A Survey of HPV Variants

The following pages summarize variation observed in several HPV and animal PV types. This introduction describes how to read the variant presentations. The "raw" alignment files of variants treated herein are available electronically on our Web site (http://hpv-web.lanl.gov).

Although perhaps obvious, it bears emphasizing that the variation presented here cannot pretend to correspond to all of the variation which could be observed in a novel isolate. For many types only limited regions have been sequenced, and for many regions only small numbers of isolates from a given type have been sequenced. The summary tables at the end of each type (described below) indicate roughly the extent to which a region has been studied, although in some cases duplicate sequences are represented only once, so the tables may underrepresent the amount of sequencing which has been done.

For each HPV type an abbreviated GenBank-like file is presented for the reference isolate. The FEATURES TABLE section of the file has been edited to show only the protein translations of coding sequences in which nucleotide variation has been observed (Figure 1).

```
CDS   4237..5658
      /note="L2 ORF from 4135 to 5658"
      /product="minor capsid protein"
      /gene="L2"
      /note="putative"
      /codon_start=1
      /translation="MRHKRSAKRTKRASATQLYKTCKQAGTCPPDIIPKVEGKTIADQ
      ILQYGSMGVFFGGLGIGTGSGTGGRTGYIPLGTRPPTATDTLAPVRPPLTVDPVGPSD
      PSIVSLVEETSFIDAGAPTSVPSIPPDVSGFSITTSTDTTPAILDINNTVTTVTTHNN
      PTFTDPSVLQPPTPAETGGHFTLSSSTISTHNYEEIPMDTFIVSTNPNTVTSSTPIPG
      SRPVARLGLYSRTTQQVKVVDPAFVTTPTKLITYDNPAYEGIDVDNTLYFSSNDNSIN
      IAPDPDFLDIVALHRPALTSRRTGIRYSRIGNKQTLRTRSGKSIGAKVHYYYDLSTID
      PAEEIELQTITPSTYTTTSHAASPTSINNGLYDIYADDFITDTSTTPVPSVPSTSLSG
      YIPANTTIPFGGAYNIPLVSGPDIPINITDQAPSLIPIVPGSPQYTIIADAGDFYLHP
      SYYMLRKRRKRLPYFFSDVSLAA"
```

Figure 1. Translation of the L2 gene from the Features Table of HPV-16R. Positions at which predicted amino acids are known to vary in other isolates of HPV-16 are shown by black boxes.

The nucleotide sequence that follows the Features Table has also been edited to remove large blocks of nucleotides for which only the sequence of the reference isolate has been determined. (Figure 2). The amino acid sequence of each orf is written out in single letter code below the DNA sequence, with amino acid codes lining up to the second position of the corresponding codon. Amino acids known to vary are boxed. Nucleotide positions are numbered from the beginning of the reference sequence. Amino acid positions, which are numbered from the start of each cds, are listed to the right of the sequence with the number corresponding to the last residue on a line.

Nucleotide positions at which nonsynonymous (amino acid replacing) variation has been observed are shown in white on black; the affected residues are boxed. (Strictly speaking, nucleotide positions are shown in white on black if the codon containing them undergoes a nonsynonymous change, regardless of whether the specific nucleotide change is nonsynonymous, which can be difficult to determine if a codon undergoes changes at more than one position.) Nucleotide positions at which synonymous changes are observed (strictly, positions at which a change is observed in codons for which there are only synonymous changes) are shown in white on gray. A triangle symbol "▲" indicates the location of an insertion in at least one of the variant sequences.

The tables following each GenBank file summarize variation at each nucleotide position at which variation has been observed. The tables are organized by gene or noncoding regions with each gene represented by two tables. The first table for a given gene, labelled "nuc," (Figure 3 A) shows nucleotide variation among the variants of the HPV type, with each column in the table presenting a position at which nucleotide changes are observed. The numbers above the body of the table, read top-to-bottom, indicate the number of the position (in the reference sequence) represented by each column. The first

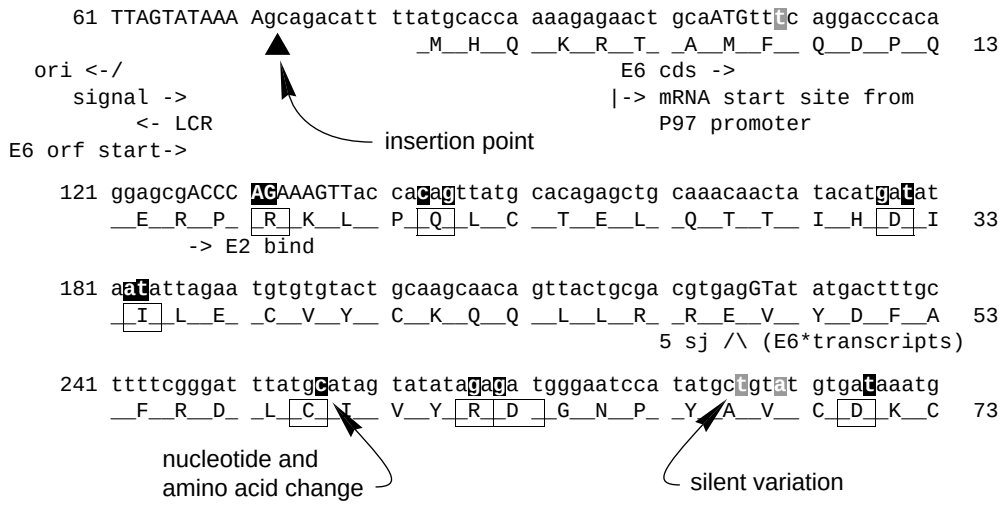


Figure 2. A portion of the E6 cds of HPV-16R. The annotated nucleotide sequence is shown with its protein translation below. Nucleotide variants that cause amino acid changes are shown in black, silent variants in gray. A "▲" symbol indicates the point where an insertion occurs in at least one variant.

row of the table represents the reference isolate, with white on black or gray to indicate nonsynonymous and synonymous changes, as described above. Dashes in the following rows indicate identity with the reference sequence, while a letter shows a change with respect to the reference and dots indicate a deletion. To the right of the table are indicated the first and last nucleotide positions over which a given variant was sequenced.

The second table in each gene-group, labelled "aa," (Figure 3 B) shows how these same nucleotide changes affect the predicted protein sequence. The numbers above the table, read vertically, indicate the amino acid residues to which the columns correspond. The reference isolate is again represented in the first row, with positions showing amino acid variation in white on black; positions shown by white on gray vary at the nucleotide but not the amino acid level. Dashes represent identity both on protein and nucleotide level to the reference isolate, letters represent changes in the predicted protein sequence relative to the reference, and asterisks indicate positions at which predicted protein sequences are identical, although there is synonymous change in the corresponding nucleotide positions. Each column in this table corresponds to a column from the preceding table (i.e. a position at which nucleotide variation is seen), in order to facilitate comparison of the two tables. This means certain amino acid residues are represented in two columns because the corresponding codon shows variability at two positions; these repeated columns are indicated by the symbol ⌒ above the relevant pairs of columns.

Occasionally, variant sequences have insertions relative to the reference sequence. These are indicated in two ways. If the insertion is short the letter "i" appears above the nucleotide position numbers (Figure 3 C). In cases of long insertions relative to the reference sequence an overbrace is used. When insertions occur the reference sequence shows a dot in the relevent nucleotide position. When the insertion is longer than one nucleotide the nucleotide position number is repeated as many times as necessary. Deletions also occur in some variant sequences relative to the reference sequence. These are shown with the same notation as just described for insertions, i.e., a "d" for short deletions and an overbrace for long deletions. In a few variant sequences frameshifts are observed which cause a relatively long string of mismatches between the variant and reference sequence. These have been noted in the GenBank file and also in the summary tables by the overbrace convention.

insertion

At the end of this section, beginning on page I-171, is a table that summarizes the sources of the variant files included in this analysis. In that table are listed the variant names, their accession numbers, and the primary literature reference.

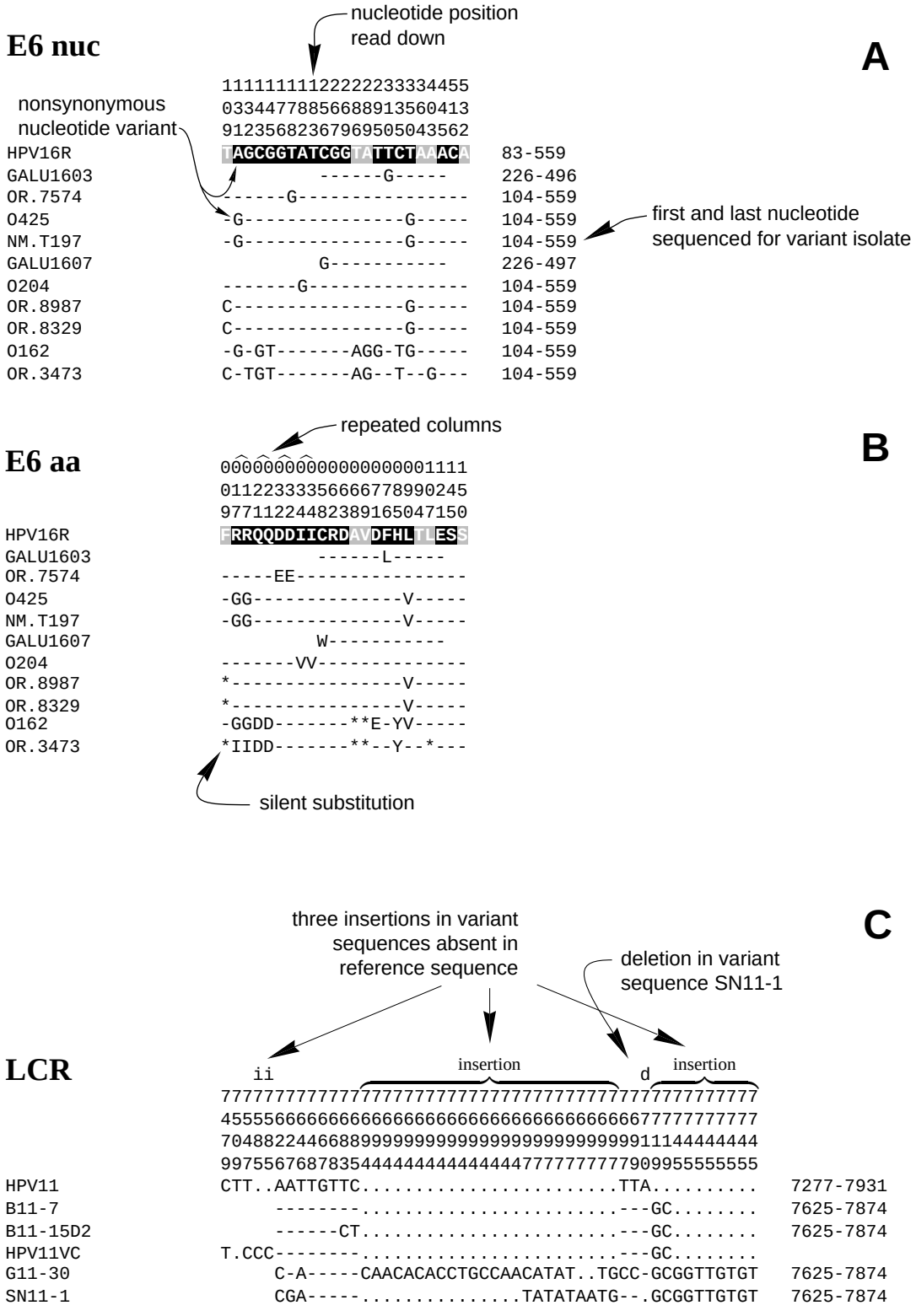


Figure 3. The nucleotide (A), and amino acid (B) tables summarizing variant sequence positions. Figure 3C illustrates insertions in variant sequences.

BPV-1R Variants

```
LOCUS       BPV1R        7946 bp ds-DNA   Circular  VRL       30-SEP-1988
DEFINITION  Bovine papillomavirus type 1 (BPV-1), complete genome.
ACCESSION   <not in GenBank>
KEYWORDS    complete genome; open reading frame.
SOURCE      Bovine papillomavirus type 1 DNA from cow, isolate 307.
  ORGANISM  Bovine papillomavirus type 1
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
COMMENT     Naturally occurring variants of BPV1R have been observed in the
            L1 region and the noncoding region 5' of the L2 ORF.
            Reference: Otten,N., von Tscharner,C., Lazary,S., Antczak,D.F.,
                and Gerber,H., Arch. Virol. 132, 121-31 (1993)
                Variant Accession Nos. L07797-L07800
            Reference: Reid,S.W., Smith,K.T., and Jarrett,W.F.,
                Vet. Rec. 135, 430-2 (1994)
                Variant Accession Nos. U23379
FEATURES             Location/Qualifiers
     CDS             5609..7096
                     /note="L1 ORF from bp 5597 to 7096"
                     /product="major capsid protein"
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /translation="MALWQQGQKLYLPPTPVSKVLCSETYVQRKSIFYHAETERLLTI
                     GHPYYPVSIGAKTVPKVSANQYRVFKIQLPDPNQFALPDRTVHNPSKERLVWPVIGVQ
                     VSRGQPLGGTVTGHPTFNALLDAENVNRKVTTQTTDDRKQTGLDAKQQQILLLGCTPA
                     EGEYWTTARPCVTDRLENGACPPLELKNKHIEDGDMMEIGFGAANFKEINASKSDLPL
                     DIQNEICLYPDYLKMAEDAAGNSMFFFARKEQVYVRHIWTRGGSEKEAPTTDFYLKNN
                     KGDATLKIPSVHFGSPSGSLVSTDNQIFNRPYWLFRAQGMNNGIAWNNLLFLTVGDNT
                     RGTNLTISVASDGTPLTEYDSSKFNVYHRHMEEYKLAFILELCSVEITAQTVSHLQGL
                     MPSVLENWEIGVQPPTSSILEDTYRYIESPATKCASNVIPAKEDPYAGFKFWNIDLKE
                     KLSLDLDQFPLCRRFLAQQGAGCSTVRKRRISQKTSSKPAKKKKK"
                                                  ▲
BASE COUNT     2270 a   1714 c   1887 g   2075 t
```

❄ ❄ ❄ ❄ Bases 1 to 4020 not shown. ❄ ❄ ❄ ❄

```
4021 acatcactgg ctattggctg tgtttttact gttgtgtgga tttgattgtt tttatatact
    -> noncoding region

4081 gtatgaagtt tttacatttg tgcttgtatt gctgtttgta agttttttac tagagtttgt

4141 attcccctg ctcagatttt atatggttTA AgctgcagcA ATAAAaATGa gtgcacGaaa
                                                    _M__ S__A__R__K      5
                           L2 orf start ->       L2 cds ->           L2 orf ->
                                         signal ->    early poly-A |
```

❄ ❄ ❄ ❄ Bases 4201 to 5521 not shown. ❄ ❄ ❄ ❄

```
5581 acggaaacat gccTAAtttt ttttgcAGAT Ggcgttgtgg caacaaggcc agaagctgta
     __R__K__H_ _A__$_             _M __A__L__W_ _Q__Q__G__ Q__K__L__Y      469/11
         L1 orf start ->          /\ 3 sj
                           L1 cds ->
                      <- L2 end

5641 tctccctcca acccctgtaa gcaaggtgct ttgcagtgaa acctatgtgc aaagaaaaag
     __L__P__P_ _T__P__V__ S__K__V__L __C__S__E_ _T__Y__V__ Q__R__K__S      31

5701 catttttat catgcagaaa cggagcgcct gctaactata ggacatccat attacccagt
     __I__F__Y_ _H__A__E__ T__E__R__L __L__T__I_ _G__H__P__ Y__Y__P__V      51

5761 gtctatcggg gccaaaactg ttcctaaggt ctctgcaaat cagtataggg tatttaaaat
     __S__I__G_ _A__K__T__ V__P__K__V __S__A__N_ _Q__Y__R__ V__F__K__I      71
```

```
5821 acaactacct gatcccaatc aatttgcact acctgacagg actgttcaca acccaagtaa
     __Q__L__P_ _D__P__N__ Q__F__A__L __P__D__R_ _T__V__H__ N__P__S__K      91

5881 agagcggctg gtgtgGgcag tcataggtgt gcaggtgtcc agagggcagc ctcttggagg
     __E__R__L_ _V__W__A__ V__I__G__V __Q__V__S_ _R__G__Q__ P__L__G__G      111
                    ^ 'c' replaced by 'g'

5941 tactgtaact gggcacccca cttttaatgc tttgcttgat gcagaaaatg tgaatagaaa
     __T__V__T_ _G__H__P__ T__F__N__A __L__L__D_ _A__E__N__ V__N__R__K      131

6001 agtcaccacc caaacaacag atgacaggaa acaaacaggc ctagatgcta agcaacaaca
     __V__T__T_ _Q__T__T__ D__D__R__K __Q__T__G_ _L__D__A__ K__Q__Q__Q      151

6061 gattctgttg ctaggctgta cccctgctga aggggaatat tggacaacag cccgtccatg
     __I__L__L_ _L__G__C__ T__P__A__E __G__E__Y_ _W__T__T__ A__R__P__C      171

6121 tgttactgat cgtctagaaa atggcgcctg ccctcctctt gaattaaaaa acaagcacat
     __V__T__D_ _R_|L_|E_| N__G__A__C __P__P__L_ _E__L__K__ N__K__H__I      191

6181 agaagatggg gatatgatgg aaattgggtt tggtgcagcc aacttcaaag aaattaatgc
     __E__D__G_ _D__M__M__ E__I__G__F __G__A__A_ _N__F__K__ E__I__N__A      211

6241 aagtaaatca gatctacctc ttgacattca aaatgagatc tgcttgtacc cagactacct
     __S__K__S_ _D__L__P__ L__D__I__Q __N__E__I_ _C__L__Y__ P__D__Y__L      231

6301 caaaatggct gaggacgctg ctggtaatag catgttcttt tttgcaagga aagaacaggt
     __K__M__A_ _E__D__A__ A__G__N__S __M__F__F_ _F__A__R__ K__E__Q__V      251

6361 gtatgttaga cacatctgga ccagaggggg ctcggagaaa gaagccccta ccacagattt
     __Y__V__R_ _H__I__W__ T__R__G__G __S__E__K_ _E__A__P__ T__T__D__F      271

6421 ttatttaaag aataataaag gggatgccac ccttaaaata cccagtgtgc attttggtag
     __Y__L__K_ _N__N__K__ G__D__A__T __L__K__I_ _P__S__V__ H__F__G__S      291

6481 tcccagtggc tcactagtct caactgataa tcaaattttt aatcggccct actggctatt
     __P__S__G_ _S__L__V__ S__T__D__N __Q__I__F_ _N__R__P__ Y__W__L__F      311

6541 ccgtgcccag ggcatgaaca atggaattgc atggaataat ttattgtttt taacagtggg
     __R__A__Q_ _G__M__N__ N__G__I__A __W__N__N_ _L__L__F__ L__T__V__G      331

6601 ggacaataca cgtggtacta atcttaccat aagtgtagcc tcagatggaa ccccactaac
     __D__N__T_ _R__G__T__ N__L__T__I __S__V__A_ _S__D__G__ T__P__L__T      351

6661 agagtatgat agctcaaaat tcaatgtata ccatagacat atggaagaat ataagctagc
     __E__Y__D_ _S__S__K__ F__N__V__Y __H__R__H_ _M__E__E__ Y__K__L__A      371

6721 ctttatatta gagctatgct ctgtggaaat cacagctcaa actgtgtcac atctgcaagg
     __F__I__L_ _E__L__C__ S__V__E__I __T__A__Q_ _T__V__S__ H__L__Q__G      391

6781 acttatgccc tctgtgcttg aaaattggga ataggtgtg cagcctccta cctcatcgat
     __L__M__P_ _S__V__L__ E__N__W__E __I__G__V_ _Q__P__P__ T__S__S__I      411

6841 attagaggac acctatcgct atatagagtc tcctgcaact aaatgtgcaa gcaatgtaat
     __L__E__D_ _T__Y__R__ Y__I__E__S __P__A__T_ _K__C__A__ S__N__V__I      431

6901 tcctgcaaaa gaagacccct atgcagggtt taagttttgg aacatagatc ttaaagaaaa
     __P__A__K_ _E__D__P__ Y__A__G__F __K__F__W_ _N__I__D__ L__K__E__K      451

6961 gctttctttg gacttagatc aatttccctt gggaagaaga tttttagcac agcaaggggc
     __L__S__L_ _D__L__D__ Q__F__P__L _|G|_R__R_ _F__L__A__ Q__Q__G__A      471

7021 aggatgttca actgtgagaa acgaagaat tagccaaaaa acttccagta agcctgcaaa
     __G__C__S_ _T_|V_| R_|_K_| R__R__I __S__Q_|K_| _T_|S_| S__ K_|P_| A__K      491
                ▲
```

❄ ❄ ❄ ❄ Bases 7081 to 7946 not shown. ❄ ❄ ❄ ❄

noncoding

```
                4444
                0001
                3694
                6748
BPV1R           GTTC   4014-4186
PPBVARA         AGAT   4014-4186
PPBVARC         AGAT   4014-4186
```

L1 nuc

```
                     i  dddddd
                6666677777777777
                1111900000000000
                3334933334666677
                5795311262015635
BPV1R           TGA▸G.ACGAAACCCT   5609-7096
BPU23379            .ACG......TC   6964-7083
PPBVARB         CATG               6062-6274
PPBVARD         CATG               6062-6274
```

L1 aa

```
                1îî1144îî444444îî44
                7777677777888888
                6779255568456699
BPV1R           LEE▸G.TTVKKTSSPP
BPU23379            XXRRXXXXXSS
PPBVARB         PNN*
PPBVARD         PNN*
```

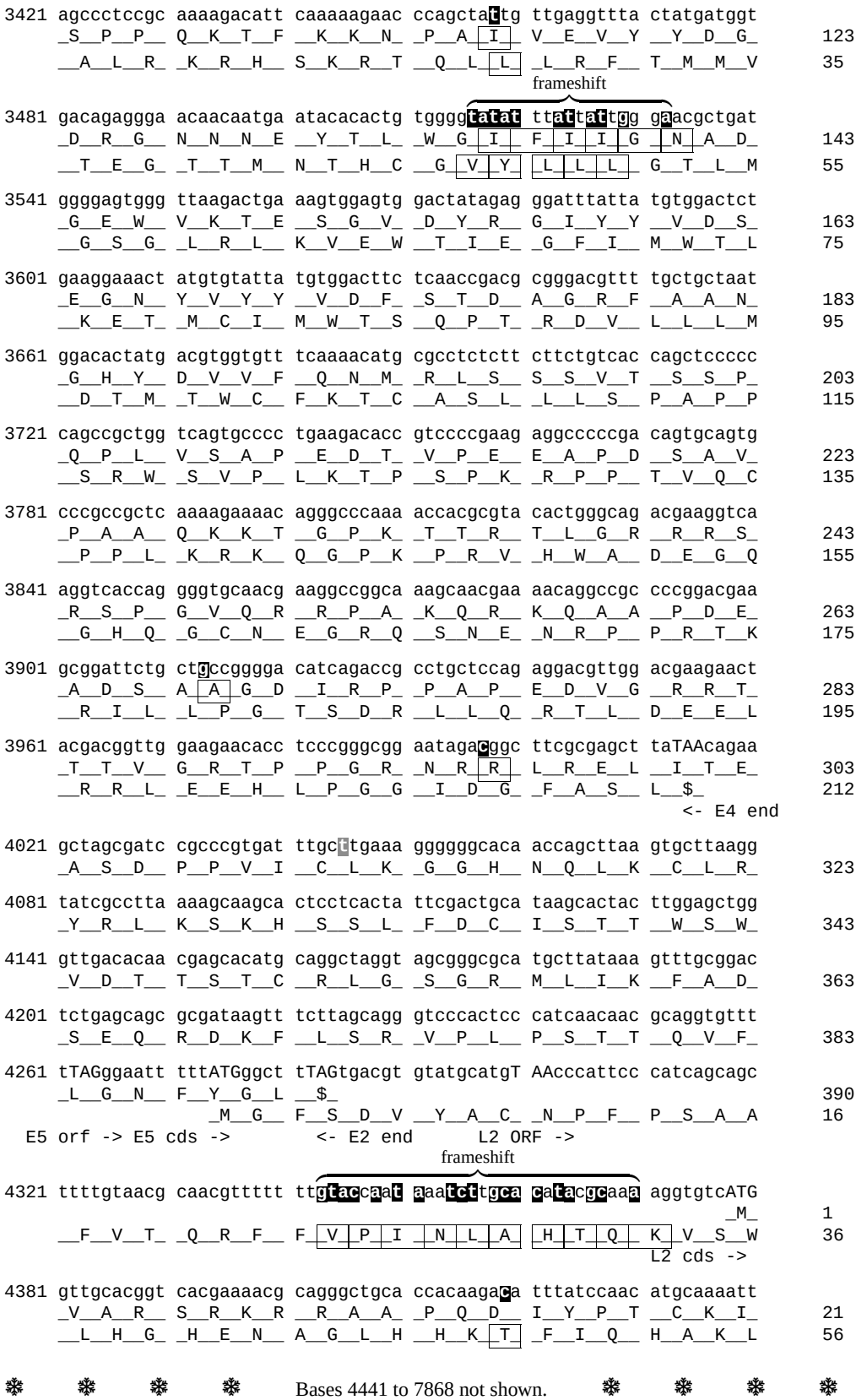
```
LOCUS       CRPV         7868 bp ds-DNA   Circular  VRL       31-AUG-1987
DEFINITION  Cottontail rabbit (Shope) papillomavirus (CRPV), complete genome.
ACCESSION   K02708
KEYWORDS    complete genome.
SOURCE      Cottontail Rabbit Papillomavirus DNA recovered from a
            papilloma of wild cottontail rabbit (provided by G.Orth),
            clone pIG30.
  ORGANISM  Cottontail rabbit papillomavirus
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
COMMENT     Naturally occurring variants of CRPV have been observed in the region
            E7, E1, E2, E4, E5, and L2.
            Reference: Wu,X., Xiao,W., and Brandsma,J.L.,  J. Virol. 68, 6097-102
                (1994). Variant Accession Nos. U09467, U09493-U09497
FEATURES             Location/Qualifiers
     CDS             1075..1359
                     /note="E7 ORF from bp 979 to 1359"
                     /product="transforming protein"
                     /gene="E7"
                     /note="putative"
                     /codon_start=1
                     /translation="MIGRTPKLSELVLGETAEALSLHCDEALENLSDDDEEDHQDRQV
                     FIERPYAVSVPCKRCRQTISFVCVCAPEAIRTLNRLLSASLSLVCPECCN"
     CDS             1362..3170
                     /note="E1 ORf from bp 1356 to 3170"
                     /product="replication protein"
                     /gene="E1"
                     /note="putative"
                     /codon_start=1
                     /translation="MAEGTDPLDDCGGFLDTEADCLDCDNLEEDLTELFDADTVSSLL
                     DDTDQVQGNSLEPFQHHEATETLKSIEHLKRKYVDSPDKSLGIDNSVNALSPRLQAFS
                     LSGQKKAVKKRLFGTDGDEAASGAESLQVESGFGSQQSVSDTPVTDILNANTARVKHL
                     LLFRQAHSVSFSELTRTFQSDKTMSWDWVGGLADIHVSVLESLQTSLRSHCVYVQYDL
                     NFAETNASSLLLLLRFKAQKCRDGVKALLSQLLGVQDLKVLLEPPKTRSVAVALFWYK
                     RAMVSGVFSYGPMPEWITQQTNVNHQMLQEKPFQLSVMVQWAYDNHLQDESSIAYKYA
                     MLAETDENARAFLASNSQAKYVRDCCNMVRLYLRAEMRQMTMSAWINYRLDGMNDDGD
                     WKVVVHFLRHQRVEFIPFMVKLKAFLRGTPKKNCMVFYGPPNSGKSYFCMSLIRLLAG
                     RVLSFANSRSHFWLQPLADAKLALVDDATSACWDFIDTYLRNALDGNPISVDLKHKAP
                     IEIKCPPLLITTNVDVKSDDRWRYLFSRICVFNFLQELPIRNGTPVYELNDANWKSFF
                     KRFWSTLELSDPEDEGDDGGSQPALRLHTGGTSQSL"
     CDS             3112..4284
                     /note="E2 ORF from bp 3088 to 4284"
                     /product="regulatory protein"
                     /gene="E2"
                     /note="putative"
                     /codon_start=1
                     /translation="MEALSQRLDSIQEELLSLYEKESTSLESQLQHWNLLRKEQVLLH
                     FCKKHGIRQLGYTPVPSLLTSQECAKQAIEMVLYIESLLRSPYSDEPWTLQDTSRERF
                     ESPPQKTFKKNPATVEVYYDGDRGNNNEYTLWGTIFIIGNADGEWVKTESGVDYRGIYY
                     VDSEGNYVYYVDFSTDAGRFAANGHYDVVFQNMRLSSSVTSSPQPLVSAPEDTVPEEA
                     PDSAVPAAQKKTGPKTTRTLGRRRSRSPGVQRRPAKQRKQAAPDEADSAAGDIRPPAP
                     EDVGRRTTTVGRTPPGRNRRLRELITEASDPPVICLKGGHNQLKCLRYRLKSKHSSLF
                     DCISTTWSWVDTTSTCRLGSGRMLIKFADSEQRDKFLSRVPLPSTTQVFLGNFYGL"
     CDS             3377..4015
                     /note="E4 ORF from bp 3332 to 4015"
                     /gene="E4"
                     /note="putative"
                     /codon_start=1
                     /translation="MSHGHCRIPVEKGSKALRKRHSKRTQLLLRFTMMVTEGTTMNTH
                     CGVYLLLGTLMGSGLRLKVEWTIEGFIMWTLKETMCIMWTSQPTRDVLLLMDTMTWCF
                     KTCASLLLSPAPPSRWSVPLKTPSPKRPPTVQCPPLKRKQGPKPRVHWADEGQGHQGC
                     NEGRQSNENRPPRTKRILLPGTSDRLLQRTLDEELRRLEEHLPGGIDGFASL"
     CDS             4274..4579
                     /note="E5 ORF from bp 4265 to 4579"
```

```
                     /gene="E5"
                     /note="putative"
                     /codon_start=1
                     /translation="MGFSDVYACNPFPSAAFVTQRFFVPINLAHTQKVSWLHGHENAG
                     LHHKTFIQHAKLLAIAQLTYRINLKTKQLQIKFCSMAALVFSLEDLESVVPEVLGVD"

BASE COUNT     2252 a   1725 c   1929 g   1962 t
```

