ggtgtgggttg aatccttttt ggtgttgagtt ggtgttgagtt
1081 attagctgcaagtagtgcttt ctggtggggt gggttgtttt ggtgtttttt ggtgtttttt
tatttttcat gcctctgttg ggtgtgggttg aatccttttt ggtgttgagtt ggtgttgagtt
I-H-31

HPV15

SEP 94

1141 tatcagttca caaagaggtt tcagcgttag ccctcgcctg cagtgtatat ctatttcgcc
1161 t9agctcggaa gtttctctgtttaa tgcagcttag ccctcgcctg cagtgtatat ctatttcgcc
1201 acagcataag tctaaaagga gattatttga acaagacagc ggactagaac tatcatttaa
1261 agaacccctt atgtgagtat attatagttg ctttaatggt gtaaagaagt gaaagacagt
1281 attagaattctt gattgacatc atctggaaata aaaaataaag cagaaagact gaaagacagt
1321 tgaagctcaa gattttactc agcagacttt ggaggtaccg gcgaccgatg ttgtgccgca
1381 gggtgccaag ggactgggca ttgttaaaga tcttcttaaa tgtaacaatg tgaaagcaat
1441 gttattagct aaatttaaag aagcgtttgg agtggggttt atggaattaa ctagacaata
1501 taaaagcagc aaaacatgct gtagagactg ggtactgact gtttatgctg tacaagatga
1561 gctgttagaa agttctaaac aattgttgat tcaacactgt gcatatattt ggttacatca
1621 aataccccct atgtgcttat atttattgtg ctttaatgtt ggtaaaagta gagaaacagt
1681 attaagatta cttacgaatt tgttacaagt atctgaaata caaataatag cagaaagacagt
1741 ttaagctat caggaagctt ttcctacttc tattttctatc taattattttta aaaaagttgt
1801 tagacaactg agacggggag aagctcactgt cagcattatt ttggtataaa ggaagtatga atccaaatgt
1861 ttatgctcat ggagaatatc ctgagtggat aatgacacaa acaatgataa atcaccaaac
1921 agacgagctt ttcacagaag aagctcactgt cagcattatt ttggtataaa ggaagtatga atccaaatgt
1981 gaaatctttt tttgaaaggc tttggagaca gtTAGagctc agtgatcaag aagacgaggg
2041 gttattgggt gttgaaggag aaggacattg gtcagaatgt gtaaagtttg tcagatacca
2101 aataagctt cagagtagctc gtcataaactg cagactaatg caataaacgtt ttttacataa
2161 aaaaagttgt atgttaatat aagcgtttgg agtggggttt atggaattaa ctagacaata
2221 attaataaaa gttctcaaag gtaaaggttt gtcattttca aattaaataa gtaaaggttt gtcattttca
2281 gttcattgtt ttttacataa attttaaatg gtaaaggttt gtcattttca aattaaataa gtaaaggttt gtcattttca
2341 gggtataaaa gttcattgtt ttttacataa attttaaatg gtaaaggttt gtcattttca aattaaataa gtaaaggttt gtcattttca
2401 gaaacacagg gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
2461 tatcattact aatagagcctg catattgact ctagacaata ctgagtcattt ttttacataa
2521 acataagtt ctttattgta ggtgaaattt cttcattaac atgctaactt ctaacattga
2581 gaaatctttt tttgaaaggc tttggagaca gtTAGagctc agtgatcaag aagacgaggg
2641 agacgATGga tactctcagc gaacgtttca atgcactaca agagaatcta atggacattt
2701 aTGAgtcagg tcgagacgac atagaaactc aaatattgca ctggcaatat tttgaggcaag
2761 aacaagtatt attctatttt gccaggaaac atggagttat gcgtttagat atcagcctgt
2821 acccccttTA Gccaccagtg agaccaaagc gaaagatgct attggtatgg tgattctgtt
2881 aacaagtattt attctatttt gccaggaaac atggagttat gcgtttagat atcagcctgt
2941 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3001 ttttacataa attttaaatg gtaaaggttt gtcattttca aattaaataa gtaaaggttt gtcattttca
3061 ccacagttaa GATGacaatc gaaacaaggt gaaagagatta atgtaactat atggcgcata
3121 tatcattact aatagagcctg catattgact ctagacaata ctgagtcattt ttttacataa
3181 gaaatctttt tttgaaaggc tttggagaca gtTAGagctc agtgatcaag aagacgaggg
3241 gaaatctttt tttgaaaggc tttggagaca gtTAGagctc agtgatcaag aagacgaggg
3301 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3361 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3421 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3481 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3541 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3601 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3661 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3721 tactgggcga cttgaaagat tactggaaga ggctagggac cccccagTAA tcttgctgcg
3781 cggtgatgct aacaaattaa aatgctttcg ctttagggca aagaaaaagt atcaggattt
3841 aagcgtttca atgcactaca agagaatcta atgtaactat atggcgcata
3901 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3961 aatagagcctg catattgact ctagacaata ctgagtcattt ttttacataa
4021 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
4081 aacaagtattt attctatttt gccaggaaac atggagttat gcgtttagat atcagcctgt
4141 gggtcagaatc acagaagttt attctatttt gccaggaaac atggagttat gcgtttagat atcagcctgt

E2 orf start ->
2641 agacgATGga tactctcagc gaacgtttca atgcactaca agagaatcta atggacattt
E2 cds ->
2701 aTGAgtcagg tcgagacgac atagaaactc aaatattgca ctggcaatat tttgaggcaag
<- E1 end
2761 aacaagtatt attctatttt gccaggaaac atggagttat gcgtttagat atcagcctgt
putative deletion causing premature termination of E2 cds
2821 acccccttTA Gccaccagtg agaccaaagc gaaagatgct attggtatgg tgattctgtt
<- premature termination of E2 cds
2881 aacaagtattt attctatttt gccaggaaac atggagttat gcgtttagat atcagcctgt
2941 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3001 ttttacataa attttaaatg gtaaaggttt gtcattttca aattaaataa gtaaaggttt gtcattttca
3061 ccacagttaa GATGacaatc gaaacaaggt gaaagagatta atgtaactat atggcgcata
E4 orf and cds ->
3121 tatcattact aatagagcctg catattgact ctagacaata ctgagtcattt ttttacataa
3181 gaaatctttt tttgaaaggc tttggagaca gtTAGagctc agtgatcaag aagacgaggg
3241 gaaatctttt tttgaaaggc tttggagaca gtTAGagctc agtgatcaag aagacgaggg
3301 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3361 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3421 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3481 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3541 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3601 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3661 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3721 tactgggcga cttgaaagat tactggaaga ggctagggac cccccagTAA tcttgctgcg
<- E4 end
3781 cggtgatgct aacaaattaa aatgctttcg ctttagggca aagaaaaagt atcaggattt
3841 aagcgtttca atgcactaca agagaatcta atgtaactat atggcgcata
3901 gcgccatgtc aaataaggtt tccaccatta atgctaactt ctaacattga
3961 aatagagcctg catattgact ctagacaata ctgagtcattt ttttacataa
probable E2 end
4021 gtgcttttta atcaactacag ataggtttgttttttatttctcttgct tttgtcct TA A TAA Atgtagtattt tacagcattc aagagtagctc cttctact ta aactcccA TG gcccgtgc acgcagagta aaacgtgcat ctgtaactg cat tcctacttg ttttacataa
L2 orf start ->
4081 aacaagtattt attctatttt gccaggaaac atggagttat gcgtttagat atcagcctgt
L2 cds ->
4141 gggtcagaatc acagaagttt attctatttt gccaggaaac atggagttat gcgtttagat atcagcctgt