❄ ❄ ❄ ❄ Bases 1 to 1020 not shown. ❄ ❄ ❄ ❄

```
1021 ttctttatat actactgttt ttccttctgt actggcttta tcgaattctg caagATGata
                                                          _M__I_      2
   -> noncoding region                                  E7 cds ->

1081 ggcagaactc ctaagcttag tgagctggtt ttaggtgaaa ctgctgaagc gcttagtctg
     _G__R__T__ P__K__L__S __E__L__V_ _L__G__E__ T__A__E__A __L__S__L_      22

1141 cattgcgacg aagcattaga gaatttaagt gatgatgatg aggaggatca tcaagataga
     _H__C__D__ E__A__L__E __N__L__S_ _D__D__D__ E__E__D__H __Q__D__R_      42

1201 caggtgttca tagaaaggcc ctatgcagtg tccgtgccat gtaagcgctg taggcaaact
     _Q__V__F__ I__E__R__P __Y__A__V_ _S__V__P__ C__K__R__C __R__Q__T_      62

1261 atcagcttcg tctgcgtctg tgctccagaa gccataagaa ccttgaatcg actgctatcc
     _I__S__F__ V__C__V__C __A__P__E_ _A__I__R__ T__L__N__R __L__L__S_      82

1321 gcatcgcttt ccctggtgtg cccggagtgt tgTAAcTGAa aATGgctgaa ggtacagacc
     _A__S__L__ S__L__V__C __P__E__C_ _C__N__$_                          94
                                      E7 end <-   _M__A__E_ _G__T__D__      6
                                 E1 ORF start ->     -> E1 cds

1381 ctttagatga ctgtgggggg ttcttagaca cggaagcgga ctgtttagac tgtgacaacc
     P__L__D__D __C__G__G_ _F__L__D__ T__E__A__D __C__L__D_ _C__D__N__      26

1441 ttgaggagga cctgacagag ctgtttgatg ctgacactgt aagcagttta ctagatgata
     L__E__E__D __L__T__E_ _L__F__D__ A__D__T__V __S__S__L_ _L__D__D__      46

1501 cagatcaggt gcagggaaat tccctggaac cttttcagca tcatgaggcg actgagacct
     T__D__Q__V __Q__G__N_ _S__L__E__ P__F__Q__H __H__E__A_ _T__E__T__      66

1561 tgaaaagcat agagcatctc aagagaaagt atgtcgatag tcctgataag agcctgggta
     L__K__S__I __E__H__L_ _K__R__K__ Y__V__D__S __P__D__K_ _S__L__G__      86

1621 tcgacaactc cgtcaatgcc ttgagtccaa gattacaagc tttctcactg tcaggacaaa
     I__D__N__S __V__N__A_ _L__S__P__ R__L__Q__A __F__S__L_ _S__G__Q__     106

1681 aaaaggctgt taaaaagaga cttttcggta ctgacggaga tgaagctgct tctggtgctg
     K__K__A__V __K__K__R_ _L__F__G__ T__D__G__D __E__A__A_ _S__G__A__     126

1741 agtcgttaca ggtagaatcg ggatttgggt ctcaacaaag cgtatcagat acacctgtga
     E__S__L__Q __V__E__S_ _G__F__G__ S__Q__Q__S __V__S__D_ _T__P__V__     146

1801 ctgacatttt aaatgcaaat acagcaagag tcaaacattt gttgttattt aggcaagctc
     T__D__I__L __N__A__N_ _T__A__R__ V__K__H__L __L__L__F_ _R__Q__A__     166

1861 acagtgttag cttttcggag ctcaccagaa catttcaaag tgacaagact atgagttggg
     H__S__V__S __F__S__E_ _L__T__R__ T__F__Q__S __D__K__T_ _M__S__W__     186

1921 attgggtagg tgggctggcg gacattcatg taagcgtgtt ggagagcttg cagacatctc
     D__W__V__G __G__L__A_ _D__I__H__ V__S__V__L __E__S__L_ _Q__T__S__     206

1981 tgagaagtca ttgcgtatat gttcagtatg atctcaattt tgcagagaca aatgcttcat
     L__R__S__H __C__V__Y_ _V__Q__Y__ D__L__N__F __A__E__T_ _N__A__S__     226

2041 ctctgctgct gctcctgaga tttaaagcac aaaaatgtag ggacggggtt aaagcgctgc
     S__L__L__L __L__L__R_ _F__K__A__ Q__K__C__R __D__G__V_ _K__A__L__     246

2101 tatcccaatt gttgggagtt caagatctaa aagttttatt agaacctcca aaaacaagga
     L__S__Q__L __L__G__V_ _Q__D__L__ K__V__L__L __E__P__P_ _K__T__R__     266
```

```
2161 gtgtcgctgt tgcattgttc tggtacaaaa gggcgatggt ttctggggtt tttagctacg
     S__V__A__V __A__L__F_ _W__Y__K__ R__A__M__V __S__G__V_ _F__S__Y__      286

2221 gtccaatgcc tgaatggata acgcagcaga caaatgttaa ccatcaaatg ttgcaggaaa
     G__P__M__P __E__W__I_ _T__Q__Q__ T__N__V__N __H__Q__M_ _L__Q__E__      306

2281 agccgtttca gttgtctgtc atggtccagt gggcatatga taaccacctt caggatgaaa
     K__P__F__Q __L__S__V_ _M__V__Q__ W__A__Y__D __N__H__L_ _Q__D__E__      326

2341 gtagtattgc atacaagtat gcaatgctcg ctgaaactga tgagaatgca agagcgtttc
     S__S__I__A __Y__K__Y_ _A__M__L__ A__E__T__D __E__N__A_ _R__A__F__      346

2401 tagcttctaa ttctcaggcg aagtatgtta gggactgttg caacatggtc agactctatt
     L__A__S__N __S__Q__A_ _K__Y__V__ R__D__C__C __N__M__V_ _R__L__Y__      366

2461 taagagcaga aatgagacag atgaccatgt ctgcatggat aaactacaga ttggatggga
     L__R__A__E __M__R__Q_ _M__T__M__ S__A__W__I __N__Y__R_ _L__D__G__      386

2521 tgaacgatga tggggattgg aaggtggtcg tgcattttct gcggcaccaa cgagtggagt
     M__N__D__D __G__D__W_ _K__V__V__ V__H__F__L __R__H__Q_ _R__V__E__      406

2581 tcataccttt catggtgaag ctgaaggcct tcctaagagg aacaccaaaa aaaaattgca
     F__I__P__F __M__V__K_ _L__K__A__ F__L__R__G __T__P__K_ _K__N__C__      426

2641 tggtgtttta tgggccacca aatagtggga agtcatattt ttgcatgagc ctcataagat
     M__V__F__Y __G__P__P_ _N__S__G__ K__S__Y__F __C__M__S_ _L__I__R__      446

2701 tacttgcagg acgggtcttg tcgtttgcaa acagcagaag ccatttttgg ctgcaaccat
     L__L__A__G __R__V__L_ _S__F__A__ N__S__R__S __H__F__W_ _L__Q__P__      466

2761 tagcagacgc caagctagcg ctcgtggatg atgctacatc cgcgtgctgg gatttcattg
     L__A__D__A __K__L__A_ _L__V__D__ D__A__T__S __A__C__W_ _D__F__I__      486

2821 atacatacct cagaaatgcc cttgatggca atcccatatc ggtggacctg aagcacaagg
     D__T__Y__L __R__N__A_ _L__D__G__ N__P__I__S __V__D__L_ _K__H__K__      506

2881 caccaataga gattaagtgc cctcccctcc tgataaccac aaatgtggac gtcaaatcag
     A__P__I__E __I__K__C_ _P__P__L__ L__I__T__T __N__V__D_ _V__K__S__      526

2941 atgatagatg gagatactta tttagtagaa tttgtgtgtt taacttttg caagaattgc
     D__D__R__W __R__Y__L_ _F__S__R__ I__C__V__F __N__F__L_ _Q__E__L__      546

3001 ccattagaaa tgggacacct gtgtatgaat taaatgatgc aaactggaaa tctttttta
     P__I__R__N __G__T__P_ _V__Y__E__ L__N__D__A __N__W__K_ _S__F__F__      566

3061 aaaggttctg gtccacctta gaacTAAgcg acccggaaga cgagggtgac gATGgaggct
                                                          _M__E__A_      3
     K__R__F__W __S__T__L_ _E__L__S__ D__P__E__D __E__G__D_ _D__G__G__      586
                    E2 ORF start ->                E2 cds ->

3121 ctcagccagc gcttagactc catacaggag gaacttctca gtctctaTGA gaaggagagc
     _L__S__Q__ R__L__D__S __I__Q__E_ _E__L__L__ S__L__Y__E __K__E__S_      23
     S__Q__P__A __L__R__L_ _H__T__G__ G__T__S__Q __S__L__$_                602
                                                       <- E1 end

3181 acgagtttgg agtcccagct acagcactgg aacttactaa gaaaagaaca ggtcctttta
     _T__S__L__ E__S__Q__L __Q__H__W_ _N__L__L__ R__K__E__Q __V__L__L_      43

3241 catttctgta aaaaacacgg gatcaggcaa ctgggctaca cgcctgtccc gtctcttctt
     _H__F__C__ K__K__H__G __I__R__Q_ _L__G__Y__ T__P__V__P __S__L__L_      63

3301 acctcacagg aatgtgcaaa gcaagccaTA Gaaatggtgc tgtacattga aagcctactc
     _T__S__Q__ E__C__A__K __Q__A__I_ _E__M__V__ L__Y__I__E __S__L__L_      83
                              E4 orf ->

3361 aggtccccgt attcagATGa gccatggaca ttgcaggata ccagtagaga aaggttcgaa
     _R__S__P__ Y__S__D__E __P__W__T_ _L__Q__D__ T__S__R__E __R__F__E_      103
                      _M__ S__H__G__H __C__R__I_ _P__V__E__ K__G__S__K      15
               E4 cds ->
```


frameshift
frameshift
E5 orf -> E5 cds ->
<- E2 end
L2 ORF ->
L2 cds ->
<- E4 end
Bases 4441 to 7868 not shown.

noncoding

```
                1
                0
                7
                4
CRPV            G   976-1074
CRPVvar.con     C   1020-1074
```

E7 nuc

```
                11111
                12222
                10018
                17813
CRPV            TTTTC   1075-1359
CRPVvar.con     CCAGA   1075-1359
```

E7 aa

```
                1 4 4 4 7
                35560
CRPV            LFFIA
CRPVvar.con     *HHRD
```

E1 nuc

```
                111
                559
                335
                151
CRPV            CTT   1362-3170
CRPVvar.con     TCC   1362-1639,1738-1995
```

E1 aa

```
                001
                559
                787
CRPV            PFV
CRPVvar.con     L*A
```

E2 nuc

```
                   frameshift
                3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4
                4 5 5 5 5 5 5 5 5 5 5 9 9 0
                5 1 1 1 1 2 2 2 2 2 3 1 9 4
                8 6 7 8 9 0 3 4 6 7 9 2 3 7 5
CRPV            TTATATATATGAGCT   3112-4284
CRPVvar.con     GGTATATATATGATC   3374-3674,3769-4209,4252-4284
```

E2 aa

```
                1 1 1 1 1 1 1 1 1 1 1 1 2 2 3
                1 3 3 3 3 3 3 3 3 3 4 4 6 9 1
                6 5 6 6 6 7 8 8 9 9 0 1 8 6 2
CRPV            IGIIIFIIIIGNARL
CRPVvar.con     S*YYYIYYYYWDTW*
```

E4 nuc

```
                33333333333333
                45555555555599
                51111222222319
                86789034679237
CRPV            TTATATATATGAGC   3377-4015
CRPVvar.con     GGTATATATATGAT   3377-3674,3769-4015
```

E4 aa

```
                00000000000012
                24444445555570
                87788890011297
CRPV            LVVYYYLLLLLGLD
CRPVvar.con     VGGIIIFIIII***
```

E5 nuc

```
                      frameshift
                44444444444444444444
                33333333333333333334
                44444555555566666671
                34568014568901346709
CRPV            GTACATATCTGCACTAGCAC   4274-4579
CRPVvar.con     TGTACATATCTGCTCGAACT   4274-4571
```

E5 aa

```
                22222222222223333334
                44455667889990011239
CRPV            VVVPPIINLLAAAHHTTQKT
CRPVvar.con     CCCTTNNKSSCCCYYAAKQI
```

L2 nuc

```
                4
                4
                1
                9
CRPV            C   4378-5856
CRPVvar.con     T   4378-4571
```

L2 aa

```
                1
                4
CRPV            D
CRPVvar.con     *
```

```
LOCUS       HPV1a        7815 bp    ds-DNA   Circular  VRL       06-JUL-1989
DEFINITION  Human papillomavirus 1a (HPV-1a), complete genome.
ACCESSION   V01116 X03321
KEYWORDS    circular; genome; origin of replication.
SOURCE      Human papillomavirus type 1a DNA.
  ORGANISM  Human papillomavirus type 1a
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
COMMENT     Naturally occurring variation of HPV1a has been observed throughout
            the genome in the complete genomic sequence of a variant isolate;
            however, the variant isolate is highly similar to the reference
            clone.
            Reference: Meissner,J., Virus Genes 9, 189-91 (1995).
            Variant Accession No: U06714
FEATURES             Location/Qualifiers
     CDS             812..2650
                     /note="putative"
                     /note="ORF E1 from bp 773 to 2650"
                     /product="replication protein"
                     /gene="E1"
                     /note="putative"
                     /codon_start=1
                     /translation="MADNKGTENDWFLVEATDCEETLEETSLGDLDNVSCVSDLSDLL
                     DEAPQSQGNSLELFHKQESLESEQELNALKRKLLYSPQARSADETDIASISPRLETIS
                     ITKQDKKRYRRQLFSQDDSGLELSLLQDETENIDESTQVDQQQKEHTGEVGAAGVNIL
                     KASNIRAALLSRFKDTAGVSFTDLTRSYKSNKTCCGDWVLAVWGVRENLIDSVKELLQ
                     THCVYIQLEHAVTEKNRFLFLLVRFKAQKSRETVIKLITTILPVDASYILSEPPKSRS
                     VAAALFWYKRSMSSTVFTWGTTLEWIAQQTLINHQLDSESPFELCKMVQWAYDNGHTE
                     ECKIAYYYAVLADEDENARAFLSSNSQAKYVKDCAQMVRHYLRAEMAQMSMSEWIFRK
                     LDNVEGSGNWKEIVRFLRFQEVEFISFMIAFKDLLCGKPKKNCLLIFGPPNTGKSMFC
                     TSLLKLLGGKVISYCNSKSQFWLQPLADAKIGLLDDATKPCWDYMDTYMRNALDGNTI
                     CIDLKHRAPQQIKCPPLLITSNIDVKSDTCWMYLHSRISAFKFAHEFPFKDNGDPGFS
                     LTDENWKSFFERFWQQLELSDQEDEGNDGKPQQSLRLTARAANEPI"
     CDS             2592..3797
                     /note="putative"
                     /note="ORF E2 from bp 2568 to 3797"
                     /product="regulatory protein"
                     /gene="E2"
                     /note="putative"
                     /codon_start=1
                     /translation="MENLSSRLDLLQEQLMNLYEQDSKLIEDQIKQWNLIRQEQVLFH
                     FARKNGVMRIGLQAVPSLASSQEKAKTAIEMVLHLESLKDSPYGTEDWSLQDTSRELF
                     LAPPAGTFKKSGSTLEVTYDNNPDNQTRHTIWNHVYYQNGDDVWRKVSSGVDAVGVYY
                     LEHDGYKNYYVLFAEEASKYSTTGQYAVNYRGKRFTNVMSSTSSPRAAGAPAVHSDYP
                     TLSESDTAQQSTSIDYTELPGQGETSQVRQRQQKTPVRRRPYGRRRSRSPRGGGRREG
                     ESTPSRTPGSVPSARDVGSIHTTPQKGHSSRLRRLLQEAWDPPVVCVKGGANQLKCLR
                     YRLKASTQVDFDSISTTWHWTDRKNTERIGSARMLVKFIDEAQREKFLERVALPRSVS
                     VFLGQFNGS"
     CDS             3877..4152
                     /note="putative"
                     /note="ORF E5 from bp 3877 to 4152"
                     /gene="E5"
                     /note="putative"
                     /codon_start=1
                     /translation="SLQHCNKCIAYVENALPPKIYTPHAKYQTPAHLTFKIKLSIQQLL
                                              ▲
                     IKYCNMAVWEFFWEVWALEQPEALEEELVILPSVRVVGLELLLVQLQ"
     CDS             3897..5420
                     /note="putative"
                     /note="ORF L2 from bp 3894 to 5420"
                     /product="minor capsid protein"
                     /gene="L2"
                     /note="putative"
                     /codon_start=1
```

```
                     /translation="MYRLRRKRAAPKDIYPSCKISNTCPPDIQNKIEHTTIADKILQY
                     GSLGVFLGGLGIGTARGSGGRIGYTPLGEGGGVRVATRPTPVRPTIPVETVGPSEIFP
                     IDVVDPTGPAVIPLQDLGRDFPIPTVQVIAEIHPISDIPNIVASSTNEGESAILDVLR
                     GNATIRTVSRTQYNNPSFTVASTSNISAGEASTSDIVFVSNGSGDRVVGEDIPLVELN
                     LGLETDTSSVVQETAFSSSTPIAERPSFRPSRFYNRRLYEQVQVQDPRFVEQPQSMVT
                     FDNPAFEPELDEVSIIFQRDLDALAQTPVPEFRDVVYLSKPTFSREPGGRLRVSRLGK
                     SSTIRTRLGTAIGARTHFFYDLSSIAPEDSIELLPLGEHSQTTVISSNLGDTAFIQGE
                     TAEDDLEVISLETPQLYSEEELLDTNESVGENLQLTITNSEGEVSILDLTQSRVRPPF
                     GTEDTSLHVYYPNSSKGTPIINPEESFTPLVIIALNNSTGDFELHPSLRKRRKRAYV"
     CDS             5413..6939
                     /note="putative"
                     /note="ORF L1 from bp 5347 to 6939"
                     /product="major capsid protein"
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /translation="MYNVFQMAVWLPAQNKFYLPPQPITRILSTDEYVTRTNLFYHAT
                     SERLLLVGHPLFEISSNQTVTIPKVSPNAFRVFRVRFADPNRFAFGDKAIFNPETERL
                     VWGLRGIEIGRGQPLGIGITGHPLLNKLDDAENPTNYINTHANGDSRQNTAFDAKQTQ
                     MFLVGCTPASGEHWTSSRCPGEQVKLGDCPRVQMIESVIEDGDMMDIGFGAMDFAALQ
                     QDKSDVPLDVVQATCKYPDYIRMNHEAYGNSMFFFARREQMYTRHFFTRGGSVGDKEA
                     VPQSLYLTADAEPRTTLATTNYVGTPSGSMVSSDVQLFNRSYWLQRCQGQNNGICWRN
                     QLFITVGDNTRGTSLSISMKNNASTTYSNANFNDFLRHTEEFDLSFIVQLCKVKLTPE
                     NLAYIHTMDPNILEDWQLSVSQPPTNPLEDQYRFLGSSLAAKCPEQAPPEPQTDPYSQ
                     YKFWEVDLTERMSEQLDQFPLGRKFLYQSGMTQRTATSSTTKRKTVRVSTSAKRRRKA"
BASE COUNT     2396 a   1482 c   1666 g   2271 t

        1 gttaactacc atcattcatt attctagtta caacaagaac cTAGgagtta tatgccagaa
                                            E8 orf start ->

       61 gtaagccTAT AAAAtacaca ggTAAgactc tgcacaggac cagATGgcga caccaatccg
                                                      _M__A__ T__P__I__R      6
          signal ->     E6 orf start ->           E6 cds ->

      121 gaccgtcaga cagctttccg aaagcctctg tatcccatat attgATGttt tattgccttg
          __T__V__R_ _Q__L__S__ E__S__L__C __I__P__Y_ _I__D__V__ L__L__P__C      26
                                                        _M__F_ _Y__C__L__      5
                                                 E8 cds ->

      181 taatttttgt aattattttt tgtctaatgc tgagaagctg ctttttgatc attttgattt
          __N__F__C_ _N__Y__F__ L__S__N__A __E__K__L_ _L__F__D__ H__F__D__L      46
          V__I__F__V __I__I__F_ _C__L__M__ L__R__S__C __F__L__I_ _I__L__I__      25

      241 gcatcttgtc tggagagaca atttggtgtt tggatgctgt caagggtgtg ctagaactgt
          __H__L__V_ _W__R__D__ N__L__V__F __G__C__C_ _Q__G__C__ A__R__T__V      66
          C__I__L__S __G__E__T_ _I__W__C__ L__D__A__V __K__G__V_ _L__E__L__      45

      301 tagcctattg gagtttgttt tatattatca ggagtcttat gaggtaccgg aaaTAGaaga
          __S__L__L_ _E__F__V__ L__Y__Y__Q __E__S__Y_ _E__V__P__ E__I__E__E      86
          L__A__Y__W __S__L__F_ _Y__I__I__ R__S__L__M __R__Y__R_ _K__$_          63
                                                                  <- E8 end

      361 aattttggac agacctttat tgcaaattga actccgttgt gttacatgca taaaaaaact
          __I__L__D_ _R__P__L__ L__Q__I__E __L__R__C_ _V__T__C__ I__K__K__L      106

      421 gagtgttgct gaaaaattgg aggttgtgtc aaacggagaa agagtgcata gagttagaaa
          __S__V__A_ _E__K__L__ E__V__V__S __N__G__E_ _R__V__H__ R__V__R__N      126

      481 cagacttaaa gcaaagtgTA Gtttgtgtcg cttgtatgcT ATATAAcaAT Ggtgggcgaa
                                                          _M __V__G__E_      4
          __R__L__K_ _A__K__C__ S__L__C__R __L__Y__A_ _I__$_                  140
                    E7 orf start ->                 E7 cds ->
                                                      <- E6 end
                                          signal ->

      541 atgccagcac taaaggacct ggttcttcaa cttgaaccaa gcgtcctaga tttagatctt
          _M__P__A__ L__K__D__L __V__L__Q_ _L__E__P__ S__V__L__D __L__D__L_      24
```

```
601 tattgttacg aggaggtgcc tcctgatgac atagaggagg agttagtgtc gcctcagcaa
    _Y__C__Y__ E__E__V__P __P__D__D_ _I__E__E__ E__L__V__S __P__Q__Q_      44

661 ccttatgctg tcgttgcttc ctgtgcctat tgcgagaaac tggttcgatt gaccgtcctc
    _P__Y__A__ V__V__A__S __C__A__Y_ _C__E__K__ L__V__R__L __T__V__L_      64

721 gcggatcaca gcgccattag acagctggag gaactccttc tgcgatcttT GAacatcgtg
    _A__D__H__ S__A__I__R __Q__L__E_ _E__L__L__ L__R__S__L __N__I__V_      84
                                                E1 orf start ->

781 tgcccactgt gcaccctaca gcgacagTAA aATGgcagat AATAAAgGTa ctgaaaacga
    _C__P__L__ C__T__L__Q __R__Q__$_                                      93
                                    _M__A__D_ _N__K__G__ T__E__N__D       10
                                E1 cds ->       5 sj /\
                                   <- E7 end    -> signal

841 ttggtttttg gtggaggcga cagattgtga ggaaacgtta gaggaaACCT CACTTGGTga
    __W__F__L_ _V__E__A__ T__D__C__E __E__T__L_ _E__E__T__ S__L__G__D      30
                                                      -> E2 bind

901 cctagataat gtttcttgtg ttagcgactt atctgattta ttagacgagg cgccgcaaag
    __L__D__N_ _V__S__C__ V__S__D__L __S__D__L_ _L__D__E__ A__P__Q__S      50

961 ccaggggaat tccctggaat tgttccacaa gcaagaatcg ctggaaagcg aacaggaact
    __Q__G__N_ _S__L__E__ L__F__H__K __Q__E__S_ _L__E__S__ E__Q__E__L      70

1021 taatgcttta aaacgaaagt tactttacag tcctcaggcg agaagcgcgg acgaaacaga
     __N__A__L_ _K__R__K__ L__L__Y__S __P__Q__A_ _R__S__A__ D__E__T__D      90

1081 cattgctagc attagtccta gattagaaac tatttctatt acaaagcaag acaaaaaaag
     __I__A__S_ _I__S__P__ R__L__E__T __I__S__I_ _T__K__Q__ D__K__K__R      110

1141 gtatcgaagg caactgtttt ctcaggatga tagtggttta gagctatcgc tgcttcagga
     __Y__R__R_ _Q__L__F__ S__Q__D__D __S__G__L_ _E__L__S__ L__L__Q__D      130

1201 tgaaactgaa aatattgatg aatcgacaca gGTagatcaa cagcagaaag aacatactgg
     __E__T__E_ _N__I__D__ E__S__T__Q __V__D__Q_ _Q__Q__K__ E__H__T__G      150
                                 5 sj /\

1261 ggaagttggg gccgctgggg tgaacatttt gaaagctagt aatatccgcg ccgcattatt
     __E__V__G_ _A__A__G__ V__N__I__L __K__A__S_ _N__I__R__ A__A__L__L      170

1321 aagcagattt aaagatacgg ctggcgtcag ttttacagAC CTGACGCGGT cgtacaagag
     __S__R__F_ _K__D__T__ A__G__V__S __F__T__D_ _L__T__R__ S__Y__K__S      190
                                                -> E2 bind

1381 caacaaaacc tgttgtggag attgggtttt ggcagtttgg ggtgtccgtg aaaatttaat
     __N__K__T_ _C__C__G__ D__W__V__L __A__V__W_ _G__V__R__ E__N__L__I      210

1441 tgacagtgta aaagaattat tgcaaaccca ttgtgtgtat attcaattgg aacatgcagt
     __D__S__V_ _K__E__L__ L__Q__T__H __C__V__Y_ _I__Q__L__ E__H__A__V      230

1501 aactgaaaaa aatagatttt tatttttatt ggtacgattt aaagcccaga aaagtagaga
     __T__E__K_ _N__R__F__ L__F__L__L __V__R__F_ _K__A__Q__ K__S__R__E      250

1561 gactgtgata aaacttataa ccacaattct tccagttgat gctagctata ttttgtctga
     __T__V__I_ _K__L__I__ T__T__I__L __P__V__D_ _A__S__Y__ I__L__S__E      270

1621 gcctccaaaa tcaagaagtg tggctgctgc attattttgg TATAAAagat ctatgtcttc
     __P__P__K_ _S__R__S__ V__A__A__A __L__F__W_ _Y__K__R__ S__M__S__S      290
                                          signal ->

1681 aactgttttt acatggggta caactttgga gtggattgca cagcaaaccc ttattaatca
     __T__V__F_ _T__W__G__ T__T__L__E __W__I__A_ _Q__Q__T__ L__I__N__H      310

1741 tcagttagat tccgaaagtc cctttgagct ttgtaaaatg gttcagtggg cctatgataa
     __Q__L__D_ _S__E__S__ P__F__E__L __C__K__M_ _V__Q__W__ A__Y__D__N      330

1801 tggacataca gaagagtgta aaattgcata ttattatgct gttttagcag atgaggatga
     __G__H__T_ _E__E__C__ K__I__A__Y __Y__Y__A_ _V__L__A__ D__E__D__E      350
```