I-H-31

HPV15

SEP 94
**LOCUS**  
HPV17  
7427 bp ds-DNA  
VRL  
04-OCT-1993

**DEFINITION**  
Human papillomavirus type 17 (HPV-17), complete genome.

**ACCESSION**  
X74469

**SOURCE**  
Human papillomavirus type 17 DNA.

**REFERENCE**  
1 (bases 1 to 7427)

**AUTHORS**  
Delius, H. and Hofmann, B.

**TITLE**  
Primer-directed sequencing of human papillomavirus types

**JOURNAL**  

**REFERENCE**  
2 (bases 1 to 7427)

**AUTHORS**  
Delius, H.

**TITLE**  
Direct Submission

**JOURNAL**  
Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG

**COMMENT**  
HPV-17 was isolated by Kremsdorf et al. in 1984 (J Virol 52: 1013-8) and subsequently sequenced by Dr. H. Delius. It has been associated with the benign macular lesions of EV, a multifactorial disease, and subsequently from squamous cell carcinomas and the malignant melanoma of an immunosuppressed patient. HPV-17 is considered to be part of the b cluster based on phylogenetic analysis. This cluster includes HPV-15, HPV-17 and HPV-9. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

**BASE COUNT**  
2329 a 1366 c 1670 g 2062 t

**ORIGIN**  
199 bp upstream from beginning of E6 cds

1 ctggagccaa ggttttgccga gaacacaatc ttggaacagc ACCGATAACG GTAagattat  
-> E2 bind

61 atctttggaAC CGTAGGCGGT acctttctgt aggattttgct gatagtagtt aacaacaatc  
-> E2 bind

121 ttctctcata aatatatgtg ACCGCCTTCG TTaccttaaa tgatctacat acaatatgag  
-> E2 bind

181 agctcttact TAAgagcatA TGgataggcc aaaacctcaa acagtgagg agctctgatga  
E6 cds ->

E6 orf start ->

241 taccctttgt attccattag ttagatattt attaccttgag agatttttta ataggttttt  
301 agctctatanca aataggttgg cttttgattttt cagataggtt gatgtaactt  
361 agatattttgt tttgctgttt gcatagtttg tggcttgagc acagcagcag atagtcttt  
421 taaggatttc gaacagcagc aataggttgg cttttgattttt cagataggtt gatgtaactt  
481 agagagatgtg tccttgagtt ttttaaaaag ttgggtttac ttagagatg  
541 agagagatgtg cttttgagtt ttttaaaaag ttgggtttac ttagagatg  
601 cagacattgt gggtgtgatatg AATGAtggg aaaaagacta caatacagaa tattagttc  
E7 cds ->

E7 orf start ->

661 gagctgcaac agctcttggc gccacactcct atgcctggact gacaggtattttt gctctgctttc  
721 gacacacag cagagacagc cttctgttctg gatagttttt ctttggctttc  
781 ttttggcttt gcctttttg gacatcactc tatgatttc aacatggaatg  
841 gagagagatgtg tttgctgttt gcatagtttg tggcttgagc acagcagcag atagtcttt  
E1 orf start ->

901 cATGAcTGAc gacaccaag aacaataaagtttta ggagatgagc tgaagagtt  
E1 cds ->

<- E7 end

**I-H-33**

SEP 94
HPV17

961 
1021 
1081 
1141 
1201 
1261 
1321 
1381 
1441 
1501 
1561 
1621 
1681 
1741 
1801 
1861 
1921 
1981 
2041 
2101 
2161 
2221 
2281 
2341 
2401 
2461 
2521 
2581 
2641 
2701 
2761 
2821 
2881 
2941 
3001 
3061 
3121 
3181 
3241 
3301 
3361 
3421 
3481 
3541 
3601 
3661 
3721 
3781 
3841 
3901 
3961 
4021 
4081 
4141 
4201 
4261 
4321 
4381 
4441 
4501 
4561 
4621 
4681 
4741 
4801 
4861 
4921 
4981 
5041 
5101 
5161 
5221 
5281 
5341 
5401 
5461 
5521 
5581 
5641 
5701 
5761 
5821 
5881 
5941 
6001 
6061 
6121 
6181 
6241 
6301 
6361 
6421 
6481 
6541 
6601 
6661 
6721 
6781 
6841 
6901 
6961 
7021 
7081 
7141 
7201 
7261 
7321 
7381 
7441 
7501 
7561 
7621 
7681 
7741 
7801 
7861 
7921 
7981 
8041 
8101 
8161 
8221 
8281 
8341 
8401 
8461 
8521 
8581 
8641 
8701 
8761 
8821 
8881 
8941 
9001 
9061 
9121 
9181 
9241 
9301 
9361 
9421 
9481 
9541 
9601 

4321 aagggccctgg ggtacgcgtg ctcggcttcgg ggacatcctg
4381 aacaaaactgg ccagccggtg ctgataacctcttgatctgacagt cgcgcctgcct gatgctgccttg
4441 cacatttagt gcggagcactcttcag cagactaatg gatttctggtgcatgctgcggatctggtgcctg
4501 ccaccctttcg acagccgagatcgatcggacctgcacgctggctgcgctgcgtgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggctgctgggCTT
HPV19

LOCUS     HPV19    7685 bp ds-DNA  VRL  04-OCT-1993
DEFINITION Human papillomavirus type 19 (HPV-19), complete genome.
ACCESSION X74470
SOURCE Human papillomavirus type 19 DNA.
REFERENCE 1 (bases 1 to 7685)
AUTHORS Delius,H. and Hofmann,B.
TITLE Primer-directed sequencing of human papillomavirus types
REFERENCE 2 (bases 1 to 7685)
AUTHORS Delius,H.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG
COMMENT HPV-19 was isolated by Gassenmaier et al. in 1984 (J Virol 52: 1019-23) and subsequently sequenced by Dr. H. Delius. It has been associated with the benign macular lesions of EV, a multifactorial disease. HPV-19 is considered to be part of the α$_2$ cluster based on phylogenetic analysis. This cluster includes HPV-14, HPV-25, HPV-20 and HPV-21, in addition to HPV-19. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

BASE COUNT 2424 a 1460 c 1663 g 2138 t
ORIGIN 199 bp upstream from beginning of E6 cds
    1 GGTaagttta ttgatacggg cgcggttaga agttactcat tcggtgcttg ttgttgccaa
    E2 bind <-
    61 caaatcagcqt tatgaacttg tttctgcctg tatcggtatc gacacaggta ttatatatat
    121 atatatatat atatatatat atatatatat atatatatac acacagatac
    181 attttgcaagc tgcaatacctA TGGctaagcg acaggtcaca gaagagagga tagaaagtgt
    E6 cds ->
    241 agaagagagga actactgcac cacaggtcac agagccacca ttaccagcaa caattgctgg
    301 attagcagca ttgctagaaa taccgttgga tgactgttta gtgccttgta atttctgtgg
    361 caagttttta tcacatttag aagctgcggc atattgaat ataaagatta gtttgaggtg
    421 gaagagagcta ctctgtgtgat cttgctgtcg cttgctgtcg acagaaacgt casacattga
    481 atttcaaggt ttaaatgac atactgtac agtgtagaag attagtttgg taacaggtga
    541 atctgtcctt gcattgctga aaattgctatg ataaatgctt attcaattga
    601 aaagcgtttt gtctgttggaa gaagacctcc ttctcataaa aaTTAAagacat cttggaaagg
    E7 orf start ->
    661 gatctgtagg ctgtgtaagc atttctataA TGattggTAA agaggtgata ttgcaagaca
    E7 cds ->
    721 ttgtattaga atttctataa tgtcagcttg atgacctac ccagaaacag agaggaagat
    781 acacaggttga cttgctgtcg cttgctgtcg atgacctac ccagaaacag agaggaagat
    841 caatgtgtgt ttcagaggta atgtgtgtgt ttcagaggta atgtgtgtgt ttcagaggta
    901 ctcaattgg ttagatccg gttgctgtcg cttgctgtcg atgacctac ccagaaacag
    E1 orf start ->
    961 cggagtccgc tgggaactgc aatcATGgcg ggtcTAAag gtagtcgcttag taaaggaggg
    E1 cds ->
    1021 tttgtgtgg tgtgtgtgtg tgtgtgtgtag tgtgtgtgtg tgtgtgtgtag tgtgtgtgtag
    JOURNAL Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG

I-H-36
SEP 94
1321  acgggactctg  agttgacctt  aacaaatgaa  gctgaagatg  tttcttctga  ggtggaggta
1381  ccggctttag  attctcagcc  ggtagctggg  gaacaatcag  gggacataga  catacatttt
1441  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
1501  agggcttcat  ctgggtctg  ggctcttata  atagctcttt  ttaaatctgt  ggtggaggta
1561  ttcctgatct  cactgtggc  aagcttcttt  attgtacagt  ggccttgaat  aagagtttaa
1621  tgctgctgat  ggtggaggta  cactggtcct  ttgcgcggcg  ggaagttatc  aagagtttaa
1681  ccgtggctgt  tgcgctgtg  acagccggtc  ggccttgaat  aagagtttaa  aagagtttaa
1741  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
1801  agggcttcat  ctgggtctg  ggctcttata  atagctcttt  ttaaatctgt  ggtggaggta
1861  ttcctgatct  cactgtggc  aagcttcttt  attgtacagt  ggccttgaat  aagagtttaa
1921  tgctgctgat  ggtggaggta  cactggtcct  ttgcgcggcg  ggaagttatc  aagagtttaa
1981  ccgtggctgt  tgcgctgtg  acagccggtc  ggccttgaat  aagagtttaa  aagagtttaa
2041  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
2101  agggcttcat  ctgggtctg  ggctcttata  atagctcttt  ttaaatctgt  ggtggaggta
2161  ccgtggctgt  tgcgctgtg  acagccggtc  ggccttgaat  aagagtttaa  aagagtttaa
2221  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
2281  agggcttcat  ctgggtctg  ggctcttata  atagctcttt  ttaaatctgt  ggtggaggta
2341  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
2401  agggcttcat  ctgggtctg  ggctcttata  atagctcttt  ttaaatctgt  ggtggaggta
2461  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
2521  agggcttcat  ctgggtctg  ggctcttata  atagctcttt  ttaaatctgt  ggtggaggta
2581  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
2641  ccgtggctgt  tgcgctgtg  acagccggtc  ggccttgaat  aagagtttaa  aagagtttaa
2701  ccggtgcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
2761  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
2821  agggcttcat  ctgggtctg  ggctcttata  atagctcttt  ttaaatctgt  ggtggaggta
2881  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
2941  agggcttcat  ctgggtctg  ggctcttata  atagctcttt  ttaaatctgt  ggtggaggta
3001  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3061  agggcttcat  ctgggtctg  ggctcttata  atagctcttt  ttaaatctgt  ggtggaggta
3121  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3181  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3241  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3301  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3361  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3421  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3481  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3541  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3601  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3661  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3721  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3781  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3841  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3901  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
3961  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
4021  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
4081  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
4141  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
4201  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
4261  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
4321  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
4381  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
4441  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
4501  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
4561  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat
4621  ttcggtcttt  aatcgctggt  tcttttcttg  tatgtaagtt  aagctcagGC  gagaagtaat

I-H-37
SEP 94
HPV-20 is primarily associated with the benign flat warts of Epidermodysplasia Verruciformis (EV), a multifactorial disease, and infrequently with malignancies. HPV-20 is considered to be part of the a$_2$ cluster based on phylogenetic analysis. This cluster includes HPV-19, HPV-25, HPV-14 and HPV-21, in addition to HPV-20. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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-20 and HPV-47. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

Kiyono et al. (Virology 186: 628-39) analyzed the noncoding region and the E6 gene of several EV associated HPVs (HPV-5, HPV-8, HPV-14, HPV-19, HPV-20, HPV-21, HPV-25, and HPV-47). They observed that in the URRs of all eight of the EV HPVs a conserved region of 29 nucleotides described by Krubke et al. (J Gen Virol 68: 3091-103), M29, was present. Another conserved region of 33 nucleotides, M33, was common to only five of those types studied (HPV-5, HPV-8, HPV-47, HPV-25, and HPV-19). Kiyono et al. established two operational clusters based on amino acid similarity of the E6 genes. The first cluster consisted of HPV-5, HPV-8 and HPV-47; the second consisted of HPV-14, HPV-20, HPV-21, and HPV-25. In addition to genomic analysis, Kiyono et al. also determined the transforming ability of the EV types relative to HPV1a. The results indicated that HPV-47, HPV-5 and HPV-8 had a stronger transforming ability than HPV-14, HPV-21 and HPV25; while HPV-1a showed no transformation. Thus, the operational clusters hold for both genomic similarity and transforming ability.

BASE COUNT 294 a 174 c 215 g 302 t

ORIGIN

1 tttatttact ctagcataa cagatcaac caagccacccga cacataaagg tgaagtgtgtgt
gccaagagag aagctagtgtgg acgcagaa gacacggct tgcggcagat
121 cagatacaact gcgtgcacga gttggctcaa acttcatttc cccaacacgt tcgggaacagg
181 aggaaatgta aggctgccaa cgcttttggc tcttcttttt ggcacagcag aagACCGTTA

E2 bind

241 ACGGTaagtt tttatttgta tcgggcgcgg tcatacatta ctcatttggg agttgtgtgt
tgctgtggaggactaact ataaagcgaacatcgaagct tatagtaacg tgcagtaacct tccagcag
301 gccaagcctc ataaagcgaacatcgaagct tatagtaacg tgcagtaacct tccagcag

signal ->

361 ATATAATAT ATATAATAT ATATAATAT ATATAATAT ATATAATAT atataagatc atataagacag
421 atatacatgct ataaagcgaacatcgaagct tatagtaacg tgcagtaacct tccagcag

E6 cds ->

481 aggcgccgctg aagacactac ataatgttgg gggcgcagtcct atcttttctt tccagcag
541 tcggcgcagtcct atcttttctt tccagcag