HPV-1a Variants

```
1861 aaatgcaagg gcatttctaa gctctaattc acaggcaaaa tatgtgaaag actgtgcaca
     __N__A__R_ _A__F__L__ S__S__N__S __Q__A__K_ _Y__V__K__ D__C__A__Q      370

1921 aatggtaaga cactatttac gtgctgagat ggcacaaatg tctatgtcag agtggatttt
     __M__V__R_ _H__Y__L__ R__A__E__M __A__Q__M_ _S__M__S__ E__W__I__F      390

1981 tagaaaacta gataatgtag aaggttctgg taattggaaa gaaattgtaa gatttttaag
     __R__K__L_ _D__N__V__ E__G__S__G __N__W__K_ _E__I__V__ R__F__L__R      410

2041 atttcaagaa gttgaattta taagctttat gattgcattt aaagatttgt tatgtggtaa
     __F__Q__E_ _V__E__F__ I__S__F__M __I__A__F_ _K__D__L__ L__C__G__K      430

2101 gccaaagaaa aactgtttgt taatatttgg acctccaaat acaggaaaat caatgttttg
     __P__K__K_ _N__C__L__ L__I__F__G __P__P__N_ _T__G__K__ S__M__F__C      450

2161 tacaagttta ttaaagttgt taggagggaa agtgatttca tactgtaaca gtaaaagtca
     __T__S__L_ _L__K__L__ L__G__G__K __V__I__S_ _Y__C__N__ S__K__S__Q      470

2221 gttttggttg cagcctctgg ctgatgctaa gatagggcta ttagatgatg caacaaagcc
     __F__W__L_ _Q__P__L__ A__D__A__K __I__G__L_ _L__D__D__ A__T__K__P      490

2281 atgttgggat tatatggaca ttttatatgag aaatgcattg gatggtaaca ctatttgtat
     __C__W__D_ _Y__M__D__ _I__Y__M__R __N__A__L_ _D__G__N__ T__I__C__I      510

2341 tgatttaaaa catagagctc ctcaacaaat taaatgccca cctttactta ttactagtaa
     __D__L__K_ _H__R__A__ P__Q__Q__I __K__C__P_ _P__L__L__ I__T__S__N      530

2401 tattgatgtt aaatcagata cctgttggat gtatttgcat agtagaatat cagcttttaa
     __I__D__V_ _K__S__D__ T__C__W__M __Y__L__H_ _S__R__I__ S__A__F__K      550

2461 atttgctcat gagtttccat ttaaagacaa tggtgatcca ggattttcct taacagacga
     __F__A__H_ _E__F__P__ F__K__D__N __G__D__P_ _G__F__S__ L__T__D__E      570

2521 aaattggaaa tctttctttg aaAGgttttg gcaacagtta gaatTAAgtg accaagaaga
     __N__W__K_ _S__F__F__ E__R__F__W __Q__Q__L_ _E__L__S__ D__Q__E__D      590
                              /\ 3 sj     E2 orf start ->

2581 cgagggaaac gATGgaaaac ctcagcagtc gcttagactt actgcaagag cagctaatga
     __E__G__N_ _D__G__K__ P__Q__Q__S __L__R__L_ _T__A__R__ A__A__N__E      610
                _M__E__N_ _L__S__S__ R__L__D__L __L__Q__E_ _Q__L__M__      16
          E2 cds ->

2641 acctataTGA acaggacagt aaattgatag aagatcaaat taagcagtgg aatctaatta
     __P__I__$_                                                          612
     N__L__Y__E __Q__D__S_ _K__L__I__ E__D__Q__I __K__Q__W_ _N__L__I__      36
             <- E1 end

2701 gacaagaaca agttcttttc catttcgcca gaaaaaatgg ggtaatgaga attggattgc
     R__Q__E__Q __V__L__F_ _H__F__A__ R__K__N__G __V__M__R_ _I__G__L__      56

2761 aggcagttcc atctttagcg tcctcacagg agaaggcaaa gacagctatt gaaatggtgt
     Q__A__V__P __S__L__A_ _S__S__Q__ E__K__A__K __T__A__I_ _E__M__V__      76

2821 tacattaga gtctttaaag gactcacctt atggcacaga ggattggtca cttcaagaca
     L__H__L__E __S__L__K_ _D__S__P__ Y__G__T__E __D__W__S_ _L__Q__D__      96

2881 ctagcagaga gctgtttttg gcacccccag ctggcacctt caagaagagt ggcagcacac
     T__S__R__E __L__F__L_ _A__P__P__ A__G__T__F __K__K__S_ _G__S__T__      116

2941 ttgaggttac ctatgacaat aaccctgata atcagacaag gcacacaatt tggaatcatg
     L__E__V__T __Y__D__N_ _N__P__D__ N__Q__T__R __H__T__I_ _W__N__H__      136

3001 tgtattatca aaatggggac gatgtatgga gaaaagtatc cagtggtgtt gatgctgtag
     V__Y__Y__Q __N__G__D_ _D__V__W__ R__K__V__S __S__G__V_ _D__A__V__      156

3061 gagtgtacta tttagaacac gatggcTATA AAaattatta tgtgttattt gctgaggagg
     G__V__Y__Y __L__E__H_ _D__G__Y__ K__N__Y__Y __V__L__F_ _A__E__E__      176
                          signal ->

3121 cctctaagta cagcacaaca ggacaatatg ctgTAAatta caggggtaaa aggtttacaa
     A__S__K__Y __S__T__T_ _G__Q__Y__ A__V__N__Y __R__G__K_ _R__F__T__      196
                              E4 orf start ->
```

```
  3181 ATGttatgtc ttccactAGc tccccaaggg ctgctggggc tcctgcagta cactccgact
       _M__L__C__ L__P__L__A __P__Q__G_ _L__L__G__ L__L__Q__Y __T__P__T_      20
       N__V__M__S __S__T__S_ _S__P__R__ A__A__G__A __P__A__V_ _H__S__D__      216
E4 cds ->                 /\ 3 sj

  3241 acccaaccct atccgagagT GAcaccgccc agcaatcgac gtccatcgac tacaccgaac
       _T__Q__P__ Y__P__R__V __T__P__P_ _S__N__R__ R__P__S__T __T__P__N_      40
                                _H__R__P __A__I__D_ _V__H__R__ L__H__R__T      13
       Y__P__T__L __S__E__S_ _D__T__A__ Q__Q__S__T __S__I__D_ _Y__T__E__      236
                  E3 orf start ->

  3301 tcccaggaca gggggagacc tcgcaggtcc gacaaagaca gcagaaaaca cctgtacgca
       _S__Q__D__ R__G__R__P __R__R__S_ _D__K__D__ S__R__K__H __L__Y__A_      60
       __P__R__T_ _G__G__D__ L__A__G__P __T__K__T_ _A__E__N__ T__C__T__Q      33
       L__P__G__Q __G__E__T_ _S__Q__V__ R__Q__R__Q __Q__K__T_ _P__V__R__      256

  3361 gacggcctta cggacggcga agatccagaa gtcccagagg tggaggacga agagaaggag
       _D__G__L__ T__D__G__E __D__P__E_ _V__P__E__ V__E__D__E __E__K__E_      80
       __T__A__L_ _R__T__A__ K__I__Q__K __S__Q__R_ _W__R__T__ K__R__R__R      53
       R__R__P__Y __G__R__R_ _R__S__R__ S__P__R__G __G__G__R_ _R__E__G__      276

  3421 aatcaacgcc ctcTAGgaca cccggatctg tcccttctgc gcgagacgtt ggaagtatac
       _N__Q__R__ P__L__G__H __P__D__L_ _S__L__L__ R__E__T__L __E__V__Y_      100
       __I__N__A_ _L__$_                                                      57
       E__S__T__P __S__R__T_ _P__G__S__ V__P__S__A __R__D__V_ _G__S__I__      296
                      <- E3 orf end

  3481 acacaacgcc tcaaaaggga cattcttcaa gacttagacg acttctgcag gaagcttggg
       _T__Q__R__ L__K__R__D __I__L__Q_ _D__L__D__ D__F__C__R __K__L__G_      120
       H__T__T__P __Q__K__G_ _H__S__S__ R__L__R__R __L__L__Q_ _E__A__W__      316

  3541 atccacccgt ggtctgtgTA Aaagggggtg ccaatcagct taagtgtctc agGTacagac
       _I__H__P__ W__S__V__$ _                                                126
       D__P__P__V __V__C__V_ _K__G__G__ A__N__Q__L __K__C__L_ _R__Y__R__      336
                            <- E4 end                    5 sj /\

  3601 ttaaagcatc tactcaagtt gactttgaca gcataagcac cacatggcat tggacagata
       L__K__A__S __T__Q__V_ _D__F__D__ S__I__S__T __T__W__H_ _W__T__D__      356

  3661 gaaaaaacac cgagaggata ggtagtgcta gaatgttagt aaagtttatt gatgaggctc
       R__K__N__T __E__R__I_ _G__S__A__ R__M__L__V __K__F__I_ _D__E__A__      376

  3721 aacgagagaa gtttcttgag agagttgctt tgcccagatc agtgtctgtg tttttgggac
       Q__R__E__K __F__L__E_ _R__V__A__ L__P__R__S __V__S__V_ _F__L__G__      396

  3781 agtttaatgg gtctTAAaat taatggaagt tgattttgct tggacgtgtg tacatagtcc
       Q__F__N__G __S__$_                                                     401
                      <- E2 end

  3841 ctgTATATAT tcccctccta cccccacata cctTGAagct tgcaacattg TAAcaaATGt
                                              _S__ L__Q__H__C __N__K__C_      8
                                                                  _M__       1
                                                        ▲
   signal ->                     E5 orf start ->    L2 orf start ->
                                                             L2 cds ->

  3901 atcgcctacg tagaaaacgc gctgccccca aagaTATATA cccctcatgc aaaatatcaa
       _I__A__Y__ V__E__N__A __L__P__P_ _K__I__Y__ T__P__H__A __K__Y__Q_      28
       Y__R__L__R __R__K__R_ _A__A__P__ K__D__I__Y __P__S__C_ _K__I__S__      21
                                     signal ->

  3961 acacctgccc acctgacatt caaAATAAAa ttgagcatac aacaattgct gataaaatat
       _T__P__A__ H__L__T__F __K__I__K_ _L__S__I__ Q__Q__L__L __I__K__Y_      48
       N__T__C__P __P__D__I_ _Q__N__K__ I__E__H__T __T__I__A_ _D__K__I__      41
                         signal ->

  4021 tgcaatatgg cagtctggga gtttttttgg gaggtttggg cattggaaca gccagaggct
       _C__N__M__ A__V__W__E __F__F__W_ _E__V__W__ A__L__E__Q __P__E__A_      68
       L__Q__Y__G __S__L__G_ _V__F__L__ G__G__L__G __I__G__T_ _A__R__G__      61
```

```
  4081 ctggaggaag aattggttat actcccctcg gtgagggtgg tggggttaga gttgctactc
       _L__E__E__ E__L__V__I __L__P__S_ _V__R__V__ V__G__L__E __L__L__L_      88
       S__G__G__R __I__G__Y_ _T__P__L__ G__E__G__G __G__V__R_ _V__A__T__      81

  4141 gtccaactcc agTAAggcct acaatacctg tggaaacagt aggccccagt gaaattttcc
       _V__Q__L__ Q__$_                                                     92
       R__P__T__P __V__R__P_ _T__I__P__ V__E__T__V __G__P__S_ _E__I__F__     101
                     <- E5 orf end

  4201 ccatagatgt tgtagatcct acaggccctg ctgttattcc cctacaagat ttaggtagag
       P__I__D__V __V__D__P_ _T__G__P__ A__V__I__P __L__Q__D_ _L__G__R__     121

  4261 acttcccaat accaactgtg caggttattg cagaaattca ccctatttct gacataccaa
       D__F__P__I __P__T__V_ _Q__V__I__ A__E__I__H __P__I__S_ _D__I__P__     141

  4321 acattgttgc atcttcaaca aatgaaggag aatctgccat attagatgtg ttacgagggga
       N__I__V__A __S__S__T_ _N__E__G__ E__S__A__I __L__D__V_ _L__R__G__     161

  4381 atgcaaccat acgcactgtt tcaagaacac aatacaataa cccctctttc actgttgcat
       N__A__T__I __R__T__V_ _S__R__T__ Q__Y__N__N __P__S__F_ _T__V__A__     181

  4441 ctacatctaa tataagtgct ggagaagcat caacatcaga tattgtattt gttagcaatg
       S__T__S__N __I__S__A_ _G__E__A__ S__T__S__D __I__V__F_ _V__S__N__     201

  4501 gttcaggtga cagggtggtg ggcgaggata tccccttggt agaattaaac ttaggccttg
       G__S__G__D __R__V__V_ _G__E__D__ I__P__L__V __E__L__N_ _L__G__L__     221

  4561 aaacagacac atcttctgtt gtacaagaaa cagcattttc cagcagcaca CCAATtgctg
       E__T__D__T __S__S__V_ _V__Q__E__ T__A__F__S __S__S__T_ _P__I__A__     241
                                                    signal ->

  4621 aaagaccctc ttttaggccc tcaagattcT ATAATAGgcg tctatatgaa caggtgcaag
       E__R__P__S __F__R__P_ _S__R__F__ Y__N__R__R __L__Y__E_ _Q__V__Q__     261
                                signal ->

  4681 tacaagaccc taggttcgtt gagcagccac agtcaatggt cacttttgat aatccagcat
       V__Q__D__P __R__F__V_ _E__Q__P__ Q__S__M__V __T__F__D_ _N__P__A__     281

  4741 ttgagccaga gcttgatgag gtgtctatta tcttccaaag agacttagat gctcttgctc
       F__E__P__E __L__D__E_ _V__S__I__ I__F__Q__R __D__L__D_ _A__L__A__     301

  4801 agacaccagt gcctgaattt agagatgtag tttatctgag caagcccaca ttttcgcggg
       Q__T__P__V __P__E__F_ _R__D__V__ V__Y__L__S __K__P__T_ _F__S__R__     321

  4861 aaccaggggg acggttaagg gttagccgcc ttggcaaaag ttcaactatt cgtacacgcc
       E__P__G__G __R__L__R_ _V__S__R__ L__G__K__S __S__T__I_ _R__T__R__     341

  4921 tgggcacagc aattggcgcc agaacccact ttttctatga tttaagttct attgctccag
       L__G__T__A __I__G__A_ _R__T__H__ F__F__Y__D __L__S__S_ _I__A__P__     361

  4981 aagactcaat tgaattattg cctttaggtg agcatagtca aacaacagtc attagttcca
       E__D__S__I __E__L__L_ _P__L__G__ E__H__S__Q __T__T__V_ _I__S__S__     381

  5041 acttaggtga cacagcattt atacaaggtg agacagcaga ggatgactta gaagttatct
       N__L__G__D __T__A__F_ _I__Q__G__ E__T__A__E __D__D__L_ _E__V__I__     401

  5101 ctttagaaac accacaatta tattcagaag aagagctttt agacacaaac gaaagtgtgg
       S__L__E__T __P__Q__L_ _Y__S__E__ E__E__L__L __D__T__N_ _E__S__V__     421

  5161 gcgaaaattt gcaacttact attactaact cagagggtga ggtttctata ctagatttaa
       G__E__N__L __Q__L__T_ _I__T__N__ S__E__G__E __V__S__I_ _L__D__L__     441

  5221 cacaaagcag agtcaggcca ccttttggca ctgaagatac tagcttgcat gtatattacc
       T__Q__S__R __V__R__P_ _P__F__G__ T__E__D__T __S__L__H_ _V__Y__Y__     461

  5281 caaattcttc taaagggact ccaataatta atcctgaaga atcatttaca cctttggtta
       P__N__S__S __K__G__T_ _P__I__I__ N__P__E__E __S__F__T_ _P__L__V__     481

  5341 ttaTAGctct taacaactca acaggggatt ttgagttaca tcctagtctt agaaagcgtc
       I__I__A__L __N__N__S_ _T__G__D__ F__E__L__H __P__S__L_ _R__K__R__     501
L1 orf start ->
```

```
5401 gtaaaagagc ttATGtaTAA tgtttttcAG atggctgtct ggttaccagc gcagaataag
                  _M__Y__N __V__F__Q_ _M__A__V__ W__L__P__A __Q__N__K_      16
     R__K__R__A __Y__V__$_                                                 507
           L1 cds ->    <- L2 end   /\ 3 sj

5461 ttctatcttc ctccccagcc catcactaga atcctgtcca ctgatgaata tgtaaccaga
     _F__Y__L__ P__P__Q__P __I__T__R_ _I__L__S__ T__D__E__Y __V__T__R_      36

5521 accaatctct tctaccatgc aacatctgaa cgtctactgc tggtcggaca tcctttgttt
     _T__N__L__ F__Y__H__A __T__S__E_ _R__L__L__ L__V__G__H __P__L__F_      56

5581 gagatctcca gtaatcaaac tgtaactata ccaaaagtgt caccaaatgc atttagagtt
     _E__I__S__ S__N__Q__T __V__T__I_ _P__K__V__ S__P__N__A __F__R__V_      76

5641 tttagggtgc gttttgctga tccaaataga tttgcatttg gggataaggc aatttttaat
     _F__R__V__ R__F__A__D __P__N__R_ _F__A__F__ G__D__K__A __I__F__N_      96

5701 ccagaaacag aaagattagt ttggggccta agagggatag agataggtag aggccagcct
     _P__E__T__ E__R__L__V __W__G__L_ _R__G__I__ E__I__G__R __G__Q__P_     116

5761 ttaggtatag gaataacggg ccaccctctt ttaaataagt tagatgatgc agaaaatcca
     _L__G__I__ G__I__T__G __H__P__L_ _L__N__K__ L__D__D__A __E__N__P_     136

5821 acaaattata ttaatactca tgcaaatgga gattctagac aaaatactgc ttttgatgca
     _T__N__Y__ I__N__T__H __A__N__G_ _D__S__R__ Q__N__T__A __F__D__A_     156

5881 aaacagacac aaatgttcct cgtcggctgt actcctgctt caggtgaaca ctggacaagt
     _K__Q__T__ Q__M__F__L __V__G__C_ _T__P__A__ S__G__E__H __W__T__S_     176

5941 agtcgttgcc caggggaaca agtgaaactt ggggactgcc ccagggtgca aatgatagag
     _S__R__C__ P__G__E__Q __V__K__L_ _G__D__C__ P__R__V__Q __M__I__E_     196

6001 tctgtcatag aagatggtga catgatggat attggttttg gggctatgga ttttgctgct
     _S__V__I__ E__D__G__D __M__M__D_ _I__G__F__ G__A__M__D __F__A__A_     216

6061 ttacagcaag acaagtctga tgtcccttta gatgttgttc aagcaacatg caaatatcct
     _L__Q__Q__ D__K__S__D __V__P__L_ _D__V__V__ Q__A__T__C __K__Y__P_     236

6121 gattatatca gaatgaacca tgaagcctat ggcaactcta tgttttttt tgcacgtcgc
     _D__Y__I__ R__M__N__H __E__A__Y_ _G__N__S__ M__F__F__F __A__R__R_     256

6181 gagcaaatgt ataccaggca ctttttttact cgcgggggtt cggtgggtga taaggaggca
     _E__Q__M__ Y__T__R__H __F__F__T_ _R__G__G__ S__V__G__D __K__E__A_     276

6241 gtcccacaaa gcctgtattt aacagcagat gctgaaccaa gaacaacttt agcaacaaca
     _V__P__Q__ S__L__Y__L __T__A__D_ _A__E__P__ R__T__T__L __A__T__T_     296

6301 aattatgtag gcacaccaag tggctctatg gtttcatctg atgtccaatt gtttaataga
     _N__Y__V__ G__T__P__S __G__S__M_ _V__S__S__ D__V__Q__L __F__N__R_     316

6361 tcttactggc ttcagcgatg tcaaggccag aataatggca tttgctggag aaaccagtta
     _S__Y__W__ L__Q__R__C __Q__G__Q_ _N__N__G__ I__C__W__R __N__Q__L_     336

6421 tttattacag ttggagataa taccagagga acaagtttat ctatcagtat gaaaaacaat
     _F__I__T__ V__G__D__N __T__R__G_ _T__S__L__ S__I__S__M __K__N__N_     356

6481 gcaagtacta catattccaa tgctaatttt aatgattttc taagacatac tgaagaattt
     _A__S__T__ T__Y__S__N __A__N__F_ _N__D__F__ L__R__H__T __E__E__F_     376

6541 gatctttctt ttatagttca gctttgtaaa gtaaagttaa ctcccgaaaa tctagcctac
     _D__L__S__ F__I__V__Q __L__C__K_ _V__K__L__ T__P__E__N __L__A__Y_     396

6601 attcatacaa tggaccctaa tattttagag gattggcaac tatctgtatc tcaaccacct
     _I__H__T__ M__D__P__N __I__L__E_ _D__W__Q__ L__S__V__S __Q__P__P_     416

6661 accaatcctc tagaagatca atataggttt ttagggtctt ccttggcagc aaaatgtcca
     _T__N__P__ L__E__D__Q __Y__R__F_ _L__G__S__ S__L__A__A __K__C__P_     436

6721 gaacaggcgc ctcctgagcc ccagactgat ccttatagtc aaTATAAATt ctgggaagtc
     _E__Q__A__ P__P__E__P __Q__T__D_ _P__Y__S__ Q__Y__K__F __W__E__V_     456
                                        signal ->
```

```
6781 gatctcacag aaaggatgtc cgaacaatta gaccaatttc cactaggaag gaaatttcta
     _D__L__T__ E__R__M__S __E__Q__L_ _D__Q__F__ P__L__G__R __K__F__L_      476

6841 tatcaaagtg gcatgacaca acgtactgct actagttcca ccacaaagcg caaaacagtg
     _Y__Q__S__ G__M__T__Q __R__T__A_ _T__S__S__ T__T__K__R __K__T__V_      496

6901 cgtgtatcta cgtcagccaa gcgcaggcgt aaggctTAGT ATATAtTATA TATAactata
     _R_ V _S__ T__S__A__K __R__R__R_ _K__A__$_                            508
                                            <- L1 end
                                        signal ->

6961 tttattagta gattatttat TATATAtttt tatattttta tactttttat acttgtttag
                    signal ->

7021 ttctaaatag acatgtaaga tttacattag tataagtagg catgtattta cataaaatag

7081 tcttggaaac cttttattag tgaaccatca tttacaatag tgacatcata gttcatctgc
                                                          deletion
7141 aattgctatt ccatcgttct tcacatattc tacagtagtg ttctctagattgtattgcta

7201 ttttcctgtt aggcaaacaa caacatctgt acatggacca aacaacccac tttcatttta

7261 ttgtgctgca tatattccag attgttgagg atttatttgt ttagactccg gtgcattata

7321 cacaagtgtg catttttgt gttctctgat tgattgtgtg ttattttcct gcaatatgcA
                                                              signal ->

7381 ATAAAagtga gctgtccttt ctttttgtta atccctccct actccAATAA Aaaatcccta
                                              signal ->

7441 cccctaaaat ctgtttgtgc tggttttatt aataattgcg ctctttTATA TAAtaagtac
                                               signal ->
                                                          |-> mRNA start site
                                                              from P(7490)
                                                              promoter

7501 tattaacACC GCACCCGTTg tggctaatcc cttatggtat ttaaaagact acacctacag
            -> E2 bind

7561 gatgtattgt ctttattgtt tatggtttac cgcgctCCAA AGACGGTTTG CCCAAAGACG
                                            ->    repeat region start

7621 GTTTGCcaac cgcggttagg acttgtttca atttgctgcc aaacttatct ggtcgtgctc
         <- repeat region end

7681 caacgggttt ggtgccaagc acctaaaacg GTaggtgtgt actcttttca agaattaaca
                                5 sj /\
                            -> E2 bind

7741 aaggagattt ctcccgccaa attagtttcg agcgACCGAA TTCGGTcgta aaaatctaaa
                                          -> E2 bind

7801 gtgatgattg ttgtt
```

E6 nuc

```
                1
                4
                2
HPV1a           A   104-526
HPV1var         G   104-526
```

E6 aa

```
                1
                3
HPV1a           E
HPV1var         *
```

E1 nuc

```
                12
                23
                80
                31
HPV1a           AT   812-2650
HPV1var         GC   812-2650
```

E1 aa

```
                14
                59
                87
HPV1a           NI
HPV1var         DT
```

E2 nuc

```
                2
                8
                2
                5
HPV1a           T   2592-3797
HPV1var         C   2592-3797
```

E2 aa

```
                7
                8
HPV1a           H
HPV1var         *
```

E5 nuc

```
                i
                3
                8
                8
                6
HPV1a           .   3877-4155
HPV1var         A   3877-4155
```

E5 aa

```
                4
HPV1a           .
HPV1var         X
```

L2 nuc

```
                 44444
                 33333
                 37778
                 15691
HPV1a            AGAGA   3897-5420
HPV1var          TAGAG   3897-5420
```

L2 aa

```
                 11111
                 46666
                 50012
HPV1a            ARRGN
HPV1var          *QQ*S
```

L1 nuc

```
                 56
                 79
                 90
                 34
HPV1a            AG   5413-6939
HPV1var          CT   5413-6939
```