I-H-39
SEP 94
601 ataccttgta acttctgcgg taatttcctt acacatttag aagtttgtga gtttgatgag
661 aagaagctta cttaaatttt gaaagatcat ttggtttttg catgctgtcg tgtttgtgc
721 tcggcaacag cgacatatga gtttaatcaa ttttatgaga gtactgttttt aagcagagac
781 atagagcaag taacaggca aatctgttttt gatatacttg tcaggtgcta cacctgtatg
841 aaatttttag actcaattga aaagcttagac aatctgtggca gaaagcttagc atttttatita
901 gtgagaggtt cttgaaaggg aatctgtatgg cttgtgtaagc attttcaaaTA Atgattgtta
<- E6 end

961 aagaggtcac attgcaagat attgt

//
HPV-21 has been associated with the flat wart-like lesions of EV, a multifactorial disease. HPV-21 is considered to be part of the a$_2$ cluster based on phylogenetic analysis. This cluster includes HPV-19, HPV-25, HPV-14, HPV-20 and HPV-21. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

Kiyono et al. (Virology 186: 628-39) analyzed the noncoding region and the E6 gene of several EV associated HPVs (HPV-5, HPV-8, HPV-14, HPV-19, HPV-20, HPV-21, HPV-25, and HPV-47). They observed that in the URRs of all eight of the EV HPVs a conserved region of 29 nucleotides described by Krubke et al. (J Gen Virol 68: 3091-103), M29, was present. Another conserved region of 33 nucleotides, M33, also described by Krubke, was common to only five of those types studied (HPV-5, HPV-8, HPV-47, HPV-25, and HPV-19). Kiyono et al. established two operational clusters based on amino acid similarity of the E6 genes. The first cluster consisted of HPV-5, HPV-8 and HPV-47; the second consisted of HPV-14, HPV-20, HPV-21, and HPV-25. In addition to genomic analysis, Kiyono et al. also determined the transforming ability of the EV types relative to HPV1a. The results indicated that HPV-47, HPV-5 and HPV-8 had a stronger transforming ability than HPV-14, HPV-21 and HPV25; while HPV-1a showed no transformation. Thus, the operational clusters hold for both genomic similarity and transforming ability.
HPV21E6

601 ccttgtaact tttgcggtaa tttctttact catttagaag tttgtgagtt tgatgagaaa
661 aagcttagtt tactttggaagacatcattgtgtgtttgctggtgtcgtggttgttgca
721 gcaacagcga catatgata taatgaattt tatgaatcta ctgttgtagg tagagatata
781 gaagaataaa caggcaaatc tattttttgatatgtgtca ggtgctacac ttgcagaaaa
841 ttttttagact caatagaaaa gctagatatt tgtggtagga agcatttttt tcataaagtgt
901 agagccattt ggaaaggaac tgtagggcgtgctgaagcttc ttcAAAtg atttgttaaag
  <- E6 end
961 aggtacatt gcaagatatt gt
DEFINITION  Human papillomavirus type 25 (HPV-25), complete genome.

ACCESSION  X74471

SOURCE  Human papillomavirus type 25 DNA.

REFERENCE  1 (bases 1 to 7713)

AUTHORS  Delius,H. and Hofmann,B.

TITLE  Primer-directed sequencing of human papillomavirus types


REFERENCE  2 (bases 1 to 7713)

AUTHORS  Delius,H.

TITLE  Direct Submission

JOURNAL  Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG

COMMENT  HPV-25 was isolated by Gassenmaier et al. in 1984 (J Virol 52: 1019-23) and subsequently sequenced by Dr. H. Delius. It has been associated with the benign macular lesions of EV, a multifactorial disease. HPV-25 is considered to be part of the a$_2$ cluster based on phylogenetic analysis. This cluster includes HPV-14, HPV-19, HPV-20 and HPV-21, in addition to HPV-25. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. These types are rarely detected in lesions afflicted with 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 shorter compared to the viruses in other groups and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

BASE COUNT  2357 a 1529 c 1699 g 2128 t

ORIGIN  199 bp downstream from beginning of E6 cds

1 TAACGGTaaag tcataatgata cgggccgcgt taaatatta tcattctgta ctttgtgctg
 E2 bind <-
 61 ccaccaatca gcattagtaa cttgtttctg cctgtatcgt tatcgacaca ggtgtgttat
 121 acatatatat atatatatat atatatatat atatatatat tataactagta ctaagctaga
 181 cactgcacga ttaggactTATT GGGcaactgc aaatgctgaa cagagcatag gaccaccaga
 E6 cds ->
 241 gcacacgcag gttatacagc caccattgcc agcaacaatt actgatctag cagctttatt
 301 ggaaattcca ttagatgatt gcttagtacc ttgcaacttc tgtggcaact ttctaacata
 361 tgttctgtct ggcacctgtt gtcgagcagc aactgaccaca aacgcragag aatagacggaa
 421 gtagacgctc gtcgctgttt tgtgcgttga aactgctgag tgtgggaag aatagctgctt
 481 cagtgtagag tacaggttat gacagttgaa agaaggtgtt gatgtctgag tgttttgacat
 541 agatgttaga tgtcaaaaac gctatggaata tttctatgca aagttgaaag tttatgatgtt
 601 tggcagaga cgtcccttccc atctagTAAG aggctctttg aaggaactc tgagctctgtg
 E7 orf start ->
 661 tcagatattc tataATGatt ggTAAGagg tcacattctgca agatgttttaca ttagatgtaa
 E7 cds -> <- E6 end
 721 tgtgaattca gcctgagctg tcaaccagttt acctgttttgg tgaagagggag ggcaggctcg
 781 gacactcggga aacagaggg dgccttgcta ttagacagag tttacacaggt gttgctgacg
 841 cttgctgtc gtggcaggtc aagctcggcc ttcttggga gacagctgtac ttggtgtatt
 901 gaacactcag aaaaacttccc taagctTAAG acgctgtctt tgaaggtgg aatggagatc
 E1 orf start ->
 961 tggcagaaag aatggtgaat tcacagttc ttaagatgatc ttagatgtaa
e1 cds -> <- E7 end
 1021 tgttgtaga gcgtttgtag ttgtagtgatg aaggtgttaa acctacactt aatggagtgtg
 1081 aacagactctc tttattcaga cctttatgct gaaagggcc acattcag
t1041 gaacttttc acattacag gttgagagag tggaggttga aatggagatc
 1201 aagatcttttc gttctgtcttt gcgtggtggt gcgtgtggtt gattcagaag tatattttttt

HPV25
1261 tcacctcagc agaagtctaa acgaaggctc tttgcagagc aggacagcgg gctcgaattg
1321 actttaacaa atgaagctga agatgtttct cctgaggtgg aggtaccggc tttaaactct
1381 cagccggtag ctgagggaca atcaggggac atagacataa gttatacagc attattgcgt
1441 gccagcaata ataaagcaat attaatggca aaatttaaag aggcttttgg ggtagggttt
1501 aatgatctga cacgtcaatt taagagttac aaaacctgtt gtaatgcttg ggttatttct
1561 gtatatgcag tgcatgatga ccttatagaa agttcaaagc agcttttgca acagcattgt
1621 gactatggtg ggatccgtgg gataggagcc atgtcattgt ttttagtttg ttttaaggcg
1681 ggaaaaatc gttgactgtc tcataaattg atgacaacta tgttaaatgt gcatgaaaag
1741 caaatattat cggaaccacc aaaattaaga aatctgcttt ctcgacacta attcactaat
1801 ggtcaatgg gttccggagt atttacatat ggctcatatc cagattggat agcccaccaa
1861 acaatattgg gccatcaaag cgctgaagct agtacatttg atctatcgga catggttcaa
1921 tgggcatttg atataactaa tttagacgaa gcagatattg catatcaata tgctaaattg
1981 gcgccagaca atagcaatgc tgttgcatgg ctcgcacata ataatcaagc caaatttgtg
2041 cgagaatgtg catccatggt gagattttat aaaaagggtc agatgaaaga aatggtactt
2101 tctgtaagga ttttatataa aatctgcttt ctcgacacta attcactaat
2161 2701 Agtgaccaag aagacgaggg cgaaATGga gaatctcagc gagcgtttca atgttctaca
2221 2761 agatcagcta atgaacattt aTGAaactgc agcacaaacc cttgaggcac aaattgagca
2281 2821 ttggcagatttt ttgcgaagag aagctgtgct actatatttt gctaggcaaa agggtgttac
2341 2881 acggcttgga tatcaacctg tacctgcctt aatggtgtct gaagcaaaag ctaaggaagc
2941 2941 tatagggatg gtgctgcaac tgcagtcatt acaaaagtct gaatttggaa aagagccatg
3001 3001 gtcattggtt gacaccagta cagagactta taaaagtcca ccagaaaacc atttcaaaaa
3061 3121 catgtggaga tatatttatt atgtggatga tgatgacaaa tggcataaaa gtgcaagcgg
3181 3181 gagcagcagg gggggccgcg gcaggggcgg gagacgacga cacagattgt ctgaatcccc
3241 3241 cactgtgttt actcctgtca ccagctccac cccccccgag tcaccaggag gacaagcaga
3301 3301 caactgtgttt actcctgtca ccagctccac cccccccgag tcaccaggag gacaagcaga
3361 3361 ctcaacaccct cccttccaag cccccccaa tcacaccgag tcgccgcgac cggccagaga
3421 3421 ctcggcagaa cggctcgcac aaaaaacgag tcacccgggc gcggggcgcg gcggggcgcg
3481 3481 cagcagcagc cagcagcagc aaaaaacgag tcacccgggc gcggggcgcg gcggggcgcg
3541 3541 gtcggtcgtg cgtggtcgtg cgtggtcgtg cgtggtcgtg cgtggtcgtg cgtggtcgtg
3601 3601 gtctgaattt tgccagttct cgcggtccgg gggggccgcg gcaggggcgg gagacgacga cacagattgt ctgaatcccc
3661 3661 qagcagcagq qagcagcagq qagcagcagq qagcagcagq qagcagcagq qagcagcagq
3721 3721 acaccttcacct ccctctccct cgggtccgcc cgcggtccgg gggggccgcg gcggggcgcg
3781 3781 gactgtagag gagcagcagq qagcagcagq qagcagcagq qagcagcagq qagcagcagq
3841 3841 cagcagcagq qagcagcagq qagcagcagq qagcagcagq qagcagcagq qagcagcagq
3901 3901 cagcagcagq qagcagcagq qagcagcagq qagcagcagq qagcagcagq qagcagcagq
3961 3961 atgggaagag gctcttgatc ccccagTAAt tttagttaga ggggatgcta acacgcttcg
4021 4021 aagctttcgc aacagggcaa agcatatatg tactgggcta tttagctcat ttagtacgca
4081 4141 cttttatcctt aacactgata ttagaacttt acataatttt aagctttcgc aacagggcaa agcatatatg tactgggcta tttagctcat ttagtacgca
4141 4201 tgaccggtcct ttggtatcat ttagaacttt acttagttt ttagaacttt acttagttt
4501 gaggaccaac aggtatatga cctttgggag aaggtccgat acgtgttggt ggaACCCCCA
     -> E2 bind
4561 CGGTTagg acctctttta gttccacagca ctatgtagac atctgacata atacgctgtg
4621 aacctcttaa tccagtggaa cccacatcct cgtctatgtg cccacactaca gacgttccag
4681 gtctctgcct gtctctgtgtg gaggctggaga caatgcgca aatatcctt ggcgcctgttg
4741 taccctccac tcgacaccccg tggacacacac cttctcagaggt ggcagcaagct tttctgacag
4801 tgcacccgag gcccaccccg ccctccctgg ctcagagttg ggaggcaacat tataaatttc
4861 catcatttca ggtctataatc gaaatctcag tctgacagaa tttgatcctgacg
4921 acattcctggt tactcctgggt tcggagggac aacaaattgg ggccatggcc atgtgacctaa
4981 tggtaactca aaggttttccccc actgctttttt cattgtaaatgattgacctgca acaccccaag
5041 gacaaatgac tacactctctt caaaggaatta gagctgtcatt ggggatttaaa
5101 caaatagcag attggtgcac ggttaatggtg catttattcct tttctgacag
5161 cacgtctagct acggtttcag tttgacatct cagtattttt ggtatcttctggtct
5221 ttaggcagga ttttaaatggt tttcagggac ccctggtgat cttgacacaa taatcttatct
5281 catcgaagta gcacaccccg cctgcgctttg tggcattgctt gaggagaaattt ccgggtgaggag
5341 gctcaagagct taccctcgcc actcgtgactt ggtgctcaag gaaatattttt taatcttatct
5401 atagagactt aactgcgatata aactctatgg actcaattgatt gttgctgca ggtgcttctggtct
5461 atcttgggtta tgtcataaatc ctgcaagacc gcacaccccg ccatccctca taatattttt
5521 tagatgaaaa tccattagct gaagatttta gttatttctgc actgctctgtg attttactttt
5581 tggatgaagct ttaatgagat ttgctgtttt accactctctt ctttctgacag gcacaccccg
5641 ctgttgccag ggtacaaagc acggatgaat atgtgcaaag aactaacatct attatatttc
5701 cctatgtgat cccgcttatt tctgttggag acccatattt taatgtgtac cgggttacttct
5761 gctctaaatt gcaaattcca aaagtgtcag gaaatcaaca cagagttttt aggttaaaat
5821 attttactcc cagctctgact acaaggggcc gacgcaaaag aaatattttta TAAatattttt
     <- L2 end
5881 ttttacagAT Ggcagtttgg caagcagcta gtggtaaagt gtaattttttt
5941 ctgttgccag ggtacaaagc acggatgaat atgtgcaaag aactaacatct attatatttc
6001 cctatagtaag gcacaccccg ccatccctca taatattttt
6061 gctctaaatt gcaaattcca aaagtgtcag gaaatcaaca cagagttttt aggttaaaat
6121 taccattcgc actcgatctg gggctcaaat aggctcacaa gttcattttt
6181 atagagactt aactgcgatata aactctatgg actcaattgatt gttgctgca ggtgcttctggtct
6241 gctctaaatt gcaaattcca aaagtgtcag gaaatcaaca cagagttttt aggttaaaat
6301 ttgagcagga tttaaatgat tttcaggagc ctcctgacag ggatttttta gatattcgct
6361 cattaggaag gccacaatat tcctgaaatc tggctgtgctg tggacacacat cgctatctttt
6421 gcaatattca aacaagggca tggcattgctt ggggatttaaa
6481 gagtattttt ttaattttttt gcctctttttt tttatgcacg gcgggaacaa tgttatgcca
6541 ctgcttgatg gatttttttt ggtagtacag cctggctgctt gttgctgata ggtgctgata
6601 aaggaagcat gaaaaatgca ttttacatac cacctaacag tagtcaggct caatataata
6661 atctaggtaa ctcaatgtat ttcccaacag tcagctgctc attggtatcc agtgatgctc
6721 aattattcaag ttagtgactt gctttgctttt tttatgcacg gcgggaacaa tgttatgcca
6781 aatggcaatt aggattcgtt cctgcaccgg acaattctat tcaggatact tatcgctaca
6841 atagagactt aactgcgatata aactctatgg actcaattgatt gttgctgca ggtgcttctggtct
6901 ggtttaatca gctatttgtc actgtggtag acaacacacat cagtttactt gccagaacaa
6961 caatcaattc agatggaaca gatgtttcca aaatcactga ttataattct caaaaattta
7021 cagaatattt gagacatgta gaagaataag agttatcatt aattatacaca ctttgcaacag
7081 taccattgcag gttacaaagc acggatgaat atgtgcaaag aactaacatct attatatttc
7141 attggcaatt aggattcgtt cctgcaccgg acaattctat tcaggatact tatcgctaca
7201 tgtgttcttt tgcattgcgtt gttttgtttt tttatgcacg gcgggaacaa tgttatgcca
7261 aacactttggt gattgtagtt taaagcaagct cttttctttta gatttttgac ggggaagcaag
7321 aatatttttt ggttgttcttt tttatgcacg gcgggaacaa tgttatgcca
7381 GTacaaaaac agtttctctcc cgaatatcct tctggggaat aaaaaaaaga cgttaaaaAT
7441 AGaattacc gtttttcgata caataagctc aactctttgc caagtatcaaa gattgtttaa
     <- L1 end
7501 ttttgtgctga tctacgtagaa tccacctgcgtg aacccgataca taaaggtgg gttatgtgcc
7561 aacgctata cactctgctgcc aacaaattttt ggtgcttctt tttcttctttta gttttggcagagctgctgatgttgaactaagttccttttta aatgtgacctaa ctgacacacac tctgacacat gtcgctgttgtgtaagtttgcgctttttttt
HPV47