L1 aa

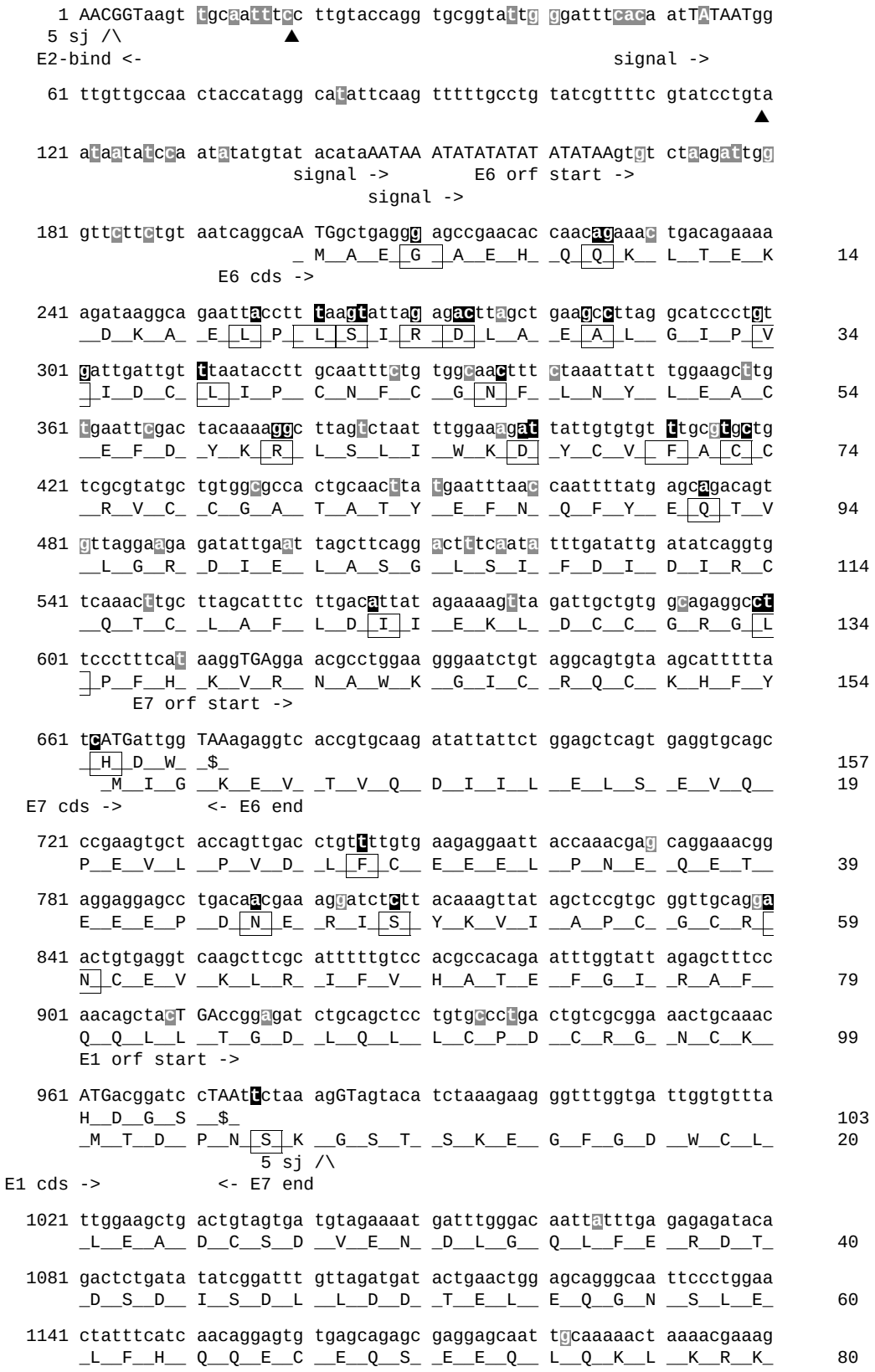
```
                 14
                 29
                 78
HPV1a            LV
HPV1var          FL
```

LCR

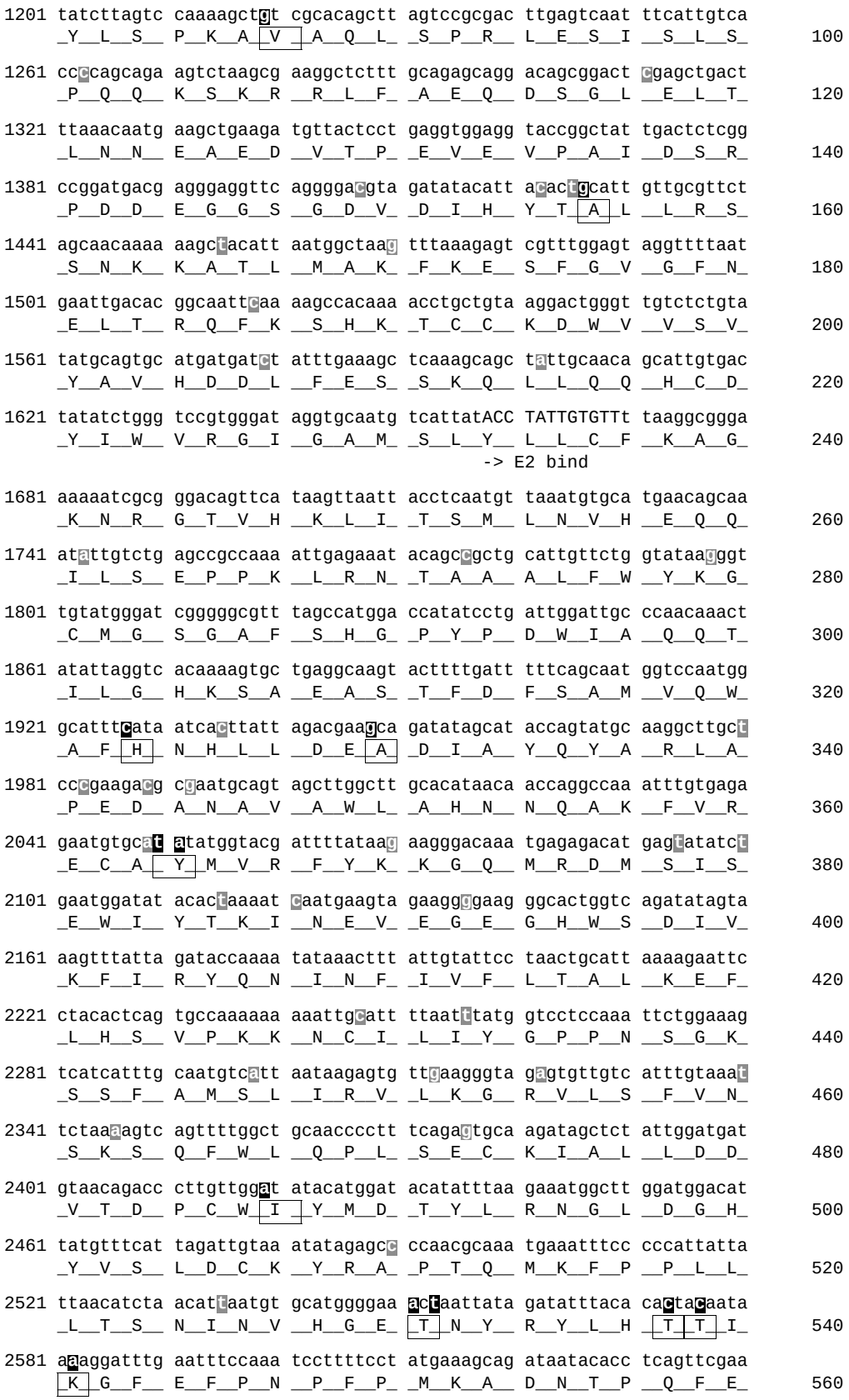
```
                 deletion
                 77777777777777777777
                 11111111111112256668
                 88899999999990073990
                 78901234567890142121
HPV1a            AGATTGTATTGCTATCGCCG   6940-7815
HPV1var          ...............TTGGC   6940-7815
```

```
LOCUS       HPV5         7746 bp ds-DNA             VRL       30-SEP-1988
DEFINITION  Human papillomavirus type 5 (HPV-5), complete genome.
ACCESSION   M17463
KEYWORDS    complete genome.
SOURCE      Human papillomavirus type 5 DNA recovered from a benign flat
            wart from an EV patient.
  ORGANISM  Human papillomavirus
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
COMMENT     Naturally occurring variants of HPV5 have been observed throughout
            the genome; the level of variation between variants appears to be
            higher than that observed in mucosal HPV types.
            Reference: Deau,M.C., Favre,M. and Orth,G.,
                Virology 184, 492-503 (1991)
                Variant Accession Nos. M73051-59, M73456-M73458, X74646-X74653, X64619-X64622
                M72884, M80460
            Reference: Deau,M.C., Favre,M., Jablonska,S., Rueda,L.A. and Orth,G.,
                J. Clin. Microbiol. 31, 2918-26 (1993)
                Variant Accession Nos. M72882, X74644, X74645, X64618, X64641-64643
            Reference: Kawase, M., Orth, G., Jablonska, S., Blanchet-Bardon, C.,
                Rueda, L.-A., Favre, M.,
                Virology 221, 189-98 (1996)
                Variant Accession Nos. U49459-U49481
            Reference: Ostrow,R.S., Zachow,K.R., and Faras,A.J.,
                Virology 158, 235-8 (1987)
                Variant Accession Nos. M18452-M18454, M22961
            Reference: Yabe,Y., Sakai,A., Hitsumoto,T., Kato,H., and Ogura,H.,
                Virology 183, 793-8 (1991)
                Variant Accession Nos. D90252
FEATURES             Location/Qualifiers
     CDS             200..673
                     /note="ORF E6 from bp 167 to 673"
                     /product="transforming protein"
                     /gene="E6"
                     /note="putative"
                     /codon_start=1
                     /translation="MAEGAEHQQKLTEKDKAELPLSIRDLAEALGIPVIDCLIPCNFC
                     GNFLNYLEACEFDYKRLSLIWKDYCVFACCRVCCGATATYEFNQFYEQTVLGRDIELA
                     SGLSIFDIDIRCQTCLAFLDIIEKLDCCGRGLPFHKVRNAWKGICRQCKHFYHDW"
     CDS             663..974
                     /note="ORF E7 from bp 618 to 974"
                     /product="transforming protein"
                     /gene="E7"
                     /note="putative"
                     /codon_start=1
                     /translation="MIGKEVTVQDIILELSEVQPEVLPVDLFCEEELPNEQETEEEPD
                     NERISYKVIAPCGCRNCEVKLRIFVHATEFGIRAFQQLLTGDLQLLCPDCRGNCKHDG
                     S"
     CDS             961..2781
                     /note="ORF E1 from bp 913 to 2781"
                     /product="replication protein"
                     /gene="E1"
                     /note="putative"
                     /codon_start=1
                     /translation="MTDPNSKGSTSKEGFGDWCLLEADCSDVENDLGQLFERDTDSDI
                     SDLLDDTELEQGNSLELFHQQECEQSEEQLQKLKRKYLSPKAVAQLSPRLESISLSPQ
                     QKSKRRLFAEQDSGLELTLNNEAEDVTPEVEVPAIDSRPDDEGGSGDVDIHYTALLRS
                     SNKKATLMAKFKESFGVGFNELTRQFKSHKTCCKDWVVSVYAVHDDLFESSKQLLQQH
                     CDYIWVRGIGAMSLYLLCFKAGKNRGTVHKLITSMLNVHEQQILSEPPKLRNTAAALF
                     WYKGCMGSGAFSHGPYPDWIAQQTILGHKSAEASTFDFSAMVQWAFHNHLLDEADIAY
                     QYARLAPEDANAVAWLAHNNQAKFVRECAYMVRFYKKGQMRDMSISEWIYTKINEVEG
                     EGHWSDIVKFIRYQNINFIVFLTALKEFLHSVPKKNCILIYGPPNSGKSSFAMSLIRV
                     LKGRVLSFVNSKSQFWLQPLSECKIALLDDVTDPCWIYMDTYLRNGLDGHYVSLDCKY
                     RAPTQMKFPPLLLTSNINVHGETNYRYLHTTIKGFEFPNPFPMKADNTPQFELTDQSW
                     KSFFTRLWTQLDLSDQEEEGEDGESQRAFQCSARSANEHL"
```

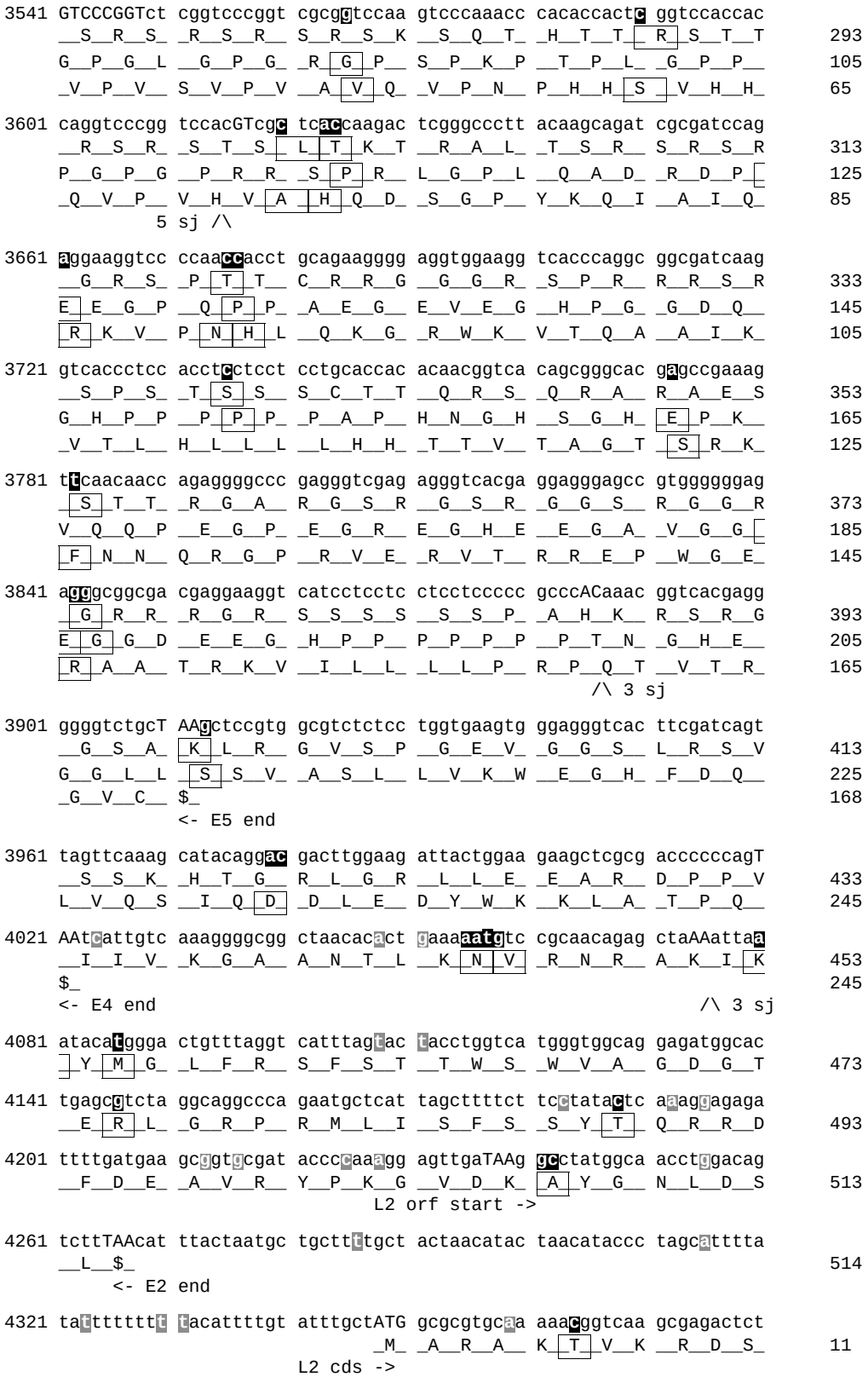
```
     CDS             2723..4267
                     /note="ORF E2 from bp 2699 to 4267"
                     /product="regulatory protein"
                     /gene="E2"
                     /note="putative"
                     /codon_start=1
                     /translation="MENLSERFNALQDQLMNIYEAAEQTLQAQIKHWQTLRKEPVLLY
                     YAREKGVTRLGYQPVPVKAVSETKAKEAIAMVLQLESLQTSDFAHEPWTLVDTSIETF
                     RSAPEGHFKKGPLPVEVIYDNDPDNANLYTMWTYVYYMDADDKWHKARSGVNHIGIYY
                     LQGTFKNYYVLFADDAKRYGTTGEWEVKVNKETVFAPVTSSTPPGSPGGQADTNTTPA
                     TPTTSTTAVDSTSRQLTTSKQPQQTETRGRRYGRRPSSKSRRSQTQQRRSRSRHRSRS
                     RSRSRSKSQTHTTRSTTRSRSTSLTKTRALTSRSRSRGRSPTTCRRGGGRSPRRRSRS
                     PSTSSSCTTQRSQRARAESSTTRGARGSRGSRGGSRGGRGRRRGRSSSSSSPAHKRSR
                     GGSAKLRGVSPGEVGGSLRSVSSKHTGRLGRLLEEARDPPVIIVKGAANTLKNVRNRA
                     KIKYMGLFRSFSTTWSWVAGDGTERLGRPRMLISFSSYTQRRDFDEAVRYPKGVDKAY
                     GNLDSL"
     CDS             <3285..4022
                     /note="ORF E4 from bp 3285 to 4022"
                     /gene="E4"
                     /note="putative"
                     /codon_start=1
                     /translation="KLIRKLCLLLSPAPRLQGRQEDKQTQTPPPRPPPPPQPPLTPRP
                     DSSPHQNSHNKPKPEEEGTDGGPPASQGDRKRSKGDQGPDTGPGLGPGRGPSPKPTPL
                     GPPPGPGPRRSPRLGPLQADRDPEEGPQPPAEGEVEGHPGGDQGHPPPPPPAPHNGHS
                     GHEPKVQQPEGPEGREGHEEGAVGGEGGDEEGHPPPPPPPTNGHEGGLLSSVASLLVK
                     WEGHFDQLVQSIQDDLEDYWKKLATPQ"
     CDS             <3406..3912
                     /note="ORF E5 from bp 3406 to 3912"
                     /gene="E5"
                     /note="putative"
                     /codon_start=1
                     /translation="LHVQTAHHIKTATTNRNQRKKVRTEALQQVKEIANAAKAIKVPT
                     PVPVSVPVAVQVPNPHHSVHHQVPVHVAHQDSGPYKQIAIQRKVPNHLQKGRWKVTQA
                     AIKVTLHLLLLHHTTVTAGTSRKFNNQRGPRVERVTRREPWGERAATRKVILLLLPRP
                     QTVTRGVC"
     CDS             4348..5904
                     /note="ORF L2 from bp 4240 to 5904"
                     /product="minor capsid protein"
                     /gene="L2"
                     /note="putative"
                     /codon_start=1
                     /translation="MARAKTVKRDSVTHIYQTCKQAGTCPPDVINKVEQTTVADNILK
                     YGSAGVFFGGLGISTGRGTGGATGYVPLGEGPGVRVGGTPTVVRPSLVPETIGPVDIL
                     PIDTVNPVEPTASSVVPLTESTGADLLPGEVETIAEIHPVPEGPSVDTPVVTTSTGSS
                     AVLEVAPEPIPPTRVRVSRTQYHNPSFQILTESTPAQGESSLADHVLVTSGSGGQRIG
                     GDITDIIELEEIPSRYTFEIEEPTPPRRSSTPLPRNQSVGRRRGFSLTNRRLVQQVQV
                     DNPLFLTQPSKLVRFAFDNPVFEEEVTNIFENDLDVFEEPPDRDFLDVRELGRPQYST
                     TPAGYVRVSRLGTRATIRTRSGAQIGSQVHFYRDLSSINTEDPIELQLLGQHSGDATI
                     VHGPVESTFIDMDISENPLSESIEAYSHDLLLDETVEDFSGSQLVIGNRRSTNSYTVP
                     RFETTRNGSYYTQDTKGYYVAYPESRNNAEIIYPTPDIPVVIIHPHDSTGDFYLHPSL
                     HRRKRKRKYL"
     CDS             5917..7467
                     /note="ORF L1 from bp 5905 to 7467"
                     /product="major capsid protein"
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /translation="MAVWHSANGKVYLPPSTPVARVQSTDEYIQRTNIYYHAFSDRLL
                     TVGHPYFNVYNINGDKLEVPKVSGNQHRVFRLKLPDPNRFALPDMSVYNPDKERLVWA
                     CRGLEIGRGQPLGVRSTGHPYFNKVKDTENSNAYITFSKDDRQDTSFDPKQIQMFIVG
                                                        ▲
                     CTPCIGEHWDKAVPCAENDQQTGLCPPIELKNTYIQDGDMADIGFGNMNFKALQDSRS
                     DVSLDIVNETCKYPDFLKMQNDIYGDACFFYARREQCYARHFFVRGGKTGDDIPRAQI
                     DNGTYKNQFYIPGADGQAQKTIGNSMYFPTVSGSLVSSDAQLFNRPFWLQRAQGHNNG
                     ILWANQMFITVVDNTRNTNFSISVYNQAGALKDVADYNADQFREYQRHVEEYEISLIL
                     QLCKVPLKAQVLAQINAMNSSLLEDWQLGFVPTPDNPIQDTYRYIDSLATRCPDKNPP
                     KEKEDPYKGLHFWDVDLTERLSLDLDQYSLGRKFLFQAGLQQTTVNGTKAVSYKGSNR
                     GTKRKRKN"
BASE COUNT     2376 a   1547 c   1736 g   2087 t
```


1 AACGGTaagt tgcaatttcc ttgtaccagg tgcggtatg ggatttcaca atTATAATgg
5 sj /\
E2-bind <-
signal ->
61 ttgttgccaa ctaccatagg cattattcaag tttttgcctg tatcgttttc gtatcctgta
121 ataatatcca atatatgtat acataAATAA ATATATATAT ATATAAgtgt ctagattgg
signal ->
E6 orf start ->
signal ->
181 gttgttgtgt aatcaggcaA TGgctgaggg agccgaacac caacagaaac tgacagaaaa
_ M__A__E__G__A__E__H__Q__Q__K__ L__T__E__K 14
E6 cds ->
241 agataaggca gaattacctt taagtattag agacttagct gaagcgttag gcatccctgt
__D__K__A__ _E__L__P__ L__S__I__R__D__L__A__ _E__A__L__ G__I__P__V 34
301 gattgattgt ttaatacctt gcaatttctg tggcaacttt ctaaattatt tggaagctg
_I__D__C__ L__I__P__ C__N__F__C __G__N__F__ _L__N__Y__ L__E__A__C 54
361 tgaattcgac tacaaaaggc ttagtctaat ttggaaagat tattgtgtgt ttgcgtgctg
__E__F__D__ _Y__K__R__ L__S__L__I __W__K__D__ _Y__C__V__ F__A__C__C 74
421 tcgcgtatgc tgtggcgcca ctgcaactta tgaatttaac caattttatg agcagacagt
__R__V__C__ _C__G__A__ T__A__T__Y __E__F__N__ _Q__F__Y__ E__Q__T__V 94
481 gttaggaaga gatattgaat tagcttcagg actttcaata tttgatattg atatcaggtg
__L__G__R__ _D__I__E__ L__A__S__G __L__S__I__ _F__D__I__ D__I__R__C 114
541 tcaaactgc ttagcatttc ttgacattat agaaaagtta gattgctgtg gcagaggcct
__Q__T__C__ _L__A__F__ L__D__I__I __E__K__L__ _D__C__C__ G__R__G__L 134
601 tccctttcat aaggTGAgga acgcctggaa gggaatctgt aggcagtgta agcattttta
_P__F__H__ _K__V__R__ N__A__W__K __G__I__C__ _R__Q__C__ K__H__F__Y 154
E7 orf start ->
661 tcATGattgg TAAagaggtc accgtgcaag atattattct ggagctcagt gaggtgcagc
_H__D__W__ _$_ 157
_M__I__G __K__E__V__ _T__V__Q__ D__I__I__L __E__L__S__ _E__V__Q__ 19
E7 cds -> <- E6 end
721 ccgaagtgct accagttgac ctgttttgtg aagaggaatt accaaacgag caggaaacgg
P__E__V__L __P__V__D__ _L__F__C__ E__E__E__L __P__N__E__ _Q__E__T__ 39
781 aggaggagcc tgacaacgaa aggatctctt acaaagttat agctccgtgc ggttgcagga
E__E__E__P __D__N__E__ _R__I__S__ Y__K__V__I __A__P__C__ _G__C__R__ 59
841 actgtgaggt caagcttcgc atttttgtcc acgccacaga atttggtatt agagctttcc
N__C__E__V __K__L__R__ _I__F__V__ H__A__T__E __F__G__I__ _R__A__F__ 79
901 aacagctacT GAccggagat ctgcagctcc tgtgccctga ctgtcgcgga aactgcaaac
Q__Q__L__L __T__G__D__ _L__Q__L__ L__C__P__D __C__R__G__ _N__C__K__ 99
E1 orf start ->
961 ATGacggatc cTAAttctaa agGTagtaca tctaaagaag ggtttggtga ttggtgttta
H__D__G__S __$_ 103
_M__T__D__ P__N__S__K __G__S__T__ _S__K__E__ G__F__G__D __W__C__L__ 20
5 sj /\
E1 cds -> <- E7 end
1021 ttggaagctg actgtagtga tgtagaaaat gatttgggac aattatttga gagagataca
_L__E__A__ D__C__S__D __V__E__N__ _D__L__G__ Q__L__F__E __R__D__T__ 40
1081 gactctgata tatcggattt gttagatgat actgaactgg agcagggcaa ttccctggaa
_D__S__D__ I__S__D__L __L__D__D__ _T__E__L__ E__Q__G__N __S__L__E__ 60
1141 ctatttcatc aacaggagtg tgagcagagc gaggagcaat tgcaaaaact aaaacgaaag
_L__F__H__ Q__Q__E__C __E__Q__S__ _E__E__Q__ L__Q__K__L __K__R__K__ 80