LOCUS HPV47 7726 bp ds-DNA VRL 15-SEP-1990
DEFINITION Human papillomavirus type 47 (HPV-47), complete genome.
ACCESSION M32305
SOURCE Human papillomavirus type 47 DNA isolated from scrapes of a benign lesion of a patient suffering from EV and skin cancer.
REFERENCE 1 (bases 1 to 7726)
AUTHORS Kiyono, T., Adachi, A. and Ishibashi, M.
TITLE Genome organization and taxonomic position of human papillomavirus type 47 inferred from its DNA sequence
JOURNAL Virology 177, 401-405 (1990)

HPV-47 is primarily associated with the benign lesions of EV, a multifactorial disease, and also has been detected in cases of malignancy. HPV-47 is considered to be part of the a5,18 cluster based on phylogenetic analysis. This cluster includes HPV-5, HPV-8, HPV-12, and HPV-47. Patients with EV tend to have depressed cell mediated immunity. In roughly one-third of EV-associated HPV infection, the sun-exposed flat wart-like or macular lesions transform into malignant squamous cell carcinomas. 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. 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 and contains two EV-specific regulatory regions: M33 and M29, both shown to be involved in protein binding.

HPV-47 has been isolated from the scraping of a benign lesion of a patient who had suffered from both EV and skin cancer. The DNA was subcloned into vector pT218R and the complete sequence was determined. Recently Kiyono et al. detected HPV-47 by PCR in small clusters of malignant cells in the lower dermis of the same patient. Also a splice donor/acceptor pair which if used can result in a E1/E4 fusion product is present in this genome. Both HPV-5 and HPV-8 also have this ability.

BASE COUNT 2369 a 1517 c 1727 g 2113 t
ORIGIN 207 bp upstream from beginning of E6 cds

E2 bind <- E2 bind ->
61 tgttgccaac taccatagtc ataatcaagt tcttgcctgt atcgttttcg taccttacct
121 acagtatttt atattaaTAT ATAAataaat aaaTATATAA atgtgtattt atttctcagg
E6 orf start -> signal -> signal ->
181 ctcagttctt tgcaattata aacagactATG gctcagaagg ctttggaaca gactacagtT
E6 cds ->
241 aaaaagaaa agctagacagct atacactaat attagaggct tagctcaatt atgttagacat
301 cctttttagag atgtttggct accttgccaa ttttgtggca gtttctttagc tattttgaa
361 gtttgggtaa ttgattatga aacagactaat atcagctctg tattttgac gtttctttagc
ttttttggac gtttctttagc ttttttggac gtttctttagc
421 tgttcttcat tgtgttcttt gaaactactgc acatcatagta tattttttgt tttctttgaa
481 agctttttag gtaatagatat taaatagatat aacagcctct ctttcttttgag tattttttag
541 aggttctata ccttctctgtc atttctttgac attttttagaa atcattctctgcttgtggaaga
601 gcaactctcct ttacacaaagT AAgaaacgcc tggagggggtc ttgtagaggca gtttaagcat
E7 orf start ->
661 ttttttacaATG attggTAAag aggtcaccct gcggaggttt gtttagaggt E7 orf start ->
241 aaaaagaaa agctagacagct atacactaat attagaggct tagctcaatt atgttagacat
301 cctttttagag atgtttggct accttgccaa ttttgtggca gtttctttagc tattttgaa
361 gtttgggtaa ttgattatga aacagactaat atcagctctg tattttgac gtttctttagc
ttttttggac gtttctttagc ttttttggac gtttctttagc
421 tgttcttcat tgtgttcttt gaaactactgc acatcatagta tattttttgt tttctttgaa
481 agctttttag gtaatagatat taaatagatat aacagcctct ctttcttttgag tattttttag
541 aggttctata ccttctctgtc atttctttgac attttttagaa atcattctctgcttgtggaaga
601 gcaactctcct ttacacaaagT AAgaaacgcc tggagggggtc ttgtagaggca gtttaagcat
E7 orf start ->
961 caaacATGgc ggattcTAAA gGTagtacat ctaagaaggg gtttgggtat tgggttatttt
(E1/E4 fused orf) 5' sj /
E1 cds -> < E7 cds end
1021 ttgaaagtctg ctgtagtgat gtggagggat atttggcaac aattttttgg aagagatcag
1081 actcagatgt ccggacctcttg ttagacaatt gtgcacttgtg atccgggcat tcccggaacc
1141 tatattcata ccagaggtgt aagaacaggg aagacagact aacaagactc aaacgaaagt
1201 atcttagtc ccacaggtggt gcgcagctta gtcccgcgtct tgcgacatatt tcatgtgctc
1261 ctcacagcag acaccaagag agggcccttg cagagcaggg cagccgacct gcagtttaacct
1321 ttaaccaatga acgtgaagag tgcctctccgt aggcttgagat aagcgtctata gactcttgcgc
5' sj ] /\n1381 cggtgatctgt tgcaggggatg ataaagacatc cttatgctag ttccttgctt
1441 cccagcaacca aaaaagcctaa ttacctgca taaacctgca aggcttggag gcagctgttta
1501 atgagcagcag aagaaactgg aaaaaagctgtc ctagttagtt tatgtgtatgt
1561 tatattgaag agatccagtt ttagaaagct gcagtttgcgt ttaagacatc tcatgtgctc
1621 ctaaattc aaaaacag ctctttgggt tgccttctgct ttcctctctct ttcctctcgtc
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
1921 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
2161 gaaaaaattt gggtagttttt gggtagttttt gggtagttttt gggtagttttt gggtagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
2221 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
2281 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
2341 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
2401 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
2461 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
2521 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
2581 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
2641 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
2701 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
ttacaaagcg agacatagtt cagcagcttc ccacagttttt gccagctgtt
2761 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
E1 end <-
2821 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
2881 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
2941 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt
ttagtgcgcag cctctgctgc tttatgtgct tcacagttttt ttataagacat
3001 tattatttttgt ggttagttttt ggttagttttt ggttagttttt ggttagttttt ggttagttttt