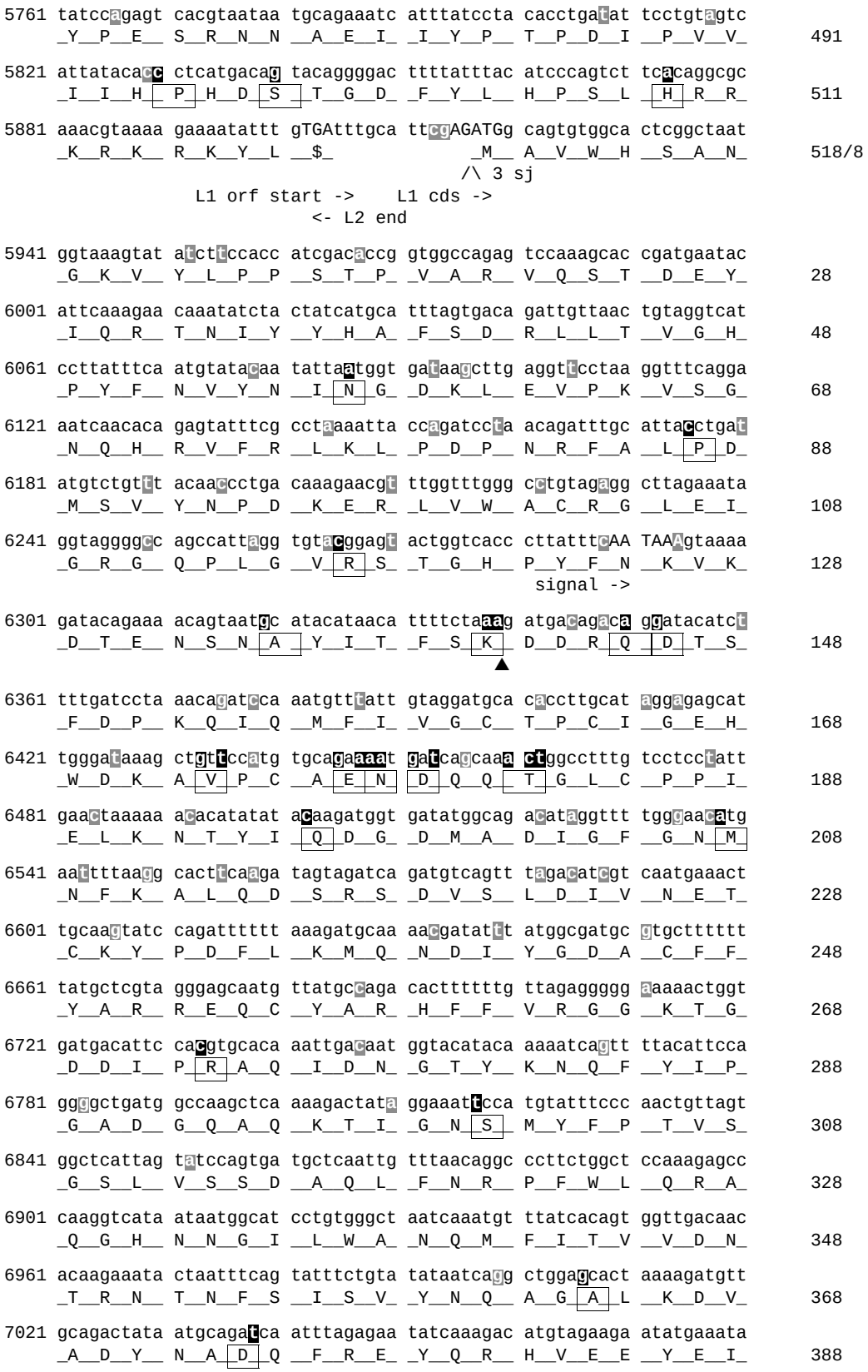
HPV-5 Variants



```
2641 ctaactgacc aaagctggaa atctttttt acaAGgcttt ggacacaatt agaccTGAgt
     _L__T__D__ Q__S__W__K __S__F__F_ _T__R__L__ W__T__Q__L __D__L__S_      580
                                        /\ 3 sj     E2 orf start ->

2701 gatcaagaag aggagggcga ggATGgagaa tctcagcgag cgtttcaatg ctctgcaaga
                             _M__E__N __L__S__E_ _R__F__N__ A__L__Q_ D      13
     _D__Q__E__ E__E__G__E __D__G__E_ _S__Q__R__ A__F__Q__C __S__A__R_      600
                       E2 cds ->

2761 tcagctaatg aacatttaTG Aagctgcaga acaaacattg caggcacaaa ttaaacattg
     _Q__L__M_ _N__I__Y__ E__A__A__E __Q__T__L_ _Q__A__Q__ I__K__H__W      33
     S_ A__N__ E__H__L__$ _                                               606
                        <- E1 end

2821 gcaaaccttA cgaaaagaac ctgtattact ctactatgct agggagaaag gtgttacaag
     __Q__T__L_ _R__K__E__ P_ _V__L__L __Y__Y__A_ _R__E__K__ G__V__T__R      53

2881 gcttggatat caacctgtgc ctgtaaaggc agtatcagaa acaaaggcta aagagcgcat
     __L__G__Y_ _Q__P__V__ P__V__K__A __V__S__E_ _T__K__A__ K__E__A__I      73

2941 agcaatggtg ctgcagcttg agtcactaca gacatctgat tttgctcatg agccatggac
     __A__M__V_ _L__Q__L__ E__S__L__Q __T__S__D_ _F_ A_ H_ E__P__W__T      93

3001 tctagttgat accagcatag aaacatttag aagcgctcca gaaggtcact tcaaaaaagg
     __L__V__D_ _T__S_ I_ E__T__F__R __S__A__P_ _E__G__H__ F__K__K__G      113

3061 ccccctccct gtagaagtta tttatgacaa tgatccagat aatgccaatt tgtatacaat
     __P_ L_ P_ _V__E__V__ I__Y__D__N __D__P__D_ _N__A__N__ L__Y__T__M      133

3121 gtggactat gtgtattata tggatgcgga tgataagtgg cataagcaa gaagtggggt
     __W__T__Y_ _V__Y__Y__ M__D__A__D __D__K__W_ _H__K__A__ R__S__G__V      153

3181 gaatcacatt ggcatttatt atttacaagg aacttttaaa aactattatg tactgtttgc
     __N__H__I_ _G__I__Y__ Y__L__Q__G __T__F__K_ _N__Y__Y__ V__L__F__A      173

3241 tgacgatgcg aaaagatatg gtacaactgg agaatgggaa gTAAaagtta ataaggaaac
     __D__D__A_ _K__R__Y__ G__T__T__G __E__W__E_ _V__K__V__ N__K_ E_ T      193
                                                 _K__L_ _I__R_ K_ _      5
                                      E4 orf start ->
                                      NH2 terminus unknown

3301 tgtgtttgct cctgtcaccA Gctccacgcc tccagggtcg ccaggaggac aagcagacac
     __V__F__A_ _P__V__T__ S__S__T__P __P__G__S_ _P__G_ G_ Q__A__D__T      213
     L__C__L__L __L__S__P_ _A__P__R__ L__Q__G__R __Q__E__D_ _K__Q__T__      25
                           /\ 3 sj

3361 aacaccaccc ccggcgaccc ccaccacctc cacaaccgcc gtTGActcca cgtccagaca
     _N_ T__T_ _P_ A_ T__ P__T__T__S __T__T__A_ _V__D__S__ T__S__R__Q      233
     Q__T__P__P __P__R__P_ _P__P__P__ P__Q__P__P __L__T__P_ _R__P__D__      45
                                                 _L__H __V__Q__T_      5
                                        E5 orf start ->
                                        NH2 terminus unknown

3421 gctcaccaca tcaaaacagc cacaacaaac cgaaaccaga ggaagaagGT acggacggag
     __L__T__T_ _S__K__Q__ P__Q__Q__T __E__T__R_ _G__R__R__ Y__G__R__R      253
     S__S__P__H __Q__N__S_ _H__N__K__ P__K__P__E __E__E__G_ _T__D__G__      65
     _A__H__H__ I__K__T__A __T__T__N_ _R__N__Q__ R__K__K__V __R__T__E_      25
                                           5 sj /\

3481 gccctccagc aagtcaagga gatcgcaaac gcagcaaagg cgatcaaggt cccgacACCG
     __P__S__S_ _K__S__R__ R__S__Q__T __Q__Q__R_ _R__S__R__ S__R__H__R      273
     G__P__P__A __S__Q__G_ _D__R__K__ R__S__K__G __D__Q__G_ _P__D__T__      85
     _A__L__Q__ Q__V__K__E __I__A__N_ _A__A__K__ A__I__K__V __P__T__P_      45
                                                    E2-bind ->
```


3541 GTCCCGGTct cggtcccggt cgcggtccaa gtcccaaacc cacaccactc ggtccaccac
293
105
65
3601 caggtcccgg tccacGTcgc tcaccaagac tcgggccctt acaagcagat cgcgatccag
313
125
85
5 sj /\
3661 aggaaggtcc ccaaccacct gcagaagggg aggtggaagg tcacccaggc ggcgatcaag
333
145
105
3721 gtcaccctcc acctcctcct cctgcaccac acaacggtca cagcgggcac gagccgaaag
353
165
125
3781 ttcaacaacc agaggggccc gagggtcgag agggtcacga ggagggagcc gtgggggggag
373
185
145
3841 agggcggcga cgaggaaggt catcctcctc ctcctccccc gcccACaaac ggtcacgagg
393
205
165
/\ 3 sj
3901 ggggtctgcT AAgctccgtg gcgtctctcc tggtgaagtg ggagggtcac ttcgatcagt
413
225
$_ 168
<- E5 end
3961 tagttcaaag catacaggac gacttggaag attactggaa gaagctcgcg acccccccagT
433
245
4021 AAtcattgtc aaaggggcgg ctaacacact gaaaaatgtc cgcaacagag ctaAAattaa
453
$_ 245
<- E4 end
/\ 3 sj
4081 atacatggga ctgtttaggt catttagtac tacctggtca tgggtggcag gagatggcac
473
4141 tgagcgtcta ggcaggccca gaatgctcat tagcttttct tcctatactc aaaggagaga
493
4201 ttttgatgaa gcggtgcgat acccgaaagg agttgaTAAg gcctatggca acctggacag
513
L2 orf start ->
4261 tcttTAAcat ttactaatgc tgcttttgct actaacatac taacataccc tagcatttta
514
<- E2 end
4321 tattttttt tacattttgt atttgctATG gcgcgtgcaa aaacggtcaa gcgagactct
11
L2 cds ->

```
4381 gtaactcata tttaccaaac ctgcaaacag gcaggcactt gccccctga tgttattAAT
     _V__T__H__ I__Y__Q__T __C__K__Q_ _A__G__T__ C__P__P__D __V__I__N_      31
                                                          signal ->

4441 AAAgtggaac aacaacagt tgctgacaat atttttaaaat atggcagtgc tggtgtattt
     _K__V__E__ Q__T__T__V __A__D__N_ _I__L__K__ Y__G__S__A __G__V__F_      51

4501 tttggtggcc ttggtattag tacaggccga ggaactgggg gtgctacagg gtacgtgcca
     _F__G__G__ L__G__I__S __T__G__R_ _G__T__G__ G__A__T__G __Y__V__P_      71

4561 cttggggaag gtcctggtgt ccgtgtcgga ggaACCCCCA CGGTTgtaag gccttccttg
     _L__G__E__ G__P__G__V __R__V__G_ _G__T__P__ T__V__V__R __P__S__L_      91
                                          -> E2 bind

4621 gttcctgaaa caatcgggcc cgttgatatt ttgcccattg atacagttaa ccccgtggaa
     _V__P__E__ T__I__G__P __V__D__I_ _L__P__I__ D__T__V__N __P__V__E_     111

4681 cctacagcat catccgtggt ccctgtaact gagtccacag gcgctgattt acttccaggt
     _P__T__A__ S__S__V__V __P__L__T_ _E__S__T__ G__A__D__L __L__P__G_     131

4741 gaagtagaaa caattgctga aatccatcct gtacctgagg ggccatcagt ggatacgcct
     _E__V__E__ T__I__A__E __I__H__P_ _V__P__E__ G__P__S__V __D__T__P_     151

4801 gtagttacca ctagcacagg ttccagtgct gttttagagg ttgccccaga gcctattcct
     _V__V__T__ T__S__T__G __S__S__A_ _V__L__E__ V__A__P__E __P__I__P_     171

4861 ccaacacggg tcaggtttc acgcacacag tatcacaatc catctttca aataataact
     _P__T__R__ V__R__V__S __R__T__Q_ _Y__H__N__ P__S__F__Q __I__I__T_     191

4921 gagtctactc cagcacaagg ggatcgtct cttgcagatc acgttttggt gacatcggt
     _E__S__T__ P__A__Q__G __E__S__S_ _L__A__D__ H__V__L__V __T__S__G_     211

4981 tctgggggggc aacgaatagg gggtgatata actgacataa ttgagttaga ggaaattcct
     _S__G__G__ Q__R__I__G __G__D__I_ _T__D__I__ I__E__L__E __E__I__P_     231

5041 agtaggtata catttgaaat tgaagaacca actcctccac gccgcagcag tactccattg
     _S__R__Y__ T__F__E__I __E__E__P_ _T__P__P__ R__R__S__S __T__P__L_     251

5101 ccacgcaatc aatctgtagg ccgtaggagg ggtttctctt tgactaatag acgtttagta
     _P__R__N__ Q__S__V__G __R__R__R_ _G__F__S__ L__T__N__R __R__L__V_     271

5161 cagcaggtac aagtggacaa tccattgttt ctaactcaAC CATCTAAGTT agttcgtttt
     _Q__Q__V__ Q__V__D__N __P__L__F_ _L__T__Q__ P__S__K__L __V__R__F_     291
                                             -> E2 bind

5221 gcatttgata atcctgtttt tgaggaagaa gtgactaata tatttgaaaa tgatctggat
     _A__F__D__ N__P__V__F __E__E__E_ _V__T__N__ I__F__E__N __D__L__D_     311

5281 gtctttgaag aacctccaga cagagatttt cttgatgtta gggaattggg acgtccacaa
     _V__F__E__ E__P__P__D __R__D__F_ _L__D__V__ R__E__L__G __R__P__Q_     331

5341 tattctacaa caccagcggg atatgttaga gtaagcaggt tggggactcg agccactatt
     _Y__S__T__ T__P__A__G __Y__V__R_ _V__S__R__ L__G__T__R __A__T__I_     351

5401 cgcactcgct ctggtgcaca gatagggtcg caagtccatt tttacagaga tcttagctct
     _R__T__R__ S__G__A__Q __I__G__S_ _Q__V__H__ F__Y__R__D __L__S__S_     371

5461 attaatactg agatcctat tgaattacaa ttattaggcc aacattcagg tgatgctact
     _I__N__T__ E__D__P__I __E__L__Q_ _L__L__G__ Q__H__S__G __D__A__T_     391

5521 atagtccacg gacctgttga aagcacattt atagatatgg atatttctga aaatccatta
     _I__V__H__ G__P__V__E __S__T__F_ _I__D__M__ D__I__S__E __N__P__L_     411

5581 tctgaaagca ttgaagcata ttcacatgat ttatttatag atgaaacggt ggaagatttc
     _S__E__S__ I__E__A__Y __S__H__D_ _L__L__L__ D__E__T__V __E__D__F_     431

5641 agtgggtctc agctggttat aggtaatcga aggagcacaa actcttacac tgttcctagg
     _S__G__S__ Q__L__V__I __G__N__R_ _R__S__T__ N__S__Y__T __V__P__R_     451

5701 tttgaaacta caagaaatgg ttcatagtat acacaagaca caaagggata ttatgttgca
     _F__E__T__ T__R__N__G __S__Y__Y_ _T__Q__D__ T__K__G__Y __Y__V__A_     471
```

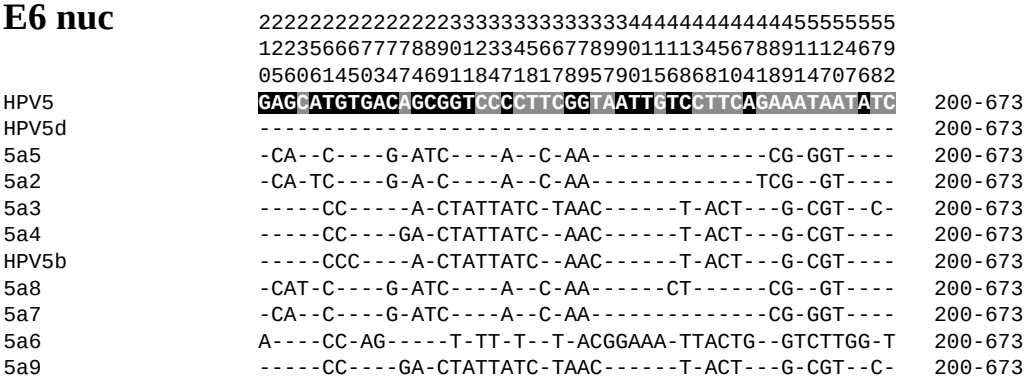

5761 tatccagagt cacgtaataa tgcagaaatc atttatccta cacctgatat tcctgtagtc
_Y__P__E__ S__R__N__N __A__E__I_ _I__Y__P__ T__P__D__I __P__V__V_ 491
5821 attatacacc ctcatgacag tacaggggac ttttatttac atcccagtct tcacaggcgc
_I__I__H__ P H__D S T__G__D_ _F__Y__L__ H__P__S__L H R__R_ 511
5881 aaacgtaaaa gaaaatattt gTGAtttgca ttcgAGATGg cagtgtggca ctcggctaat
_K__R__K__ R__K__Y__L __$_ _M__ A__V__W__H __S__A__N_ 518/8
/\ 3 sj
L1 orf start -> L1 cds ->
<- L2 end
5941 ggtaaagtat atctccacc atcgacaccg gtggccagag tccaaagcac cgatgaatac
_G__K__V__ Y__L__P__P __S__T__P_ _V__A__R__ V__Q__S__T __D__E__Y_ 28
6001 attcaaagaa caaatatcta ctatcatgca tttagtgaca gattgttaac tgtaggtcat
_I__Q__R__ T__N__I__Y __Y__H__A_ _F__S__D__ R__L__L__T __V__G__H_ 48
6061 ccttatttca atgtatacaa tattaatggt gataagcttg aggtcctaa ggtttcagga
_P__Y__F__ N__V__Y__N __I N G_ _D__K__L__ E__V__P__K __V__S__G_ 68
6121 aatcaacaca gagtatttcg cctaaaatta ccagatccta acagatttgc attacctgat
_N__Q__H__ R__V__F__R __L__K__L_ _P__D__P__ N__R__F__A __L P D_ 88
6181 atgtctgttt acaacctga caaagaacgt ttggtttggg ctgtagagg cttagaaata
_M__S__V__ Y__N__P__D __K__E__R_ _L__V__W__ A__C__R__G __L__E__I_ 108
6241 ggtaggggcc agccattagg tgtacggagt actggtcacc cttatttcAA TAAAgtaaaa
_G__R__G__ Q__P__L__G __V R S_ _T__G__H__ P__Y__F__N __K__V__K_ 128
signal ->
6301 gatacagaaa acagtaatgc atacataaca ttttctaaag atgagagaca ggatacatct
_D__T__E__ N__S__N A Y__I__T_ _F__S K D__D__R Q D T__S_ 148
▲
6361 tttgatccta aacagatcca aatgttgatt gtaggatgca ccttgcat aggagagcat
_F__D__P__ K__Q__I__Q __M__F__I_ _V__G__C__ T__P__C__I __G__E__H_ 168
6421 tgggatgaaag ctgttccatg tgcagaaaaat gatcagcaaa ctggcctttg tcctccgatt
_W__D__K__ A V P__C __A E N _D Q__Q T G__L__C __P__P__I_ 188
6481 gaagtaaaaa acacatatat acaagatggt gatatggcag acatggttt tgggaagatg
_E__L__K__ N__T__Y__I Q D__G_ _D__M__A__ D__I__G__F __G__N M_ 208
6541 aatttaaagg cactgcagga tagtagatca gatgtcagtt tagacatcgt caatgaaact
_N__F__K__ A__L__Q__D __S__R__S_ _D__V__S__ L__D__I__V __N__E__T_ 228
6601 tgcaagtatc cagatttttt aaagatgcaa aacgatattt atggcgatgc gtgctttttt
_C__K__Y__ P__D__F__L __K__M__Q_ _N__D__I__ Y__G__D__A __C__F__F_ 248
6661 tatgctcgta gggagcaatg ttatgccaga cactttttg ttagaggggg aaaaactggt
_Y__A__R__ R__E__Q__C __Y__A__R_ _H__F__F__ V__R__G__G __K__T__G_ 268
6721 gatgacattc cacgtgcaca aattgacaat ggtacataca aaaatcagtt ttacattcca
_D__D__I__ P R A__Q __I__D__N_ _G__T__Y__ K__N__Q__F __Y__I__P_ 288
6781 gggctgatg gccaagctca aaagactata ggaaattcca tgtatttccc aactgttagt
_G__A__D__ G__Q__A__Q __K__T__I_ _G__N S M__Y__F__P __T__V__S_ 308
6841 ggctcattag tatccagtga tgctcaattg tttaacaggc ccttctggct ccaaagagcc
_G__S__L__ V__S__S__D __A__Q__L_ _F__N__R__ P__F__W__L __Q__R__A_ 328
6901 caaggtcata ataatggcat cctgtgggct aatcaaatgt ttatcacagt ggttgacaac
_Q__G__H__ N__N__G__I __L__W__A_ _N__Q__M__ F__I__T__V __V__D__N_ 348
6961 acaagaaata ctaatttcag tatttctgta tataatcagg ctggagcact aaaagatgtt
_T__R__N__ T__N__F__S __I__S__V_ _Y__N__Q__ A__G A L __K__D__V_ 368
7021 gcagactata atgcagatca atttagagaa tatcaaagac atgtagaaga atatgaaata
_A__D__Y__ N__A D Q __F__R__E_ _Y__Q__R__ H__V__E__E __Y__E__I_ 388

```
7081 tctttaattc tacaactctg taaggttcct ttaaaggcac aggtattggc acagatcaat
     _S__L__I__ L__Q__L__C __K__V__P_ _L__K__A__ Q__V__L__A __Q__I__N_      408

7141 gcaatgaact cttcgttatt ggaggattgg cagttaggat ttgttcccac tcctgataat
     _A__M__N__ S__S__L__L __E__D__W_ _Q__L__G__ F__V__P__T __P__D__N_      428

7201 ccaattcagg acacctacag atatattgac tctttggcta cacggtgtcc agataagaat
     _P__I__Q__ D__T__Y__R __Y__I__D_ _S__L__A__ T__R__C__P __D__K__N_      448

7261 cctccgaaag aaaaggaaga cccttataag ggcttacatt tttgggatgt agatttaact
     _P__P__K__ E__K__E__D __P__Y__K_ _G__L__H__ F__W__D__V __D__L__T_      468

7321 gaaagattgt cattagattt agatcaatat tccttaggca gaaaattttt attccaagct
     _E__R__L__ S__L__D__L __D__Q__Y_ _S__L__G__ R__K__F__L __F__Q__A_      488

7381 ggttacaac aaacgACCGT TAACGGtaca aaagcagtgt cttataaagg gtctaataga
     _G__L__Q__ _Q__T__T__V __N__G__T_ _K__A__V__ S__Y__K__G __S__N__R_      508
                      -> E2 bind

7441 ggaacaaaac gcaaacgtaa aaatTGAggt ctgaccgaaa gtggtacatt ttataaaact
     _G__T__K__ R__K__R__K __N__$_                                         516
                               <- L1 end

7501 tttacacagt attcaaggaa tgtttgttta ctctgactaa gtataagtct tccaaggata

7561 ccgACCGCAC CCGGTacact cagtcaagtt gttgccaata tagaatcaga tcagtgccaa
        ->  E2-bind

7621 acacaccgtc ttggactcag aacagaccgt gttcgttata acatgctcgg attagggacc

7681 tccccaaaga agattgtaatc taCAATCGCT TTTGGCAATC GCATTTGGCA ctgctaaaag
                              ->    overlapping repeat    <-

7741 ACCGTT
     -> E2-bind
```

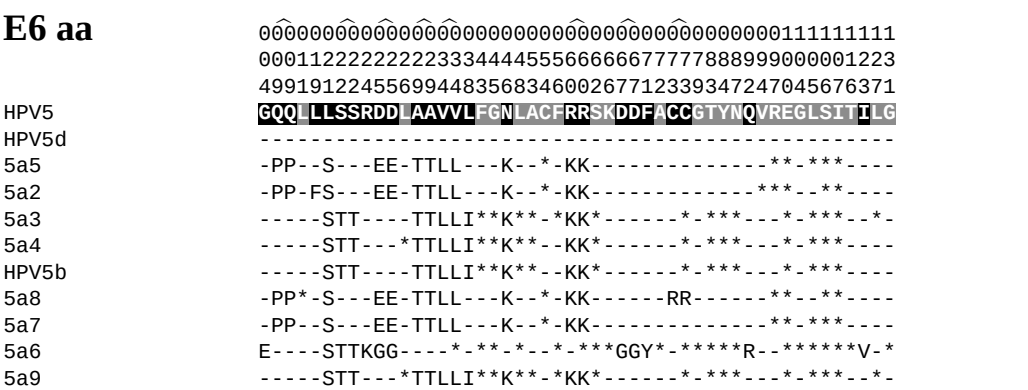
noncoding

```
                      insertion
                    ___________       i
                    0000000000000000000001111111111111
                    1111122222223444445822222367778 88
                    1467955555558017894302479393670 47
HPV5                TATTC.......TGGCACAT.TATCAGAATGCC   1-199
HPV5d               -----.....--------.------------   1-199
5a5                 ---A-.....--------TC-----G-----   1-199
5a2                 ----T.......-AA---G-TC-----G-----   1-199
5a3                 A-G--.....---TG-G-TC---------A-   1-199
5a4                 A-G--.......-AATG-GCTC---------A-   1-199
HPV5b               A-G--TCCTTGT---TG-GCTC---G-----A-   1-199
5a8                 -----.......-A----G-TC--T--G-----   1-199
5a7                 ---A-.......-AA--TG-TC-----G-----   1-199
5a6                 AT--T.......C-----G-TC-A--A-GCA-T   1-199
5a9                 A-G--.....---TG-GCTCG--------A-   1-199
```

E6 nuc



E6 aa


0ˆ0ˆ0ˆ

E7 nuc

```
                77788889999
                47900340133
                50638909758
HPV5            TGAGCGACACT   663-974
HPV5d           -----------   663-974
5a5             ----T------   663-974
5a2             ----------C   663-974
5a3             --T---CT-T-   663-974
5a4             --T---CT-T-   663-974
HPV5b           --T--ACT-T-   663-974
5a6             AATA--C-G--   663-974
```

E7 aa

```
                23444568899
                86579903512
HPV5            FENRSRNLGCP
HPV5d           -----------
5a5             ----F------
5a2             ----------*
5a3             --I---H*-*-
5a4             --I---H*-*-
HPV5b           --I--*H*-*-
5a6             Y*I*--H-*--
```

E1 nuc

```
                01111111111111111111111122222222222222222222222222
                90122344444455677799999900000111122233333445555557
                76816102225717047923488894557901234591244719355778 6
                65293172565089236775803929010405167683206690513362 1
HPV5            TAGGCCCCTGTGCCAACGCCGTCCGATAGTTTCGCTAGATAGACTATCCAT   961-2781
HPV5d           ------------------G---------------------------GG--   961-2781
HPV5b           CGAATTTTATCATTGTAAGTCATTATGCACCGTTTATAGCGAGTCGCGGG-   961-2781
5a5                                                             -   961-974,2723-2781
5a2                                                             -   961-974,2723-2781
5a4                                                             -   961-974,2723-2781
5a3                                                             A   961-974,2723-2781
5a6                                                             -   961-974,2723-2781
```

E1 aa

```
                00001111111112222233333333333333333444444444555555556
                03780145556780167722344446667788892345566781233334 0
                65471794565067412935001343440805729261402270511891 1
HPV5            SLLVPLDYTAAKFLLIAKHHAAPDAAYYKSSTIGCISLRNKEIAITTTTKS
HPV5d           ------------------D---------------------------SR--
HPV5b           P**I*****S********D*P*****AA**************V**AASRR-
5a5                                                             -
5a2                                                             -
5a4                                                             -
5a3                                                             T
5a6                                                             -
```

E2 nuc

```
                 22222222222222333333333333333333333333333344444444
                 78888999999999001122233333556666667778899900000000
                 60348233577888162655946777692226773784417824555558
                 10004958114047857604872145500341565222339048156780
HPV5             TGACTTACCGATGCTCCGGAAGACGCGCCACACCCATGGGACCAGAATGA   2723-4267
HPV5d            ---G-----------G----------------------------------   2723-4267
5a5              ---G------G----                        TA--ATGCT-   2723-3022,3968-4267
5a2              ---G------G-T--                        -----TGCTC   2723-3022,3968-4267
5a3              A-GG-------C--C                        ----ATGCT-   2723-3022,3968-4267
5a4              --GG-------C--C                        ----ATGCT-   2723-3022,3968-4267
HPV5b            --GG-------C--CGTAACTAGGAACTGGGGGTAGCTT-----AT-CT-   2723-4267
5a6              -AGGCCGTTA-C-GC                       C--AT-TGCT-   2723-3022,3854-4267
```

E2 aa

```
                 00000000000000011111122222222333333333333444444444444
                 12345677788888913477901111890001113557791234444445
                 36604912734689955868294788100113888044479042355563
HPV5             DLLPLAEALQTDAHILTKAREGNPAARRLTTRTTSRSGGKGRITLNNNVK
HPV5d            ---A-----------V----------------------------------
5a5              ---A------*----                        **--*CCCF-
5a2              ---A------*-S--                        -----CCCFT
5a3              E-*A-------*--T                        ----*CCCF-
5a4              --*A-------*--T                        ----*CCCF-
HPV5b            --*A-------*--TV****DRDAKK*WVGG*SSY*PLL-----*YYYF-
5a6              -**A******-*-DT                       N--**-CCCF-
```

E2 nuc cont'd.

```
                 444444444444444
                 011111112222222
                 801488991122445
                 681638253658125
HPV5             TTTGCCAGGGCAGCG   2723-4267
HPV5d            ---------------   2723-4267
5a5              ---TA-G--A-----   2723-3022,3968-4267
5a2              AC--A----AT----   2723-3022,3968-4267
5a3              C----G-AA------   2723-3022,3968-4267
5a4              C-A----AA------   2723-3022,3968-4267
HPV5b            C----G-AA-AG--T   2723-4267
5a6              ------G--C--CGT   2723-3022,3854-4267
```

E2 aa cont'd.

```
                 444444444455555
                 566788999900001
                 523579017812771
HPV5             MSTRSTQRAVPKAAL
HPV5d            ---------------
5a5              ---L*-*--*-----
5a2              K*--*----**----
5a3              T----S-**------
5a4              T-*----**------
HPV5b            T----S-**-**--*
5a6              ------*--*--RR*
```

E4 nuc

```
                3333333333333333333333
                2333335566666677788999
                9467776922267737844178
                8721455003415652223390
HPV5            AGACGCGCCACACCCATGGGAC   3285-4022
HPV5d           -----------------------  3285-4022
5a5                           TA        3968-4022
5a2                           --        3968-4022
5a3                           --        3968-4022
5a4                           --        3968-4022
HPV5b           TAGGAACTGGGGGTAGCTT---   3285-4022
5a6                          C--        3854-4022
```

E4 aa

```
                00000001111111̂111111222̂
                0222339011123356688133
                5169014223461113667022
HPV5            KEQPPRGLRSPEPPPEVEGSDD
HPV5d           ----------------------
5a5                           VV
5a2                           --
5a3                           --
5a4                           --
HPV5b           I*****A***AGVVTG*DC---
5a6                          T--
```

E5 nuc

```
                3333333333333
                5566666677788
                6922267737844
                5003415652223
HPV5            GCCACACCCATGG   3406-3912
HPV5d           -------------   3406-3912
HPV5b           CTGGGGGTAGCTT   3406-3912
```

E5 aa

```
                0000̂000011111̂
                5677789912244
                4223360103666
HPV5            VSAHHRNHLSFRR
HPV5d           -------------
HPV5b           LLGRRGKY*GSII
```

noncoding

```
                  dd
                44444
                23333
                81233
                65301
HPV5            TATTT   4268-4347
HPV5d           -----   4268-4347
HPV5b           CGC..   4268-4347
```

L2 nuc

```
                4444444444444444444444444444444444444445555555555
                3344445666666777777777788888889999999990011112233
                5615773333466001457788990346789001446791623575822
                9462464235168150926928173689540585475724974978338 9
HPV5            ACCATAAAACCGTCCTAATGGAGCAAAGGCGTTAAGTGAGAATTACGCGG   4348-5904
HPV5d           -G-------------------------------------------------   4348-5904
5a5                   T--T----T----A--AT----AAA----------           4500-5030,5770-5904
HPV5b           TGTGCC-GGT---TT--G-A-GA-G-C-A-TAC---GAGATGCCGTAAAA   4348-5904
5a6                   -TGTAACTTAG-GAT-A--GTT----CGGA                4531-4959,5854-5904
```

L2 aa

```
                0000ô0000ô001î111111111111111111111111222222222223333
                0023446999900122334444455667778889900111245677012 2
                4635333566877801453457802374691670906056409407227 8
HPV5            ATGQLLTIIIPVVVLTETPEGSVTVLPRRRQSFIESVSQRIERSLDVVLG
HPV5d           -R-------------------------------------------------
5a5                   S--*----*----*--**----***----------
HPV5b           *R****-*VV---**--*-*-**-*-*-*-H**---***Q********R
5a6                   -*VV*II****-***-*--***----*V**
```

L2 nuc cont'd.

```
                55555555555555555555555
                33334445556666667777888888
                67781270281224452366012347
                70982128937560957306879003
HPV5            TAGTTGAACTTAACTGCAAATACCGA   4348-5904
HPV5d           --------G----------------   4348-5904
5a5                             -G--AG     4500-5030,5770-5904
5a3                             -GTAAG     5770-5904
HPV5b           CGACGAGTGACGGTCTTCTGCG-AAG   4348-5904
5a6                               G         4531-4959,5854-5904
```

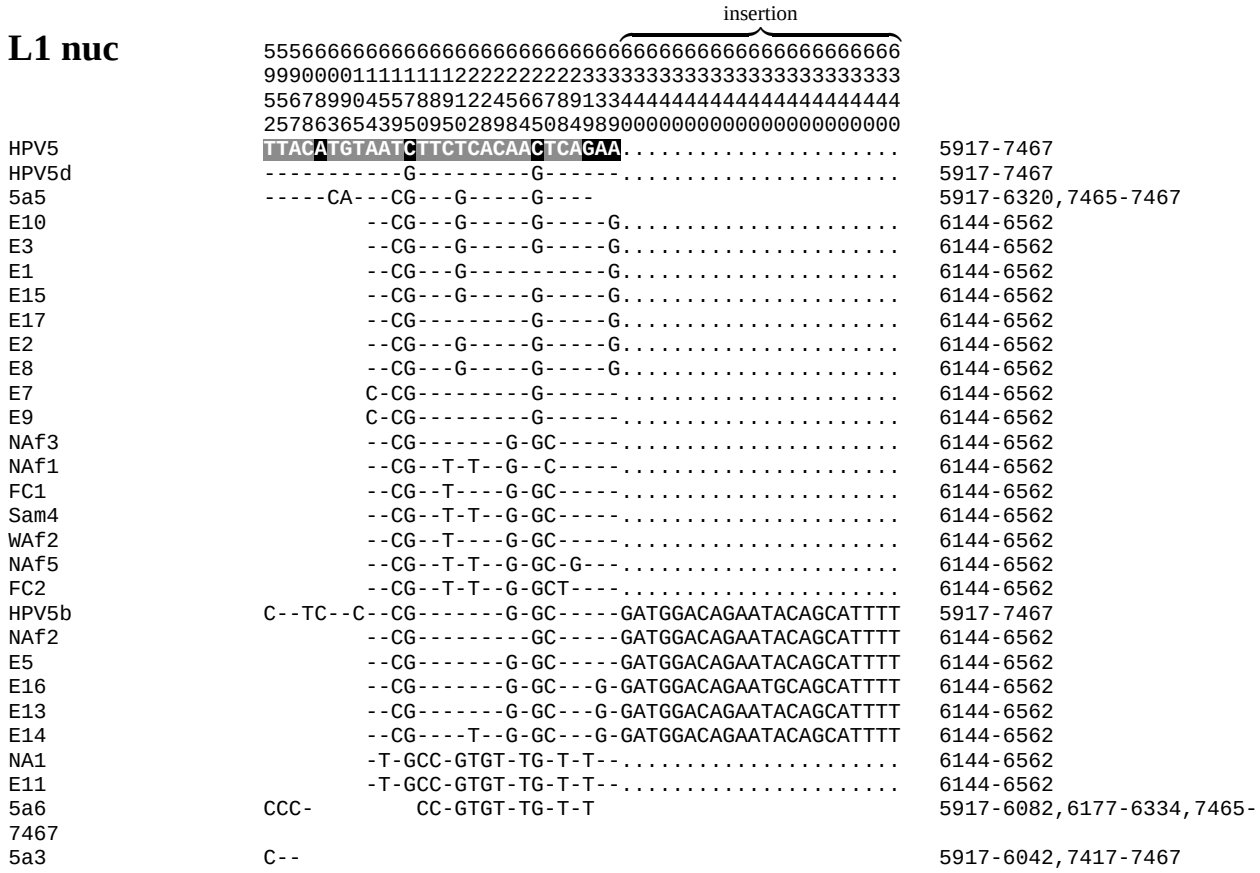
L2 aa cont'd.

```
                33333333344444444444444445
                44445578912223336677899990
                01475857424671460213704589
HPV5            VRRTSQESHSLETFSLYTAPDVHPSH
HPV5d           --------Q----------------
5a5                             -*--NR
5a3                             -**TNR
HPV5b           ********Q***A*********-TNR
5a6                               R
```

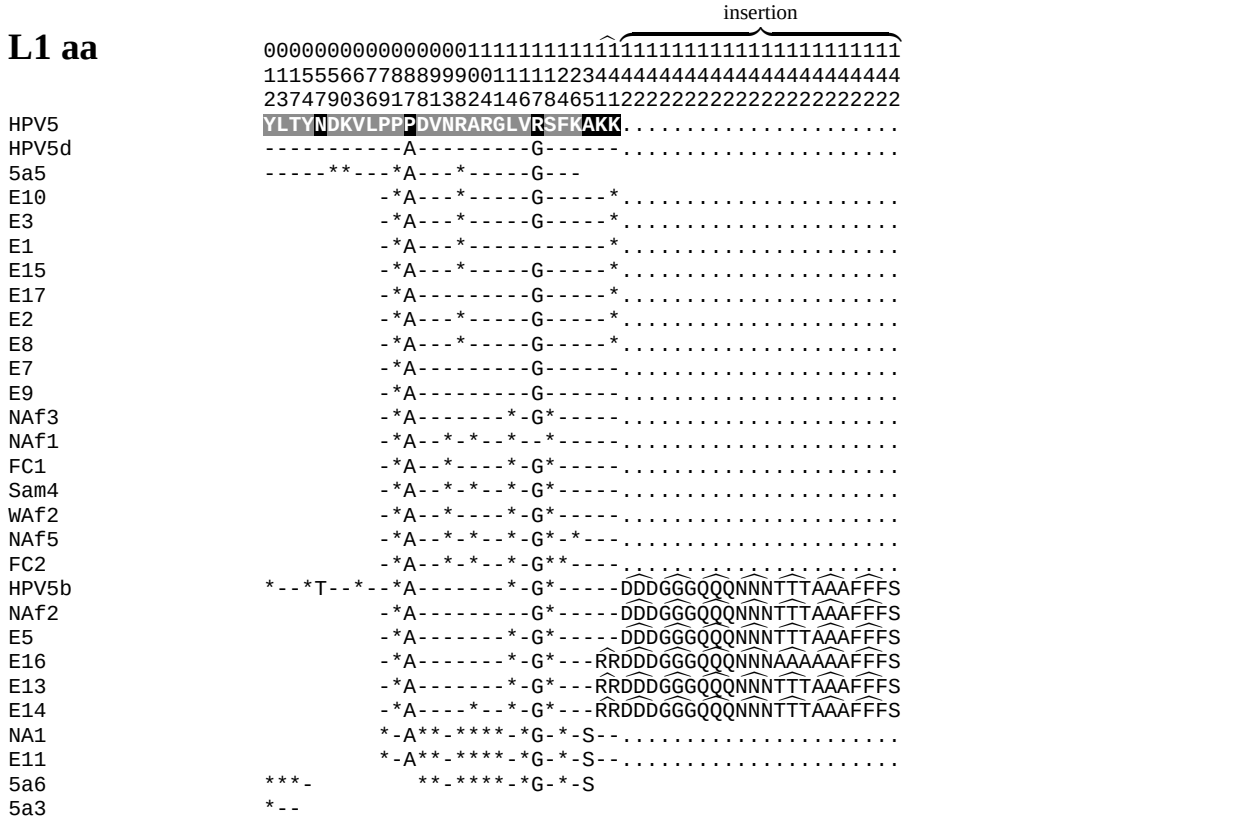
noncoding

```
                55
                99
                11
                34
HPV5            CG   5905-5916
HPV5d           --   5905-5916
5a5             --   5905-5916
5a3             GC   5905-5916
HPV5b           GC   5905-5916
5a6             GC   5905-5916
```

L1 nuc


insertion
HPV5 TTACATGTAATCTTCTCACAACTCAGAA..... 5917-7467
HPV5d ----------G---------G------..... 5917-7467
5a5 -----CA---CG---G-----G---- 5917-6320,7465-7467
E10 --CG---G-----G-----G..... 6144-6562
E3 --CG---G-----G-----G..... 6144-6562
E1 --CG---G----------G..... 6144-6562
E15 --CG---G-----G-----G..... 6144-6562
E17 --CG---------G-----G..... 6144-6562
E2 --CG---G-----G-----G..... 6144-6562
E8 --CG---G-----G-----G..... 6144-6562
E7 C-CG---------G------..... 6144-6562
E9 C-CG---------G------..... 6144-6562
NAf3 --CG-------G-GC-----..... 6144-6562
NAf1 --CG--T-T--G--C-----..... 6144-6562
FC1 --CG--T----G-GC-----..... 6144-6562
Sam4 --CG--T-T--G-GC-----..... 6144-6562
WAf2 --CG--T----G-GC-----..... 6144-6562
NAf5 --CG--T-T--G-GC-G---..... 6144-6562
FC2 --CG--T-T--G-GCT----..... 6144-6562
HPV5b C--TC--C--CG-------G-GC-----GATGGACAGAATACAGCATTTT 5917-7467
NAf2 --CG---------GC-----GATGGACAGAATACAGCATTTT 6144-6562
E5 --CG-------G-GC-----GATGGACAGAATACAGCATTTT 6144-6562
E16 --CG-------G-GC---G-GATGGACAGAATGCAGCATTTT 6144-6562
E13 --CG-------G-GC---G-GATGGACAGAATACAGCATTTT 6144-6562
E14 --CG----T--G-GC---G-GATGGACAGAATACAGCATTTT 6144-6562
NA1 -T-GCC-GTGT-TG-T-T--..... 6144-6562
E11 -T-GCC-GTGT-TG-T-T--..... 6144-6562
5a6 CCC- CC-GTGT-TG-T-T 5917-6082,6177-6334,7465-7467
5a3 C-- 5917-6042,7417-7467

L1 aa


insertion
HPV5 YLTYNDKVLPPPDVNRARGLVRSFKAKK.....
HPV5d ----------A---------G------.....
5a5 -----**---*A---*-----G---
E10 -*A---*-----G-----*.....
E3 -*A---*-----G-----*.....
E1 -*A---*----------*.....
E15 -*A---*-----G-----*.....
E17 -*A---------G-----*.....
E2 -*A---*-----G-----*.....
E8 -*A---*-----G-----*.....
E7 -*A---------G------.....
E9 -*A---------G------.....
NAf3 -*A-------*-G*-----.....
NAf1 -*A--*-*--*--*-----.....
FC1 -*A--*----*-G*-----.....
Sam4 -*A--*-*--*-G*-----.....
WAf2 -*A--*----*-G*-----.....
NAf5 -*A--*-*--*-G*-*---.....
FC2 -*A--*-*--*-G**----.....
HPV5b *--*T--*--*A-------*-G*-----DDDGGGQQQNNNTTTAAAFFFS
NAf2 -*A---------G*-----DDDGGGQQQNNNTTTAAAFFFS
E5 -*A-------*-G*-----DDDGGGQQQNNNTTTAAAFFFS
E16 -*A-------*-G*---RRDDDGGGQQQNNNAAAAAAFFFS
E13 -*A-------*-G*---RRDDDGGGQQQNNNTTTAAAFFFS
E14 -*A----*--*-G*---RRDDDGGGQQQNNNTTTAAAFFFS
NA1 -*A**-****-*G-*-S--.....
E11 -*A**-****-*G-*-S--.....
5a6 ***- **-****-*G-*-S
5a3 *--

L1 nuc cont'd.

```
                                 insertion
                                 ⌒⌒⌒
                                 66666666666666666666666666666666666666666666666666
                                 33333333333333444444444444444444445555555555555556666
                                 44444445567780112333444455566678902233344558880335
                                 00000580205872146358578913601274222547839582586391
HPV5                             .....CAAGTGCTAAATGTAGAAAGTGACTTCCCCAGCATGTAACCGCTG   5917-7467
HPV5d                            .....----------------------------------------------   5917-7467
E10                              .....----------G---------A-------G---------          6144-6562
E3                               .....----------G---------A-------G--A------          6144-6562
E1                               .....---A------G----------A----T-G--A------          6144-6562
E15                              .....----------G----------A----T-G--A------          6144-6562
E17                              .....---A------G----------A----T-G--A------          6144-6562
E2                               .....---------TG---------A-------G--A------          6144-6562
E8                               .....---A------G----------A----T-G--A------          6144-6562
E7                               .....----------G-----C-----------GT-AT-C---          6144-6562
E9                               .....----C-----G-----C-----------GT-AT-C---          6144-6562
NAf3                             .....---ACA-C-----GT-----C--T--T-GT-AT-C---          6144-6562
NAf1                             .....----CA-C--G-CGT-----C-----TTGT-AT-C---          6144-6562
FC1                              .....----CA-C--G--GT-----C-----TTGTCAT-C---          6144-6562
Sam4                             .....----CA-C--G--GT-----C-----TTGTCAT-C---          6144-6562
WAf2                             .....----CA-C--G--GT-----C-----TTGTCATTC---          6144-6562
NAf5                             .....----CA-C--G--GT-----C-----TTGTCAT-C---          6144-6562
FC2                              .....----CA-C--G--GT-----C-----TTGTCAT-CA--          6144-6562
5a4                                                                      A-CC   6600-6942,7399-7467
HPV5b                            CTAAA--TACA-C-----GTA----C-----TTGT-AT-C---GTTATCC   5917-7467
NAf2                             CTAAAT-TACA-C-----GTA----C-----TTGT-AT-C---          6144-6562
E5                               CTAAA--TACA-C-----GTA----C-C---TTGT-AT-C---          6144-6562
E16                              CTAAA--TACA-C-----GTA----C-C---TTGT-AT-C---          6144-6562
E13                              CTAAA--TACA-C-----GTA-G--C-----TTGT-AT-C---          6144-6562
E14                              CTAAA-GTACA-C-----GTA-G--C-----TTGT-AT-C---          6144-6562
NA1                              .....----CA-CT--CA---C--AC-GTACTTGT-AT-CAAG          6144-6562
E11                              .....----CAACT--CA---C-CAC-TTACTTGT-AT-CAAG          6144-6562
```

L1 aa cont'd.

```
                                 insertion
                                 ⌒⌒⌒
                                 11111111111111111111111111111111111122222222222222222
                                 44444444444555666777777777788888999000000111222334 4
                                 22222345683472560334778899022270262367891342340915
HPV5                             .....DRQDSQIFTIGDVVPEENNDDQTTTPLNQDIGNMNKLQLDIKNIA
HPV5d                            .....----------------------------------------------
E10                              .....----------*--------EE-------E---------
E3                               .....----------*--------EE-------E--*------
E1                               .....---N------*----------*----*-E--*------
E15                              .....----------*----------*----*-E--*------
E17                              .....---N------*----------*----*-E--*------
E2                               .....---------**--------EE-------E--*------
E8                               .....---N------*----------*----*-E--*------
E7                               .....----------*----DD-----------E*-**-*---
E9                               .....----*-----*----DD-----------E*-**-*---
NAf3                             .....---N**-*-----**-----*-III-*-E*-**-*---
NAf1                             .....----**-*--*-LL*-----*-----**E*-**-*---
FC1                              .....----**-*--*--**-----*-----**E****-*---
Sam4                             .....----**-*--*--**-----*-----**E****-*---
WAf2                             .....----**-*--*--**-----*-----**E****L*---
NAf5                             .....----**-*--*--**-----*-----**E****-*---
FC2                              .....----**-*--*--**-----*-----**E****-**--
5a4                                                                      *-**
HPV5b                            SSKKK--LN**-*-----**KK---*-----**E*-**-*---*******
NAf2                             SSKKK*-LN**-*-----**KK---*-----**E*-**-*---
E5                               SSKKK--LN**-*-----**KK---*-PPP-**E*-**-*---
E16                              SSKKK--LN**-*-----**KK---*-PPP-**E*-**-*---
E13                              SSKKK--LN**-*-----**KKDD-*-----**E*-**-*---
E14                              SSKKK-*LN**-*-----**KKDD-*-----**E*-**-*---
NA1                              .....----**-**--*II-DD--NN-VVV***E*-**-****
E11                              .....----*****--*II-DDTTNN-LLL***E*-**-****
```

L1 nuc cont'd.

```
                6666666666777777777777777777
                6777778889000011111222233344
                8134681159039912556125696890 4
                7137830729680890254217602327 9
HPV5            CACCGGATAGGTCCACTGGCAGGGAGATA   5917-7467
HPV5d           --G-----------G-------------   5917-7467
5a3                                    G       5917-6042,7417-7467
5a4             TTGTAA-GT                CG    6600-6942,7399-7467
HPV5b           TTG-AACGTACGTTGGCTATGAAAGTC-G   5917-7467
```

L1 aa cont'd.

```
                2222222333333344444444444444 5
                5677889016679900111334558899 1
                7537498121442412236257082927 1
HPV5            AGRDQGISVQADLLAQSSEDRKPKRGQGK
HPV5d           --G-----------E-------------
5a3                                    *
5a4             **G***-A*                **
HPV5b           **G-***A**PE***E**********H-*
```