HPV47

3061 gggccctgtac ctgtggaggt GTatatATGac aaagatgaag caaattgctaa tttgtatact
E4 cds ->
E4 orf start ->
3121 atgtggcatct ttgtgtatata catggatcca gtagatgtgt ggcataaac aacaagttgg
3181 gcaattcaaa ctgaggatcat taacactatat gagaatg gagaatttaa taataagggaa
3241 atctgctagtt ctgcgtgcac tAGctccaca ccaccaaggtg caccaggag cacaacagac
3301 (E1/E4 fused orf) /
3361 ccagacacct cctccaacagg ccaccacacct ccaactcgac ctcgccccaga
3421 ccgcaacctt ctcacaaacct tccacaaaac accgacagac ggcagccagg
3481 agagcccatca gcggcaggcag gggcaacggc gaagctcctgg gcctggaggt
gtggaggggt gttAATAAAg ttgagcaaac signal ->
3541 cgtctgcctg ccctccacact ccaacactct ccaccacacc cccacccaga
gtgagtttttg gccttgtgcgt gggatttttg gagcgccttc
3601 ccctccacta ctggtgatcc cttcaaccaaa cccacagctt gcgggcctgc
3661 ctgggcaagtt gcagctggcttt caagttcgcct ctgctggaggg ctccacaccacc
3721 tcgctgagct gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
3781 tcggctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
3841 ttgctgttaag ctgctgcttt ccgctgttaag ctgctgcttt ccgctgttaag
3901 ggagggcgtt cggctgcttt ccgctgttaag ctgctgcttt ccgctgttaag
3961 tcgggaggtt cggctgcttt ccgctgttaag ctgctgcttt ccgctgttaag
---- E4 end

4021 aacacattaata aatgcttttcgc ccacagacca agggaaacatt atagaggggcat ttttagatca
ttttagatca
4081 ttcgactcata cttttttttgg gggaggttt ccaggggattt ggaggaggtt
4141 atgctgttaag ctgctgcttt ccgctgttaag ctgctgcttt ccgctgttaag
4201 ccaaaaggggt cgggaggttt cggctgcttt ccgctgttaag ctgctgcttt ccgctgttaag
4261 tcgctgagct gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
4321 3' sj

L2 orf -> L2 cds ->
4381 tcgactgctgc aacaggcag cacccttgccc tctggagccc ctcggacgtt gttAATAAAg ttgagcaaac signal ->
4441 aacaggttgt tcgagatcct ctccttctct ttcctctctct ttttagatca
ttttagatca
4501 ttcgactcata cttttttttgg gggaggttt ccaggggattt ggaggaggtt
4561 ttcgactcata cttttttttgg gggaggttt ccaggggattt ggaggaggtt
4621 ttcgactcata cttttttttgg gggaggttt ccaggggattt ggaggaggtt
4681 ttcgactcata cttttttttgg gggaggttt ccaggggattt ggaggaggtt
4741 ttcgactcata cttttttttgg gggaggttt ccaggggattt ggaggaggtt
4801 ttcgactcata cttttttttgg gggaggttt ccaggggattt ggaggaggtt
4861 ttcgactcata cttttttttgg gggaggttt ccaggggattt ggaggaggtt
5041 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5101 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5161 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5221 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5281 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5341 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5401 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5461 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5521 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5581 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5641 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5701 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5761 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
5821 gggcctggag gcgggcctgc cgggagcgcg gtcgctggag gcgggcctgc
HPV49