LCR

```
                777777777777777777
                445556666666666777
                792580145667889223
                027275308090035159
HPV5            TTTCAAAGAAGGCCTGTA   7468-7746
HPV5d           ------------GG----   7468-7746
5a5             ------------GG----   7468-7746
5a2             ------------GG---C   7468-7746
5a3             C---G-G-----TG---G   7468-7746
5a4             C---G-G--G-CTG---G   7468-7746
HPV5b           C---G-G-----TG-A-G   7468-7746
5a8                             C    7736-7746
5a7                             C    7736-7746
5a6             -CGA-G-AG-AATGC-CG   7468-7746
5a9                             G    7737-7746
```

```
LOCUS       HPV6b        7902 bp ds-DNA   circular  VRL       11-MAR-1994
DEFINITION  Human papillomavirus type 6b (HPV-6b), complete genome.
ACCESSION   X00203
KEYWORDS    complete genome; overlapping genes.
SOURCE      Human papilloma virus type 6b DNA.
  ORGANISM  Human papillomavirus type 6b
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
COMMENT     Naturally occurring variants of HPV6b have been observed throughout
            the genome, including those observed in the complete genome of HPV6a
            (Hofman et al, 1995).
            Reference: Farr,A., Wang,H., Kasher,M.S., and Roman,A.,
                J. Infect. Dis. 171, 697-700 (1995)
                Variant Accession Nos. L22693, L22694
            Reference: Heinzel,P.A., Chan,S.Y., Ho,L., O'Connor,M., et al.
                J. Clin. Microbiol. 33, 1746-54 (1995)
                Variants obtained from authors.
            Reference: Hofmann,K.J., Cook,J.C., Joyce,J.G., Brown,D.R., Schultz,L.D.,
                George,H.A., Rosolowsky,M., Fife,K.H., and Jansen,K.U.
                Virology 209, 506-18 (1995)
                Variant Accession No. L41216
            Reference: Icenogle,J.P., Sathya,P., Miller,D.L., Tucker,R.A., and Rawls,W.E.
                Virology 184, 101-7 (1991)
                Sequences constructed by editing reference sequence according to figures
                in article.
            Reference: Kasher,M.S., and Roman,A., Virology 165, 225-33 (1988)
                Variant Accession Nos. M22106-M22108, M20561
            Reference: Roman,A., and Brown,D., J. Infect. Dis. 171, 697-700 (1995)
                Variant Accession Nos. M36837-M36842
            FEATURES             Location/Qualifiers
     CDS             join(7746..7902,1..5)
                     /note="probably not functional"
                     /note="E8 from bp 7611 to  5"
                     /gene="E8"
                     /note="putative"
                     /codon_start=1
                     /translation="MHCSQLLKAFLACSSTFFCSYCLVYNNIKMSNLRSHTCDRFRLS
                     TPYIFPSYS"
     CDS             102..554
                     /note="ORF E6 from bp 30 to 554"
                     /product="transforming protein"
                     /gene="E6"
                     /note="putative"
                     /codon_start=1
                     /translation="MESANASTSATTIDQLCKTFNLSMHTLQINCVFCKNALTTAEIY
                     SYAYKHLKVLFRGGYPYAACACCLEFHGKINQYRHFDYAGYATTVEEETKQDILDVLI
                     RCYLCHKPLCEVEKVKHILTKARFIKLNCTWKGRCLHCWTTCMEDMLP"
     CDS             530..826
                     /note="ORF E7 from 440 to 826"
                     /product="transforming protein"
                     /gene="E7"
                     /note="putative"
                     /codon_start=1
                     /translation="MHGRHVTLKDIVLDLQPPDPVGLHCYEQLVDSSEDEVDEVDGQD
                     SQPLKQHFQIVTCCCGCDSNVRLVVQCTETDIREVQQLLLGTLNIVCPICAPKT"
     CDS             832..2781
                     /note="ORF E1 from bp 715 to 2781"
                     /product="replication protein"
                     /gene="E1"
                     /note="putative"
                     /codon_start=1
                     /translation="MADDSGTENEGSGCTGWFMVEAIVQHPTGTQISDDEDEEVEDSG
                     YDMVDFIDDSNITHNSLEAQALFNRQEADTHYATVQDLKRKYLGSPYVSPINTIAEAV
                     ESEISPRLDAIKLTRQPKKVKRRLFQTRELTDSGYGYSEVEAGTGTQVEKHGVPENGG
```

```
                DGQEKDTGRDIEGEEHTEAEAPTNSVREHAGTAGILELLKCKDLRAALLGKFKECFGL
                SFIDLIRPFKSDKTTCLDWVVAGFGIHHSISEAFQKLIEPLSLYAHIQWLTNAWGMVL
                LVLLRFKVNKSRSTVARTLATLLNIPENQMLIEPPKIQSGVAALYWFRTGISNASTVI
                GEAPEWITRQTVIEHGLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
                FLNSNMQAKYVKDCATMCRHYKHAEMRKMSIKQWIKHRGSKIEGTGNWKPIVQFLRHQ
                NIEFIPFLTKFKLWLHGTPKKNCIAIVGPPDTGKSYFCMSLISFLGGTVISHVNSSSH
                FWLQPLVDAKVALLDDATQPCWIYMDTYMRNLLDGNPMSIDRKHKALTLIKCPPLLVT
                SNIDITKEDKYKYLHTRVTTFTFPNPFPFDRNGNAVYELSNTNWKCFFERLSSSLDIQ
                DSEDEEDGSNSQAFRCVPGTVVRTL"
CDS             2723..3829
                /note="ORF E2 from bp 2696 to 3829"
                /product="regulatory protein"
                /gene="E2"
                /note="putative"
                /codon_start=1
                /translation="MEAIAKRLDACQEQLLELYEENSTDLHKHVLHWKCMRHESVLLY
                KAKQMGLSHIGMQVVPPLKVSEAKGHNAIEMQMHLESLLRTEYSMEPWTLQETSYEMW
                QTPPKRCFKKRGKTVEVKFDGCANNTMDYVVWTDVYVQDNDTWVKVHSMVDAKGIYYT
                CGQFKTYYVNFVKEAEKYGSTKHWEVCYGSTVICSPASVSSTTQEVSIPESTTYTPAQ
                TSTLVSSSTKEDAVQTPPRKRARGVQQSPCNALCVAHIGPVDSGNHNLITNNHDQHQR
                RNNSNSSATPIVQFQGESNCLKCFRYRLNDRHRHLFDLISSTWHWASSKAPHKHAIVT
                VTYDSEEQRQQFLDVVKIPPTISHKLGFMSLHLL"
CDS             3255..3584
                /note="ORF E4 from bp 3240 to 3584"
                /gene="E4"
                /note="putative"
                /codon_start=1
                /translation="MGAPNIGKYVMAAQLYVLLHLYLALHKKYPFLNLLHTPPHRPPP
                LCPQAPRKTQCKRRLGNEHEESNSPLATPCVWPTLDPWTVETTTSSLTITTSTKDGTT
                VTVQLRL"
CDS             3887..4162
                /note="ORF E5a from bp 3872 to 4162"
                /gene="E5a"
                /note="putative"
                /codon_start=1
                /translation="MEVVPVQIAAGTTSTFILPVIIAFVVCFVSIILIVWISEFIVYT
                SVLVLTLLLYLLLWLLLTTPLQFFLLTLLVCYCPALYIHYYIVTTQQ"
CDS             4159..4377
                /note="ORF E5b from bp 4003 to 4377"
                /gene="E5b"
                /note="putative"
                /codon_start=1
                /translation="MMLTCQFNDGDTWLGLWLLCAFIVGMLGLLLMHYRAVQGDKHTK
                CKKCNKHNCNDDYVTMHYTTDGDYIYMN"
CDS             4423..5802
                /note="ORF L2 from bp 4378 to 5802"
                /product="minor capsid protein"
                /gene="L2"
                /note="putative"
                /codon_start=1
                /translation="MAHSRARRRKRASATQLYQTCKLTGTCPPDVIPKVEHNTIADQI
                LKWGSLGVFFGGLGIGTGSGTGGRTGYVPLQTSAKPSITSGPMARPPVVVEPVAPSDP
                SIVSLIEESAIINAGAPEIVPPAHGGFTITSSETTTPAILDVSVTSHTTTSIFRNPVF
                TEPSVTQPQPPVEANGHILISAPTVTSHPIEEIPLDTFVVSSSDSGPTSSTPVPGTAP
                RPRVGLYSRALHQVQVTDPAFLSTPQRLITYDNPVYEGEDVSVQFSHDSIHNAPDEAF
                MDIIRLHRPAIASRRGLVRYSRIGQRGSMHTRSGKHIGARIHYFYDISPIAQAAEEIE
                MHPLVAAQDDTFDIYAESFEPGINPTQHPVTNISDTYLTSTPNTVTQPWGNTTVPLSL
                PNDLFLQSGPDITFPTAPMGTPFSPVTPALPTGPVFITGSGFYLHPAWYFARKRRKRI
                PLFFSDVAA"
CDS             5789..7291
                /note="ORF L1 from bp 5678 to 7291"
                /product="major capsid protein"
                /gene="L1"
                /note="putative"
```

```
                     /codon_start=1
                     /translation="MWRPSDSTVYVPPPNPVSKVVATDAYVTRTNIFYHASSSRLLAV
                     GHPYFSIKRANKTVVPKVSGYQYRVFKVVLPDPNKFALPDSSLFDPTTQRLVWACTGL
                     EVGRGQPLGVGVSGHPFLNKYDDVENSGSGGNPGQDNRVNVGMDYKQTQLCMVGCAPP
                     LGEHWGKGKQCTNTPVQAGDCPPLELITSVIQDGDMVDTGFGAMNFADLQTNKSDVPI
                     DICGTTCKYPDYLQMAADPYGDRLFFFLRKEQMFARHFFNRAGEVGEPVPDTLIIKGS
                     GNRTSVGSSIYVNTPSGSLVSSEAQLFNKPYWLQKAQGHNNGICWGNQLFVTVVDTTR
                     STNMTLCASVTTSSTYTNSDYKEYMRHVEEYDLQFIFQLCSITLSAEVMAYIHTMNPS
                     VLEDWNFGLSPPPNGTLEDTYRYVQSQAITCQKPTPEKEKPDPYKNLSFWEVNLKEKF
                     SSELDQYPLGRKFLLQSGYRGRSSIRTGVKRPAVSKASAAPKRKRAKTKR"
BASE COUNT     2438 a   1530 c   1699 g   2235 t

        1 GTTAATAACA ATCttggttt aaaaaaTAGg agggaccgaa aacggttcaa ccgaaaacgg
                                    ▲
             <- E8 end      E6 orf start ->
          ->  CAAT-box  <-
        (begins at bp 7898)

       61 ttgTATATAA accagcccta aaatttagca aacgaggCAT TATGgaaagt gcaaatgcct
                                                   _M__E__S_ _A__N__A__      6
            -> signal                          E6 cds ->
                                         cap site ->   <-

      121 ccacgtctgc aacgaccatA Gaccagttgt gcaagacgtt taatctatct atgcatacgt
          S__T__S__A __T__T__I_ _D__Q__L__ C__K__T__F __N__L__S_ _M__H__T__      26
                                /\ 3 sj

      181 tgcaaattaa ttgtgtgttt tgcaagaatg cactgaccac agcagagatt tattcatatg
          L__Q__I__N __C__V__F_ _C__K__N__ A__L__T__T __A__E__I_ _Y__S__Y__      46

      241 catataaaca cctaaaggtc ctgtttcgag gcggctatcc atatgcagcc tgcgcgtgct
          A__Y__K__H__ _L__K__V_ _L__F__R__ G__G__Y__P __Y__A__A_ _C__A__C__      66

      301 gcctagaatt tcatggaaaa atAaaccaat atagacactt tgattatgct ggatatgcaa
          C__L__E__F __H__G__K_ _I__N__Q__ Y__R__H__F __D__Y__A_ _G__Y__A__      86

      361 caacagttga agaagaaact aaacaagaca tcttagacgt gctaattcgg tgctacctgt
          T__T__V__E __E__E__T_ _K__Q__D__ I__L__D__V __L__I__R_ _C__Y__L__      106

      421 gtcacaaacc gctgtgTGAa gtagaaaagG Taaaacatat actaaccaag gcgcggttca
          C__H__K__P __L__C__E_ _V__E__K__ V__K__H__I __L__T__K_ _A__R__F__      126
                 E7 orf start ->        /\ 5 sj

      481 taaagctaaa ttgtacgtgg aAGggtcgct gcctacactg ctggacaacA TGcatggaag
          I__K__L__N __C__T__W_ _K__G__R__ C__L__H__C __W__T__T_ _C__M__E__      146
                                                        _ M__H__G__R      4
                                  /\ 3 sj              E7 cds ->

      541 acatgttacc cTAAaggata ttgtattaga cctgcaacct ccagaccctg tagggttaca
          D__M__L__P __$_                                                   150
          __H__V__T_ _L__K__D__ I__V__L__D __L__Q__P_ _P__D__P__ V__G__L__H      24
                       <- E6 end

      601 ttgctatgag caattagtag acagctcaga agatgaggtg gacgaagtgg acggacaaga
          __C__Y__E_ _Q__L__V__ D__S__S__E __D__E__V_ _D__E__V__ D__G__Q__D      44

      661 ttcacaacct ttaaaacaac atttccaaat agtgacctgt tgctgtggat gTGAcagcaa
          __S__Q__P_ _L__K__Q__ H__F__Q__I __V__T__C_ _C__C__G__ C__D__S__N      64
                                                    E1 orf start ->

      721 cgttcgactg gttgtgcagt gtacagaaac agacatcaga gaagtgcaac agcttctgtt
          __V__R__L_ _V__V__Q__ C__T__E__T __D__I__R_ _E__V__Q__ Q__L__L__L      84

      781 gggaacacta aacatagtgt gtcccatctg cgcaccgaAG accTAAcaac gATGgcggac
                                                             _M__A__D_      3
          __G__T__L_ _N__I__V__ C__P__I__C __A__P__K_ _T__$_                 98
                                           3 sj /\     E1 cds ->
                                                  <- E7 end
```

```
 841 gattcaggta cagaaaatga ggggtctggg tgtacaggat ggtttatggt agaagctata
     _D__S__G__ T__E__N__E __G__S__G_ _C__T__G__ W__F__M__V __E__A__I_      23

 901 gtgcaacacc caacaggtac acaaatatca gacgatgagg atgaggaggt ggaggacagt
     _V__Q__H__ P__T__G__T __Q__I__S_ _D__D__E__ D__E__E__V __E__D__S_      43

 961 gggtatgaca tggtggactt tattgatgac agcaatatta cacacaattc actggaagca
     _G__Y__D__ M__V__D__F __I__D__D_ _S__N__I__ T__H__N__S __L__E__A_      63

1021 caggcattgt ttaacaggca ggaggcggac acccattatg cgactgtgca ggacctaaaa
     _Q__A__L__ F__N__R__Q __E__A__D_ _T__H__Y__ A__T__V__Q __D__L__K_      83

1081 cgaaagtatt taggtagtcc atatgttagt cctataaaca ctatagccga ggcagtggaa
     _R__K__Y__ L__G__S__P __Y__V__S_ _P__I__N__ T__I__A__E __A__V__E_     103

1141 agtgaaataa gtccacgatt ggacgccatt aaacttacaa gacagccaaa aaaggtaaag
     _S__E__I__ S__P__R__L __D__A__I_ _K__L__T__ R__Q__P__K __K__V__K_     123

1201 cgacggctgt ttcaaaccag ggaactaacg gacagtggat atggctattc tgaagtggaa
     _R__R__L__ F__Q__T__R __E__L__T_ _D__S__G__ Y__G__Y__S __E__V__E_     143

1261 gctggaacgg gaacgcagGT agagaaacat ggcgtaccgg aaaatggggg agatggtcag
     _A__G__T__ G__T__Q__V __E__K__H_ _G__V__P__ E__N__G__G __D__G__Q_     163
                  5 sj /\

1321 gaaaaggaca caggaaggga catagagggg gaggaacata cagaggcgga agcgcccaca
     _E__K__D__ T__G__R__D __I__E__G_ _E__E__H__ T__E__A__E __A__P__T_     183

1381 aacagtgtac gggagcatgc aggcacagca ggaatattgg aattgttaaa atgtaaagat
     _N__S__V__ R__E__H__A __G__T__A_ _G__I__L__ E__L__L__K __C__K__D_     203

1441 ttacgggcag cattacttgg taagtttaaa gaatgctttg ggctgtcttt tatagattta
     _L__R__A__ A__L__L__G __K__F__K_ _E__C__F__ G__L__S__F __I__D__L_     223

1501 attaggccat ttaaaagtga taaaacaaca tgttttagatt gggtggtagc aggttttggt
     _I__R__P__ F__K__S__D __K__T__T_ _C__L__D__ W__V__V__A __G__F__G_     243

1561 atacatcata gcatatcaga ggcatttcaa aaattaattg agccattaag tttatatgca
     _I__H__H__ S__I__S__E __A__F__Q_ _K__L__I__ E__P__L__S __L__Y__A_     263

1621 catatacaat ggctaacaaa tgcatgggga atggtattgt tagtattatt aagatttaaa
     _H__I__Q__ W__L__T__N __A__W__G_ _M__V__L__ L__V__L__L __R__F__K_     283

1681 gtaaataaaa gtagaagtac cgttgcacgt acacttgcaa cgctattaaa tatacctgaa
     _V__N__K__ S__R__S__T __V__A__R_ _T__L__A__ T__L__L__N __I__P__E_     303

1741 aacaaatgt taatagagcc accaaaaata caaagtggtg ttgcagccct gtattggttt
     _N__Q__M__ L__I__E__P __P__K__I_ _Q__S__G__ V__A__A__L __Y__W__F_     323

1801 cgtacaggta tatcaaatgc cagtacagtt ataggggaag caccagaatg gataacacgc
     _R__T__G__ I__S__N__A __S__T__V_ _I__G__E__ A__P__E__W __I__T__R_     343

1861 caaacagtta ttgaacagg gttggcagac agtcagttta aattaacaga aatggtgcag
     _Q__T__V__ I__E__H__G __L__A__D_ _S__Q__F__ K__L__T__E __M__V__Q_     363

1921 tgggcgtatg ataatgacat atgcgaggag agtgaaattg catttgaata tgcacaaagg
     _W__A__Y__ D__N__D__I __C__E__E_ _S__E__I__ A__F__E__Y __A__Q__R_     383

1981 ggagattttg attctaatgc acgagcattt ttaaatagca atatgcaggc aaaatatgtg
     _G__D__F__ D__S__N__A __R__A__F_ _L__N__S__ N__M__Q__A __K__Y__V_     403

2041 aaagattgtg caactatgtg tagacattat aaacatgcag aaatgaggaa gatgtctata
     _K__D__C__ A__T__M__C __R__H__Y_ _K__H__A__ E__M__R__K __M__S__I_     423

2101 aaacaatgga taaaacatag gggttctaaa atagaaggca caggaaattg gaaaccaatt
     _K__Q__W__ I__K__H__R __G__S__K_ _I__E__G__ T__G__N__W __K__P__I_     443

2161 gtacaattcc tacgacatca aaatatagaa ttcattccgt ttttaacgtaa atttaaatta
     _V__Q__F__ L__R__H__Q __N__I__E_ _F__I__P__ F__L__T__K __F__K__L_     463

2221 tggctgcacg gtacgccaaa aaaaaactgc atagccatag taggccctcc agatactggg
     _W__L__H__ G__T__P__K __K__N__C_ _I__A__I__ V__G__P__P __D__T__G_     483
```

```
2281 aaatcgtact tttgtatgag tttaataagc tttctaggag gtacagttat tagtcatgta
     _K__S__Y__ F__C__M__S __L__I__S_ _F__L__G__ G__T__V__I __S__H__V_      503

2341 aattccagca gccatttttg gttgcaaccg ttagtagatg ctaaggtagc attgttagat
     _N__S__S__ S__H__F__W __L__Q__P_ _L__V__D__ A__K__V__A __L__L__D_      523

2401 gatgcaacac agccatgttg gatatatatg gatacatata tgagaaattt gttagatggt
     _D__A__T__ Q__P__C__W __I__Y__M_ _D__T__Y__ M__R__N__L __L__D__G_      543

2461 aatcctatga gtattgacag aaagcataaa gcattgacat taattaaatg tccacctctg
     _N__P__M__ S__I__D__R __K__H__K_ _A__L__T__ L__I__K__C __P__P__L_      563

2521 ctagtaacgt ccaacataga tattactaaa gaagatgaaat ataagtattt acatactaga
     _L__V__T__ S__N__I__D __I__T__K_ _E__D__K__ Y__K__Y__L __H__T__R_      583

2581 gtaacaacat ttacatttcc aaatccattc ccttttgaca gaaatgggaa tgcagtgtat
     _V__T__T__ F__T__F__P __N__P__F_ _P__F__D__ R__N__G__N __A__V__Y_      603

2641 gaactgtcaa atacaaactg gaaatgtttt tttgaaAGac tgtcgtcaag ccTAGacatt
     _E__L__S__ N__T__N__W __K__C__F_ _F__E__R__ L__S__S__S __L__D__I_      623
                                                  E2 orf start ->
                                            /\ 3 sj

2701 caggattctg aggacgagga agATGgaagc aatagccaag cgtttagatg cgtgccagga
     _Q__D__S__ E__D__E__E __D__G__S_ _N__S__Q__ A__F__R__C __V__P__G_      643
                           _M__E__A __I__A__K_ _R__L__D__ A__C__Q__E      13
                  E2 cds ->

2761 acagttgtta gaactttaTG Aagaaaacag tactgaccta cacaaacatg tattgcattg
     _T__V__V__ R__T__L__$ _                                            649
     __Q__L__L_ _E__L__Y__ E__E__N__S __T__D__L_ _H__K__H__ V__L__H__W      33
                        <- E1 end

2821 gaaatgcatg agacatgaaa gtgtattatt atataaagca aaacaaatgg gcctaagcca
     __K__C__M_ _R__H__E__ S__V__L__L __Y__K__A_ _K__Q__M__ G__L__S__H      53

2881 cataggaatg caagtagtgc caccattaaa ggtgtccgaa gcaaaaggac ataatgccat
     __I__G__M_ _Q__V__V__ P__P__L__K __V__S__E_ _A__K__G__ H__N__A__I      73

2941 tgaaatgcaa atgcatttag aatcattatt aaggactgag tatagtatgg aaccgtggac
     __E__M__Q_ _M__H__L__ E__S__L__L _R__T__E_ _Y__S__M__ E__P__W__T      93

3001 attacaagaa acaagttatg aaatgtggca aacaccacct aaacgctgtt ttaaaaaacg
     __L__Q__E_ _T__S__Y__ E__M__W__Q __T__P__P_ _K__R__C__ F__K__K__R      113

3061 gggcaaaact gtagaagtta aatttgatgg ctgtgcaaac aatacaatgg attatgtggt
     __G__K__T_ _V__E__V__ K__F__D__G __C__A__N_ _N__T__M__ D__Y__V__V      133

3121 atggacagat gtgtatgtgc aggacaatga cacctgggta aaggtgcata gtatggtaga
     __W__T__D_ _V__Y__V__ Q__D__N__D _T__W__V_ _K__V__H__ S__M__V__D      153

3181 tgctaagggt atatattaca catgtggaca atttaaaaca tattatgtaa actttgTAAa
     __A__K__G_ _I__Y__Y__ T__C__G__Q __F__K__T_ _Y__Y__V__ N__F__V__K      173
                                                  E4 orf start ->

3241 AGaggcagaa aagtATGgga gcaccaaaca ttgggaagta tgttatggca gcacagttat
                   _M__G_ _A__P__N__I__G__K__Y __V__M__A_ _A__Q__L__      15
     __E__A__E_ _K__Y__G__ S__T__K__H__W__E__V_ _C__Y__G__ S__T__V__I      193
            E4 cds ->
      /\ 3 sj

3301 atgttctcct gcatctgtat ctagcactac acaagaagta tccattcctg aatctactac
     Y__V__L__L __H__L__Y_ _L__A__L__ H__K__K__Y __P__F__L_ _N__L__L__      35
     __C__S__P_ _A__S__V__ S__S__T__T __Q__E__V_ _S__I__P__ E__S__T__T      213

3361 atacaccccc gcacagacct ccacccctgt gtcctcaagc accaaggaag acgcagtgca
     H__T__P__P __H__R__P_ _P__P__L__ C__P__Q__A __P__R__K_ _T__Q__C__      55
     __Y__T__P_ _A__Q__T__ S__T__L__V __S__S__S_ _T__K__E__ D__A__V__Q      233
```

```
3421 aacgccgcct aggaaacgag cacgaggagt ccaacagtcc ccttgcaacg ccttgtgtgt
     K__R__R__L _ G _N__E_ _H__E__E__ S__N__S_ P _L__A__T_ _P__C__V__      75
     __T__P__P_ _R__K__R__ A__R__G__V __Q__Q__S_ _P__C__N__ A__L__C__V     253

3481 ggcccacatt ggacccgtgg acagtggaaa ccacaacctc atcactaaca atcacgacca
     W__P__T__L __D__P__W_ _T__V__E__ T__T__T__S __S__L__T_ _I__T__T__      95
     __A__H__I_ _G__P__V__ D__S__G__N __H__N__L_ _I__T__N__ N__H__D__Q     273

3541 gcaccaaaga cggaacaaca gtaacagttc agctacgcct aTAGtgcaat ttcaagGTga
     S__T__K_ D _G__T__T_ _V__T__V__ Q__L__R__L __$_                       109
     __H__Q__R_ _R__N__N__ S__N__S__S __A__T__P_ _I__V__Q__ F__Q__G__E     293
                                                 <- E4 end
                                                        5 sj /\

3601 atcgaattgt ttaaagtgtt ttagatatag gctaaatgac agacacagac atttatttga
     __S__N__C_ _L__K__C__ F__R__Y__R __L__N__D_ R _H__R__ H__L__F__D      313

3661 tttaatatca tcaacgtggc actgggcctc ctcaaaggca ccacataaac atgccattgt
     __L__I__S_ _S__T__W__ H__W__A__S _ S _K__A_ _P__H__K__ H__A__I__V      333

3721 aactgtaaca tatgatagtg aggaacaaag gcaacagttt ttagatgttg taaaaatacc
     __T__V__T_ _Y_ D _S__ E__E__Q__R __Q__Q__F_ _L_ D _V__ V__K__I__P      353

3781 ccctacgatt agcacaaac tgggatttat gtcactgcac ctattgTAAt ttgtatatat
     __P__T__I_ _S_ H__K__ L__G__F__M __S__L__H_ _L__L__$_                 368
                                                 <- E2 end

3841 gtaaatgtgt aaatatatgg tattggtgTA Atacaactgt acatgtATGg aagtggtgcc
                                                    _M__ E__V__V__P      5
                          E5a orf start ->      E5a cds ->

3901 tgtacaaata gctgcaggaa caaccagcac attcatactg cctgttataa ttgcatttgt
     __V__Q__I_ _A__A__G__ T__T__S__T _ F _I__L_ _P__V__I__ I__A__F__V      25

3961 tgtatgtttt gttagcatca tacttattgt atggatatcT GAgtttattg tgtacacatc
     __V__C__F_ _V__S__I__ I__L__I__V __W__I__S_ _E_ F__I__ V__Y__T__S      45
                                    E5b orf start ->

4021 tgtgctagta ctaacactgc ttttatattt actattgtgg ctgctattaa caaccccctt
     __V__L__V_ _L__T__L__ L__L__Y__L __L__L__W_ _L__L__L__ T__T__P__L      65

4081 gcaatttttc ctactaactc tacttgtgtg ttactgtccc gcattgtata tacactacta
     __Q__F__F_ _L__L__T__ L__L__V__C __Y__C__P_ _A__L__Y__ I__H_ Y _Y      85

4141 tattgttacc acacagcaAT GAtgctaaca tgtcaattta atgatggaga tacctggctg
                        _M __M__L__T_ _C__Q__F__ N__D__G__D __T__W__L_      14
     __I__V_ T _T__Q__Q__ $_                                               91
              E5b cds -> <- E5a end

4201 ggtttgtggt tgttatgtgc ctttattgta gggatgttgg ggttattatt gatgcactat
     _G__L__W__ L__L__C__A __F__I__V_ _G_ M _L__ G__L__L__L __M__H__Y_      34

4261 agagctgtac aaggggataa acacaccaaa tgtaagaagt gtaacaaaca cactgtaat
     _R__A__V__ Q__G__D__K __H__T__K_ _C_ K _K__ C__N__K__H _ N _C__N_      54

4321 gatgattatg taactatgca tatatactact gatggtgatt atatatatat gaatTAGagt
     D _D__Y__ V__T__M__H _ Y _T_ T_ D _G__D__ Y__I__Y__M __N__$_          72
                                                          <- E5b end
                                                 L2 orf start ->

4381 aaacgtttt ttatatttgt aacAGtgtat gtttgtata ccATGgcaca tagtagggcc
                                                  _M__A__H __S__R__A_      6
                              /\ 3 sj     L2 cds ->

4441 cgacgacgca agcgtgcgtc agctacacag ctatatcaaa catgtaaact cactggaaca
     _R__R__R__ K__R__A__S __A__T__Q_ _L__Y__Q__ T__C__K__L __T__G__T_      26

4501 tgccccccag atgtaattcc taaggtggag cacaacacca ttgcagatca aatATTAAAA
     _C__P__P__ D__V__I__P __K__V__E_ _H__N__T__ I__A__D__Q __I__L__K_      46
                                                    signal ->
```

```
4561 tggggaagtt tgggggtgtt ttttggaggg ttgggtatag gcacggttc cggcactggg
     _W__G__S__ L__G__V__F __F__G__G_ _L__G__I__ G__T__G__S __G__T__G_      66

4621 ggtcgtactg gctatgttcc cttacaaact tctgcaaaac cttctattac tagtgggcct
     _G__R__T__ G__Y__V__P __L__Q__T_ _S__A__K__ P__S__I__T __S__G__P_      86

4681 atggctcgtc ctcctgtggt ggtggagcct gtggcccctt cggatccatc tattgtgtct
     _M__A__R__ P__P__V__V __V__E__P_ _V__A__P__ S__D__P__S __I__V__S_     106

4741 ttaattgaag aatcggcaat cattaacgca ggggcgcctg aaattgtgcc ccctgcacac
     _L__I__E__ E__S__A__I __I__N__A_ _G__A__P__ E__I__V__P __P__A__H_     126

4801 ggtgggttta caattacatc ctctgaaaca actacccctg caatattgga tgtatcagtt
     _G__G__F__ T__I__T__S __S__E__T_ _T__T__P__ A__I__L__D __V__S__V_     146

4861 actagtcaca ctactactag tatatttaga aatcctgtct ttacagaacc ttctgtaaca
     _T__S__H__ T__T__T__S __I__F__R_ _N__P__V__ F__T__E__P __S__V__T_     166

4921 caaccccaac cacccgtgga ggctaatgga catatattaa tttctgcacc cactgtaacg
     _Q__P__Q__ P__P__V__E __A__N__G_ _H__I__L__ I__S__A__P __T__V__T_     186

4981 tcacacccta tagaggaaat tcctttagat acttttgtgg tatcatctag tgatagcggt
     _S__H__P__ I__E__E__I __P__L__D_ _T__F__V__ V__S__S__S __D__S__G_     206

5041 cctacatcca gtacccctgt tcctggtact gcacctcggc ctcgtgtggg cctatatagt
     _P__T__S__ S__T__P__V __P__G__T_ _A__P__R__ P__R__V__G __L__Y__S_     226

5101 cgtgcattgc accaggtgca ggttacagac cctgcatttc tttccactcc tcaacgctta
     _R__A__L__ H__Q__V__Q __V__T__D_ _P__A__F__ L__S__T__P __Q__R__L_     246

5161 attacatatg ataaccctgt atatgaaggg gaggatgtta gtgtacaatt tagtcatgat
     _I__T__Y__ D__N__P__V __Y__E__G_ _E__D__V__ S__V__Q__F __S__H__D_     266

5221 tctatacaca atgcacctga tgaggctttt atggacataa ttcgtttgca cagacctgcc
     _S__I__H__ N__A__P__D __E__A__F_ _M__D__I__ I__R__L__H __R__P__A_     286

5281 attgcgtccc gacgtggcct tgtgcggtac agtcgcattg gacaacgggg gtctatgcac
     _I__A__S__ R__R__G__L __V__R__Y_ _S__R__I__ G__Q__R__G __S__M__H_     306

5341 actcgcagcg gaaagcacat aggggcccgc attcattatt tttatgatat ttcacctatt
     _T__R__S__ G__K__H__I __G__A__R_ _I__H__Y__ F__Y__D__I __S__P__I_     326

5401 gcacaggctg cagaagaaat agaaatgcac cctcttgtgg ctgcacagga tgatacattt
     _A__Q__A__ A__E__E__I __E__M__H_ _P__L__V__ A__A__Q__D __D__T__F_     346

5461 gatatttatg ctgaatcttt tgaacctggc attaacccta cccaacaccc tgttacaaat
     _D__I__Y__ A__E__S__F __E__P__G_ _I__N__P__ T__Q__H__P __V__T__N_     366

5521 atatcagata catatttaac ttccacacct aatacagtta cacaaccgtg gggtaacacc
     _I__S__D__ T__Y__L__T __S__T__P_ _N__T__V__ T__Q__P__W __G__N__T_     386

5581 acagttccat tgtcagttcc taatgacctg tttttacaat ctggccctga tataactttt
     _T__V__P__ L__S__L__P __N__D__L_ _F__L__Q__ S__G__P__D __I__T__F_     406

5641 cctactgcac ctatgggaac acccttttagt cctgTAActc ctgctttacc tacaggccct
     _P__T__A__ P__M__G__T __P__F__S_ _P__V__T__ P__A__L__P __T__G__P_     426
                                  L1 orf start ->

5701 gttttcatta caggttctgg attttatttg catcctgcat ggtattttgc acgtaaacgc
     _V__F__I__ T__G__S__G __F__Y__L_ _H__P__A__ W__Y__F__A __R__K__R_     446

5761 cgtaaacgta ttcccttatt tttttcAGAT GtggcggccT AGcgacagca cagtatatgt
     _R__K__R__ I__P__L__F __F__S__D_ _V__A__A__ $_                       459
                                _M __W__R__P_ _S__D__S__ T__V__Y__V      11
                          L1 cds ->           <- L2 end
                                /\ 3 sj

5821 gcctcctcct aaccctgtat ccaaagttgt tgccacggat gcttatgtta ctcgcaccaa
     __P__P__P_ _N__P__V__ S__K__V__V __A__T__D_ _A__Y__V__ T__R__T__N      31

5881 catattttat catgccagca gttctagact tcttgcagtg ggacatcctt atttttccat
     __I__F__Y_ _H__A__S__ S__S__R__L __L__A__V_ _G__H__P__ Y__F__S__I      51
```

```
 5941 aaaacgggct aacaaaactg ttgtgccaaa ggtgtcagga tatcaataca gggtatttaa
      __K__R__A_ _N__K__T__ V__V__P__K __V__S__G_ _Y__Q__Y__ R__V__F__K      71

 6001 ggtggtgtta ccagatccta acaaatttgc attgcctgac tcgtctcttt tcgatcccac
      __V__V__L_ _P__D__P__ N__K__F__A __L__P__D_ _S__S__L__ F__D__P__T      91

 6061 aacacaacgt ttagtatggg catgcacagg cctagaggtg ggcaggggac agccattagg
      __T__Q__R_ _L__V__W__ A__C__T__G __L__E__V_ _G__R__G__ Q__P__L__G     111

 6121 tgtgggtgta agtggacatc ctttcctaaa taaatatgat gatgttgaaa attcagggag
      __V__G__V_ _S__G__H__ P__F__L__N __K__Y__D_ _D__V__E__ N__S__G__S     131

 6181 tggtggtaac cctggacagg ataacagggt taatgtaggt atggattata aacaaacaca
      __G__G__N_ _P__G__Q__ D__N__R__V __N__V__G_ _M__D__Y__ K__Q__T__Q     151

 6241 attatgcatg gttggatgtg cccccccttt gggcgagcat tggggtaaag gtaaacagtg
      __L__C__M_ _V__G__C__ A__P__P__L __G__E__H_ _W__G__K__ G__K__Q__C     171

 6301 tactaataca cctgtacagg ctggtgactg cccgccctta gaacttatta ccagtgttat
      __T__N__T_ _P__V__Q__ A__G__D__C __P__P__L_ _E__L__I__ T__S__V__I     191

 6361 acaggatggc gatatggttg acacaggctt tggtgctatg aattttgctg atttgcagac
      __Q__D__G_ _D__M__V__ D__T__G__F __G__A__M_ _N__F__A__ D__L__Q__T     211

 6421 cAATAAAtca gatgttccta ttgacatatg tggcactaca tgtaaatatc cagattattt
      __N__K__S_ _D__V__P__ I__D__I__C __G__T__T_ _C__K__Y__ P__D__Y__L     231
signal ->

 6481 acaaatggct gcagacccat atggtgatag attatttttt tttctacgga aggaacaaat
      __Q__M__A_ _A__D__P__ Y__G__D__R __L__F__F_ _F__L__R__ K__E__Q__M     251

 6541 gtttgccaga catttttta acagggctgg cgaggtgggg gaacctgtgc ctgatacct
      __F__A__R_ _H__F__F__ N__R__A__G __E__V__G_ _E__P__V__ P__D__T__L     271

 6601 tataattaag ggtagtggaa atcgcacgtc tgtagggagt agtatatatg ttaacacccc
      __I__I__K_ _G__S__G__ N__R__T__S __V__G__S_ _S__I__Y__ V__N__T__P     291

 6661 gagcggctct ttggtgtcct ctgaggcaca attgtttaat aagccatatt ggctacaaaa
      __S__G__S_ _L__V__S__ S__E__A__Q __L__F__N_ _K__P__Y__ W__L__Q__K     311

 6721 agcccaggga cataacaatg gtatttgttg gggtaatcaa ctgtttgtta ctgtggtaga
      __A__Q__G_ _H__N__N__ G__I__C__W __G__N__Q_ _L__F__V__ T__V__V__D     331

 6781 taccacacgc agtaccaaca tgacattatg tgcatccgta actacatctt ccacatacac
      __T__T__R_ _S__T__N__ M__T__L__C __A__S__V_ _T__T__S__ S__T__Y__T     351

 6841 caattctgat tataaagagt acatgcgtca tgtggaagag tatgatttac aatttatttt
      __N__S__D_ _Y__K__E__ Y__M__R__H __V__E__E_ _Y__D__L__ Q__F__I__F     371

 6901 tcaattatgt agcattacat tgtctgctga agtaatggcc tatattcaca caatgaatcc
      __Q__L__C_ _S__I__T__ L__S__A__E __V__M__A_ _Y__I__H__ T__M__N__P     391

 6961 ctctgttttg gaagactgga actttgggtt atcgcctccc ccaaatggta cattagaaga
      __S__V__L_ _E__D__W__ N__F__G__L __S__P__P_ _P__N__G__ T__L__E__D     411

 7021 tacctatagg tatgtgcagt cacaggccat tacctgtcaa aagcccactc ctgaaaagga
      __T__Y__R_ _Y__V__Q__ S__Q__A__I __T__C__Q_ _K__P__T__ P__E__K__E     431

 7081 aaagccagat ccctataaga accttagttt ttgggaggtt aatttaaaag aaaagttttc
      __K__P__D_ _P__Y__K__ N__L__S__F __W__E__V_ _N__L__K__ E__K__F__S     451

 7141 tagtgaattg gatcagtatc ctttgggacg caagtttttg ttacaaagtg gatatagggg
      __S__E__L_ _D__Q__Y__ P__L__G__R __K__F__L_ _L__Q__S__ G__Y__R__G     471

 7201 acggtcctct attcgtacag gtgttaagcg ccctgctgtt tccaaagcct ctgctgcccc
      __R__S__S_ _I__R__T__ G__V__K__R __P__A__V_ _S__K__A__ S__A__A__P     491
```

```
7261 taaacgtaag cgcgccaaaa ctaaaaggTA ATATATGTGT ATATGTACTG TTATATATAT
     __K__R__K_ _R__A__K__ T__K__R__$ _                                  500
                                      ▲
                                    <- L1 end
                                      -> 24 bp tandem repeat  <-->
                                                        deletion
7321 GTGTGTATGT ACTGTTATGt atatgtgtgt gtgtgttctg tgtgtaatGT aagttatttg
    24 bp tandem repeat<-             ▲            5 sj /\
              deletion                                deletion
7381 tgtaatgtgt atgtgtgttt atgtgcAATA AAcaattacc tcttgttaca ccctgtgact
                          signal ->                ▲

7441 cagtggctgT TGCACGCGtt ttggtTTGCA CGCGccttac acacataagt aatatacatg
       repeat ->         repeat ->

7501 cacaatatat atatttttgt ttagaatact atactttat atttgcaacc gttttcggtt

7561 gcccttagca tacactttcc accatttgt tacaacgtgt ttcctgtTAA tcctatatat
                                              E8 orf start ->

7621 tttgtgccag gtacacattg ccctgccaag ttgcttgcca agtgcatcat atcctgccaa

7681 ccacacacct ggcgccaggg tgcggtattg ccttactcat aaacgtgtct ttgtgttata
                                              ▲

7741 ctttTATGca ctgtagccaa ctcttaaaag cattttggc ttgtagcagc acatttttt
   E8 cds ->

7801 gctcttactg tttggtatac aataacataa aaatgagtaa cctaaggtca cacacctgcg

7861 accggtttcg gttatccaca ccctacatat ttccttcTTA TA
                                        -> CAAT-box start
```