LOCUS HPV49 7560 bp ds-DNA VRL 04-OCT-1993
DEFINITION Human papillomavirus type 49 (HPV-49), complete genome.
ACCESSION X74480
SOURCE Human papillomavirus type 49 DNA.
REFERENCE 1 (bases 1 to 7560)
AUTHORS Delius,H. and Hofmann,B.
TITLE Primer-directed sequencing of human papillomavirus types
REFERENCE 2 (bases 1 to 7560)
AUTHORS Delius,H.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1993) to the EMBL/GenBank/DDBJ databases. H.
Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenerheimer Feld 506, W 6900 Heidelberg, FRG
COMMENT HPV-49 was originally isolated from pooled flat warts of a Polish renal transplant patient (Favre et al. J Virol 63: 4909 (1989) and subsequently sequenced by Dr. H. Delius. Favre et al. screened benign cutaneous lesions from 134 patients, including 51 immunosuppressed patients and 35 epidermodysplasia verruciformis patients, premalignant cutaneous lesions from 64 patients, and invasive skin carcinomas from 48 patients, for HPV-49 DNA. Despite its similarity to other EV-related papillomaviruses, HPV-49 was detected only in the flat warts of two other Polish renal transplant patients, and in none of the EV patients. HPV-49 forms a remote branch off of the b cluster of the EV- associated HPV types.
BASE COUNT 2366 a 1436 c 1672 g 2086 t
ORIGIN 199 bp upstream from beginning of E6 cds
1 CCACATTGCT Tccagctaca ttttgccgcc aacctcttttg cagcaacacc agaagctatt E2 bind <-
61 ccggtaggtt caatcgggcc cggtcacatt atacttagtc atctcttgtg gttgttaaca E2 orf start ->
121 aacaatcTGA aacagatata catgtaaccg cttggctgtct gtaacttcttt tattcttgga E6 orf start ->
181 aagaatacag acaggacaca TGgtcagACC TGTAAAGGTA tgtgagctag cccaccacct E6 cds -> -> E2 bind
241 aatatatact attttgaggag ttttgcttttc ttttaatattt tgtcagaggt tgttaatata E7 orf start ->
301 ttcaaggttgt ttaagatttt tgcagttcttt tcgagccact ttggaataac E7 cds -><- E6 end
361 cttggctgtt tgtgtagctg atacttagct cagctgattt tagtaatattt tgtgagctag gacaggttaggt gaagcgaggtgt gTACCacatc E7 orf start ->
421 cttggccggc gcgtgatttt tgtgttagatt ttggaatattt gtagctagct ggtagctag E7 orf start ->
481 aagtagtgata tgtctctattt ttggtatttt tgtcaggtttt gcagttctgt TAGccactgt cgaagctggt E7 orf start ->
541 tgcagcagcac aagagactt cttgcgagcc agaagtttcc cggagagcctt cggagagctt gcagttctgt TAGccactgt cgaagctggt E7 orf start ->
601 cagagttata gaATGATggtt cagagtttct cctgtcttttc tgtgagctag cggagagcctt E7 orf start ->
661 gtagctgttc gcagctgttg cagactgttg cagactgttg cggagagcctt E7 orf start ->
721 tgtgttagtt tgcagaggg cgcctttttc tgtgttagtt tgcagaggg cggagagcctt E7 orf start ->
781 tgtgttagtt tgcagaggg cgcctttttc tgtgttagtt tgcagaggg cggagagcctt E7 orf start ->
841 aagagattggt cctgttcttt tgcagagttt tgcagagttt tgcagagttt tgcagagttt E7 orf start ->
901 gaaatccgga ATGgagggg cTAAGatgct ggtcagcccca gaaggggtga gcgagaggttt E7 orf start ->
961 tatagataat gaagagcact tttgagggc tttgagggc tttgagggc tttgagggc E7 orf start ->
1021 aagccccggg cttgcgagctt ctttctttct tttctttct tttctttct tttctttct E7 orf start ->
1081 gtagagtttt tgcagaggg cgcctttttc tgtgtgtagtt tgcagaggg cggagagcctt E7 orf start ->
1141 aagagattt gtagagtttt tgcagaggg cgcctttttc tgtgtgtagtt tgcagaggg cggagagcctt E7 orf start ->
1201 tgtgttggtt tgttttggtt tgttttggtt tgttttggtt tgttttggtt tgttttggtt E7 orf start ->
1261 attacgactt gccgtggctt gcctgtgttt gcctgtgttt gcctgtgttt gcctgtgttt E7 orf start ->
1321 gtagagtttt gtagagtttt gtagagtttt gtagagtttt gtagagtttt gtagagtttt E7 orf start ->
1381 gtagagtttt gtagagtttt gtagagtttt gtagagtttt gtagagtttt gtagagtttt E7 orf start ->
1441 gtagagtttt gtagagtttt gtagagtttt gtagagtttt gtagagtttt gtagagtttt E7 orf start ->
1501 gtagagtttt gtagagtttt gtagagtttt gtagagtttt gtagagtttt gtagagtttt E7 orf start ->
1561 gtagagtttt gtagagtttt gtagagtttt gtagagtttt gtagagtttt gtagagtttt E7 orf start ->
1621 tgtgtttttt tgtgtttttt tgtgtttttt tgtgtttttt tgtgtttttt tgtgtttttt E7 orf start ->
1681 tgtgtttttt tgtgtttttt tgtgtttttt tgtgtttttt tgtgtttttt tgtgtttttt E7 orf start ->
1741 atactgtgatc aaagggcagta cggacatcgc tgggtgtatct catgagagtc atcttggtatg
1801 gaggttatgct atagtgcttt ggcagataca tccacagaata cacaagatag aaggggagaag
1861 cgcacgttgc gtaaatcagac ctaatgctt ggcagcagca gatgctatgc aatttgacct
1921 ttctgaaatg atacaatggg cctatgatag cgatctcaca gatgaagctg acattgcata
1981 tctttatgct aaaatggcaa atagtgactc taatgcaaga gcttggttag cacataataa
2041 tcaggcaagg tacttaagag aatgtgctca aatggttaga cattacagac ggggagaat
2101 gagggatatg agtatgtctg agtggataca tcacagaata caacaagatag aaggggagaag
2161 aaacacaaat acaaacgttg aagttcactt cccgagatcc ctgggtttggt gtagatggca
2221 gtcggtgttc gcgaggagac ccggcgagaa cctatgatag cgatctcaca gatgaagctg acattgcata
2281 tctttatgct aaaatggcaa atagtgactc taatgcaaga gcttggttag cacataataa
2341 tcagaggctg tacttaagag aatgtgctca aatggttaga cattacagac ggggagaat
2401 gagggatatg agtatgtctg agtggataca tcacagaata caacaagatag aaggggagaag
2461 ccattggtct gaaatagtta agtttataag atttcaagaa ataaacttta taatatttct
2521 gccgccgac tgtggcaagt caatgtttgc tatgtcatta ttaaaagttt taaaaggcaa
2581 gacacaaatTA Gagctcagtg acccagaaga cgaggcagac aATGgaggca ctcaacgctc
2641 E2 orf start -> E2 cds ->
2701 gtttcaatgt actacaagag acttttaaag cattacaagag actatttaagtg gtagatggca
2761 aaacacaaat acaaacgttg aagttcactt cccgagatcc ctgggtttggt gtagatggca
2821 gttcagtaaa acaaatgctt ggtgcttattt gataagttt taaaatgtta cagatacaga gataagaagc
2881 taatatttta taaaatgtta cagatacaga gataagaagc
2941 tctttatgct aaaatggcaa atagtgactc taatgcaaga gcttggttag cacataataa
3001 cagatgctt cggagagaag cccttatttt taaaatgtta cagatacaga gataagaagc
3061 atctttaaag atatactgct tggaaagaga tttattttgT AGactcagAT Gatatgtggc
3121 gccttttagt actacaagag acttttaaag cattacaagag actatttaagtg gtagatggca
3181 acaaatgctt ggtgcttattt gataagttt taaaatgtta cagatacaga gataagaagc
3241 tcagaatgtg actatactgct tggaaagaga tttattttgT AGactcagAT Gatatgtggc
3301 gatgcttattt gataagttt taaaatgtta cagatacaga gataagaagc
3361 actatactgct tggaaagaga tttattttgT AGactcagAT Gatatgtggc
3421 ggacaatgactcagtg acccagaaga cgaggcagac aATGgaggca ctcaacgctc
3481 E4 orf start ->
3541 gtttcaatgt actacaagag acttttaaag cattacaagag actatttaagtg gtagatggca
3601 aaacacaaat acaaacgttg aagttcactt cccgagatcc ctgggtttggt gtagatggca
3661 cagatgctt cggagagaag cccttatttt taaaatgtta cagatacaga gataagaagc
3721 gccttttagt actacaagag acttttaaag cattacaagag actatttaagtg gtagatggca
3781 acaaatgctt ggtgcttattt gataagttt taaaatgtta cagatacaga gataagaagc
3841 taatatttta taaaatgtta cagatacaga gataagaagc
3901 E4 end
3961 gtaaattagg tttagtaaaa cattataqta ccaccctgtgct atgggtgttct gtgaatgctgca
4021 taatatttta taaaatgtta cagatacaga gataagaagc
4081 cagatgctt cggagagaag cccttatttt taaaatgtta cagatacaga gataagaagc
4141 taatatttta taaaatgtta cagatacaga gataagaagc
4201 L2 orf start ->
4261 atctttaaag atatactgct tggaaagaga tttattttgT AGactcagAT Gatatgtggc
4321 L2 cds ->
4381 tgtttttttg tgttttggtta atagtgcttg ccgcttgtct gcgtggcagt actggtcattg
4441 tatacttggg gataagtttt taaaatgtta cagatacaga gataagaagc
4501 taatatttta taaaatgtta cagatacaga gataagaagc
4561 cagatgctt cggagagaag cccttatttt taaaatgtta cagatacaga gataagaagc
4621 gccttttagt actacaagag acttttaaag cattacaagag actatttaagtg gtagatggca
4681 cagatgctt cggagagaag cccttatttt taaaatgtta cagatacaga gataagaagc
4741 gccttttagt actacaagag acttttaaag cattacaagag actatttaagtg gtagatggca
4801 cagatgctt cggagagaag cccttatttt taaaatgtta cagatacaga gataagaagc
4861 gccttttagt actacaagag acttttaaag cattacaagag actatttaagtg gtagatggca
4921 cagatgctt cggagagaag cccttatttt taaaatgtta cagatacaga gataagaagc
4981 taatatttta taaaatgtta cagatacaga gataagaagc
5041 taatatttta taaaatgtta cagatacaga gataagaagc

HPV49

SEP 94