LCR

```
                idd
                222
                178
HPV6B           .TA   1-101
HPV6A           T--   1-101
W50             ...   1-101
T70             .--   1-101
```

E6 nuc

```
                223334
                252697
                113529
HPV6B           ACAACC   102-554
HPV6A           TGCTTT   102-554
```

E6 aa

```
                000001
                457892
                004876
HPV6B           THITIF
HPV6A           *Q****
```

E7 nuc

```
                78
                92
                13
HPV6B           AC   530-826
HPV6A           GA   530-826
```

E7 aa

```
                89
                88
HPV6B           NT
HPV6A           D*
```

E1 nuc

```
                1111111111111112222222
                0024455556778891123567
                1192933356446729901550
                1266445949036863974639
HPV6B           ACATATTTGTACACGCTCCTAT   832-2781
HPV6A           CTCCTGCCAGCTTTATAGTGGA   832-2781
```

E1 aa

```
                00112^22222333334444566
                6659233348004465559702
                0159155610345954695586
HPV6B           SLVLILLDGLENTHAFPTLDTS
HPV6A           *****AA**VD******S*EA*
```

E2 nuc

```
                2233333333333333333
                8911234445666777778
                0745780365049368890
                1372174301422441735
HPV6B           CGAATTAGCCCGTGGCCCA   2723-3829
HPV6A           AACTACCAAATACCAATGG   2723-3829
```

E2 aa

```
                00111222222333333333
                28448223479023455556
                74243287674748835571
HPV6B           HRNTHLKRSRSRSDDPTSG
HPV6A           NKTSQPQ****KPHN**R*
```

E4 nuc

```
                333333
                234445
                780365
                174301
HPV6B           TTAGCC   3255-3584
HPV6A           ACCAAA   3255-3584
```

E4 aa

```
                045669
                650099
HPV6B           ILPGPD
HPV6A           N**EHE
```

E5a nuc

```
                3344444
                8900111
                9304344
                8438619
HPV6B           GCGTTC   3887-4162
HPV6A           AACCCCA  3887-4162
```

E5a aa

```
                0135888
                4694458
HPV6B           VFEYYT
HPV6A           *LD*H*N
```

E5b nuc

```
                44444444444
                12222333333
                93579124445
                44156322562
HPV6B           CAGGGAATACA   4159-4377
HPV6A           AGACCCCCGAC   4159-4377
```

E5b aa

```
                12334556ĉ66
                26196252335
HPV6B           TMLGKNDYTTD
HPV6A           *V**NTAHDDA
```

noncoding

```
                444
                334
                881
                462
HPV6B           CGC   4378-4422
HPV6A           ATT   4378-4422
```

L2 nuc

```
                44444444455555555
                45666778900024456
                97044356722780891
                18556159505706969
HPV6B           CGGCATGCGGACCGGCA   4423-5802
HPV6A           TTCGGCATAACATAAAG   4423-5802
```

L2 aa

```
                0000ĉ0111122223333
                25677014800182599
                32155319501968629
HPV6B           LVTQQSSHVVSRAQGLQ
HPV6A           ***GG***II****DI*
```

L1 nuc

```
                556666667
                990015662
                295709261
                322348519
HPV6B           AGCAAACGA   5789-7291
HPV6A           TATGCT-AC   5789-7291
India-D9             T--    6486-6739
Georgia-B5           T--    6507-6661
India-D4             T--    6486-6739
India-D5             T--    6486-6739
India-D7             T--    6486-6739
Philippines-A4       T--    6486-6739
Georgia-B4           T-A    6486-6739
Georgia-B6           T-A    6486-6739
Georgia-G6           T-A    6486-6739
Alaska-C36           T-A    6507-6661
Georgia-G4           TA-    6486-6739
Georgia-G7           TA-    6486-6739
Georgia-B1           TAA    6486-6739
Georgia-G1           TAA    6486-6739
Georgia-G2           TAA    6486-6739
Georgia-G3           TAA    6486-6739
Georgia-G5           TAA    6486-6739
Philippines-A6       TAA    6486-6739
```

L1 aa

```
                000012224
                468907797
                588560917
HPV6B           GRFLRTRPT
HPV6A           ******-**
India-D9             *--
Georgia-B5           *--
India-D4             *--
India-D5             *--
India-D7             *--
Philippines-A4       *--
Georgia-B4           *-*
Georgia-B6           *-*
Georgia-G6           *-*
Alaska-C36           *-*
Georgia-G4           **-
Georgia-G7           **-
Georgia-B1           ***
Georgia-G1           ***
Georgia-G2           ***
Georgia-G3           ***
Georgia-G5           ***
Philippines-A6       ***
```

LCR

```
                                  insertion                 insertion
                  ┌──────────────────────────┐┌──────────────────────────┐
                  77777777777777777777777777777777777777777777777777777
                  22222222222222222222222233333333333333333333333333333
                  99999999999999999999999234555555555555555555555555555
                  22222222222222222222225991111111111111111111111111111
HPV6B             ........................GGG.......................   7292-7902
HPV6A             ..........................---ATGTACTGTTATATGTATGTGTG   7292-7902
W50                                  ---ATGTACTGTTATATGTATGTGTG   7292-7902
ML2               TATATGTGTATATGTACTG.TATA---ATGTACTGTTATATGTATGTGTG   7292-7420
BAM               TATATGTGTATATGTACTGTTATA---ATGTACTGTTATATGTATGTGTG   7292-7350
6R                ..........................---ATGTACTGTTATATGTATGTGTG   7292-7420
1082              ..........................--TATGTAGTGTTATATGTATGTGTG   7292-7436
1083              ..........................---ATGTACTGTTATATGTATGTGTG   7292-7436
1084              ..........................--TATGTAGTGTTATATGTATGTGTG   7292-7436
1086              ..........................---ATGTACTGTTATATGTATGTGTG   7292-7434
1094              ..........................---ATGTACTGTTATATGTATGTGTG   7292-7436
T70                                AA-ATGTACTGTTATGTATATGTGTG   7292-7902
PPHRRA                             AA-ATGTACTGTTATGTATATGTGTG   7296-7420
```

LCR cont'd.

```
                                        insertion
                  ┌──────────────────────────────────────────────────┐
                  77777777777777777777777777777777777777777777777777777
                  33333333333333333333333333333333333333333333333333333
                  55555555555555555555555555555555555555555555555555555
                  11111111111111111111111111111111111111111111111111111
HPV6B             ...................................................   7292-7902
HPV6A             TTGTATATATGTGTGTATATATGTGTATGTGTGTATATGTATATGTATGT   7292-7902
W50               TTGTATATATGTGTGTATATATGTGTATGTGTGTATATGTATATGTATGT   7292-7902
ML2               TTGTATATATGTGTGTATATATGTGTCTGTGTGTATATGTATATGTATGT   7292-7420
BAM               TTGTATATATGTGTGTATATATGTGTCTGTGTGTA                  7292-7350
6R                TTGTATATATGTGTGTATATATGTGTCTGTGTGTATATGT..........   7292-7420
1082              TTGTATATATGTGTGTATATATGTGTATGTGTGTATATGTATATGTATGT   7292-7436
1083              TTGTATATATGTGTGTATATATGTGTATGTGTGTATATGTATATGTATGT   7292-7436
1084              TTGTATATATGTGTGTATATATGTGTATGTGTGTATATGTATATGTATGT   7292-7436
1086              TTGTATATATGTGTGTATATATGTGTATGTGTGTATATGTATATGTATGT   7292-7434
1094              TTGTATATATGTGTGTATATATGTGTATGTGTGTATATGTATATGTATGT   7292-7436
T70               T.......ATGTAC.TGT.TATATGTATGTGTGT...TGTATATAT....   7292-7902
PPHRRA            T.......ATGTAC.TGT.TATATGTATGT..........ATATAT....   7296-7420
```

LCR cont'd.

```
                                  insertion                 deletion
                  ┌──────────────────────────┐┌──────────────────────────┐
                  77777777777777777777777777777777777777777777777777777
                  33333333333333333333333333333333333333333333333333333
                  55555555555555555555555566666666667777777777888888
                  11111111111111111111178901234567890123456789012345
HPV6B             ..................TCTGTGTGTAATGTAAGTTATTTGTGTAA   7292-7902
HPV6A             GTTGTGTATATATATGTGTGT-A-------------T------------   7292-7902
W50               GTTGTGTATATATATGTGTGT----------------T------------   7292-7902
ML2               GTTGTGTATATATATGTGTGT----------------------------   7292-7420
6R                .......ATATATATGTGTGT----------------------------   7292-7420
1082              GTTGTGTATATATATGTGTGT-A-------------T------------   7292-7436
1083              GTTGTGTATATATATGTGTGT-A-------------T------------   7292-7436
1084              GTTGTGTATATATATGTGTGT-A-------------T------------   7292-7436
1086              GTTGTGTATATATATGTGTGT----------------T------------   7292-7434
1094              GTTGTGTATATATATGTGTGT-A-------------T------------   7292-7436
T70               ...GTGTGTATATATGTGTCT.......................... .   7292-7902
PPHRRA            ...GTGTGTATATATGTGTCT..........................     7296-7420
```

LCR cont'd.

deletion insertion deletion d iiii

```
7777777777777777777777777777777777777777777777777
3333333333333444444444444444444444555555666667777
8888999999999912222222222222222222222368005572222
6789012345678992222222222222223456724947526345222
HPV6B              TGTGTATGTGTGTTC..............CTTGTTTACCAATCGCT....   7292-7902
HPV6A              --------------A..................CATG-GGCA-TGTT   7292-7902
W50                --------------.............----------C---A-....   7292-7902
ML2                ---------------                                    7292-7420
6R                 ---------------                                    7292-7420
1082               --------------AACATTATTGTATAT------                7292-7436
1083               --------------AACATTATTGTATATG-----                7292-7436
1084               --------------AACATTATTGTATAT------                7292-7436
1086               --------------.............------                7292-7434
1094               --------------A...................                7292-7436
T70                ............A-.............--------------....   7292-7902
PPHRRA             ............A-                                    7296-7420
SN6-11                                                  C-------....   7582-7845
IN6-6                                                   C---A-....   7582-7845
G6-42                                                   ----A-TGTT   7582-7845
NY6-16                                                  C--CA-....   7582-7845
B6-5                                                    --G-A-TGTT   7582-7845
B6-15/1                                                 C--CA-....   7582-7845
G6-78                                                   ----A-TGTT   7582-7845
NY6-1                                                   C--CA-....   7582-7845
SN6-3                                                   -GG-A-TGTT   7582-7845
HPV-6C                                                  -GG-A-TGTT   7582-7845
J6-8                                                    -GG-A-TGTT   7582-7845
HPV-6A                                                  -GG-A-TGTT   7582-7845
SN6-1                                                   -GG-A-....   7582-7845
SN6-6A                                                  -GG-A-TGTT   7582-7845
B6-1                                                    C--CAATGTT   7582-7845
G6-6                                                    -GGCA-TGTT   7582-7845
NY6-19                                                  -GGCA-TGTT   7582-7845
AM6-1                                                   -GGCA-TGTT   7582-7845
```

LCR cont'd.

insertion d

```
77777777777777777777777
77777777777777777777888888
22222222222222222499001146
22222222222222225609035700
HPV6B              ...............CACTTTGAAG   7292-7902
HPV6A              TATTGCCACTGCAATA-CA------A   7292-7902
W50                ...............---------A   7292-7902
T70                ...............----------   7292-7902
SN6-11             ...............---------    7582-7845
IN6-6              ...............---------    7582-7845
G6-42              TATTGCCACTGCAATA---------    7582-7845
NY6-16             ...............---------    7582-7845
B6-5               TATTGCCACTGCAATA---------    7582-7845
B6-15/1            ...............--A------    7582-7845
G6-78              TATTGCCACTGCAATA--A------    7582-7845
NY6-1              ...............------C--    7582-7845
SN6-3              TATTGCCACTGCAATA---------    7582-7845
HPV-6C             TATTGCCACTGCAATA--A------    7582-7845
J6-8               TATTGCCACTGCAATA-------G-    7582-7845
HPV-6A             TATTGCCACTGCAATA--A----G-    7582-7845
SN6-1              ...............T----C---    7582-7845
SN6-6A             TATTGCCACTGCAATA-CA------    7582-7845
B6-1               TATTGCCACTGCAATA-CAG----G    7582-7845
G6-6               TATTGCCACTGCAATA-CA------    7582-7845
NY6-19             TATTGCCACTGCAATA-CA-.----    7582-7845
AM6-1              TATTGCCACTGCAATA-CAG-----    7582-7845
```

```
LOCUS       HPV11        7931 bp ds-DNA   circular  VRL       30-SEP-1988
DEFINITION  Human papillomavirus type 11 (HPV-11), complete genome.
ACCESSION   M14119
SOURCE      Human laryngeal papillomavirus type 11 DNA recovered from a
            laryngeal papilloma.
  ORGANISM  Human papillomavirus type 11
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
COMMENT     Naturally occurring variants of HPV11 have been observed in E6, E7,
            E1, E2, E4, E5A, E5B and LCR.
            Reference: Dollard,S.C., Chow,L.T., Kreider,J.W., Broker,T.R., Lill,N.L.,
                and Howett,M.K.,  Virology 171, 294-7 (1989)
                Variant Accession Nos. J04351
            Reference: Heinzel,P.A., Chan,S.Y., Ho,L., O'Connor,M., et al.
                J. Clin. Microbiol. 33, 1746-54 (1995)
                Variant sequences obtained from authors.
            Reference: Metcalfe,L., Chen,S.L., and Mounts,P.,
                Virus Genes 3, 11-27 (1989)
                Variant Accession Nos. M26656
            Reference: McGlennen,R.C., Ghai,J., Ostrow,R.S., LaBresh,K.,
                Schneider,J.F., and Faras,A.J.
                Cancer Res. 52, 5872-8 (1992)
                Variant sequences constructed by editing reference sequence
                according to figures in article, with corrections from authors.
FEATURES             Location/Qualifiers
     CDS             102..554
                     /note="ORF E6 from bp 18 to 554"
                     /product="transforming protein"
                     /gene="E6"
                     /note="putative"
                     /codon_start=1
                     /translation="MESKDASTSATSIDQLCKTFNLSLHTLQIQCVFCRNALTTAEIY
                     AYAYKNLKVVWRDNFPFAACACCLELQGKINQYRHFNYAAYAPTVEEETNEDILKVLI
                     RCYLCHKPLCEIEKLKHILGKARFIKLNNQWKGRCLHCWTTCMEDLLP"
     CDS             530..826
                     /note="ORF E7 from bp 494 to 826"
                     /product="transforming protein"
                     /gene="E7"
                     /note="putative"
                     /codon_start=1
                     /translation="MHGRLVTLKDIVLDLQPPDPVGLHCYEQLEDSSEDEVDKVDKQD
                     AQPLTQHYQILTCCCGCDSNVRLVVECTDGDIRQLQDLLLGTLNIVCPICAPKP"
     CDS             832..2781
                     /note="ORF E1 from bp 715 to 2781"
                     /product="replication protein"
                     /gene="E1"
                     /note="putative"
                     /codon_start=1
                     /translation="MADDSGTENEGSGCTGWFMVEAIVEHTTGTQISEDEEEEVEDSG
                     YDMVDFIDDRHITQNSVEAQALFNRQEADAHYATVQDLKRKYLGSPYVSPISNVANAV
                     ESEISPRLDAIKLTTQPKKVKRRLFETRELTDSGYGYSEVEAATQVEKHGDPENGGDG
                     QERDTGRDIEGEGVEHREAEAVDDSTREHADTSGILELLKCKDIRSTLHGKFKDCFGL
                     SFVDLIRPFKSDRTTCADWVVAGFGIHHSIADAFQKLIEPLSLYAHIQWLTNAWGMVL
                     LVLIRFKVNKSRCTVARTLGTLLNIPENHMLIEPPKIQSGVRALYWFRTGISNASTVI
                     GEAPEWITRQTVIEHSLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
                     FLNSNMQAKYVKDCAIMCRHYKHAEMKKMSIKQWIKYRGTKVDSVGNWKPIVQFLRHQ
                     NIEFIPFLSKLKLWLHGTPKKNCIAIVGPPDTGKSCFCMSLIKFLGGTVISYVNSCSH
                     FWLQPLTDAKVALLDDATQPCWTYMDTYMRNLLDGNPMSIDRKHRALTLIKCPPLLVT
                     SNIDISKEEKYKYLHSRVTTFTFPNPFPFDRNGNAVYELSDANWKCFFERLSSSLDIE
                     DSEDEEDGSNSQAFRCVPGSVVRTL"
     CDS             2723..3826
                     /note="ORF E2 from bp 2696 to 3826"
                     /product="regulatory protein"
                     /gene="E2"
```

```
                     /note="putative"
                     /codon_start=1
                     /translation="MEAIAKRLDACQDQLLELYEENSIDIHKHIMHWKCIRLESVLLH
                     KAKQMGLSHIGLQVVPPLTVSETKGHNAIEMQMHLESLAKTQYGVEPWTLQDTSYEMW
                     LTPPKRCFKKQGNTVEVKFDGCEDNVMEYVVWTHIYLQDNDSWVKVTSSVDAKGIYYT
                     CGQFKTYYVNFNKEAQKYGSTNHWEVCYGSTVICSPASVSSTVREVSIAEPTTYTPAQ
                     TTAPTVSACTTEDGVSAPPRKRARGPSTNNTLCVANIRSVDSTINNIVTDNYNKHQRR
                     NNCHSAATPIVQLQGDSNCLKCFRYRLNDKYKHLFELASSTWHWASPEAPHKNAIVTL
                     TYSSEEQRQQFLNSVKIPPTIRHKVGFMSLHLL"
     CDS             3255..3581
                     /note="ORF E4 from bp 3231 to 3581"
                     /gene="E4"
                     /note="putative"
                     /codon_start=1
                     /translation="MVVPIIGKYVMAAQLYVLLHLYLALYEKYPLLNLLHTPPHRPPP
                     LQCPPAPRKTACRRRLGSEHVDRPLTTPCVWPTSDPWTVQSTTSSLTITTSTKEGTTV
                     TVQLRL"
     CDS             3871..4146
                     /note="ORF E5A from bp 3862 to 4146"
                     /gene="E5A"
                     /note="putative"
                     /codon_start=1
                     /translation="MEVVPVQIAAATTTTLILPVVIAFAVCTLSIVLIILISDFVVYT
                     SVLVLTLLLYLLLWLLLTTPLQFFLLTLCVCYFPAFYIHIYIVQTQQ"
BASE COUNT     2406 a   1519 c   1736 g   2270 t

        1 cttaataaCA ATCTTAGTTT AAAAaagagg agggACCGAA AACGGTtcaA CCGAAAACGG
           signal ->        -> signal          -> E2 bind      -> E2 bind
               E6 orf start ->

       61 TtataTATAA Accagcccaa aaaattagca gacgaggcat tATGgaaagt aaagatgcct
                                                   _M__E__S_ _K__D__A__      6
       signal ->   <-                          E6 cds ->

      121 ccacgtctgc aacatcgata gaccagttgt gcaagacgtt taatctttct ttgcacactc
          S__T__S__A __T__S__I_ _D__Q__L__ C__K__T__F __N__L__S_ _L__H__T__      26

      181 tgcaaattca gtgcgtgttt tgcaggaatg cactgaccac cgcagagata tatgcatatg
          L__Q__I__Q __C__V__F_ _C__R__N__ A__L__T__T __A__E__I_ _Y__A__Y__      46

      241 cctataagaa cctaaaggtt gtgtggcgag acaactttcc ctttgcagcg tgtgcctgtt
          A__Y__K__N __L__K__V_ _V__W__R__ D__N__F__P __F__A__A_ _C__A__C__      66

      301 gcttagaact gcaagggaaa attaaccaat atagacactt taattatgct gcatatgcac
          C__L__E__L __Q__G__K_ _I__N__Q__ Y__R__H__F __N__Y__A_ _A__Y__A__      86

      361 ctacagtaga agaagaaacc aatgaagata ttttaaaagt gttaattcgt tgttacctgt
          P__T__V__E __E__E__T_ _N__E__D__ I__L__K__V __L__I__R_ _C__Y__L__      106

      421 gtcacaagcc gttgtgtgaa atagaaaaac taaagcacat attgggaaag gcacgcttca
          C__H__K__P __L__C__E_ _I__E__K__ L__K__H__I __L__G__K_ _A__R__F__      126

      481 taaaactaaa TAAccagtgg aagggtcgtt gcttacactg ctggacaacA TGcatggaag
                                                          _ M__H__G__R      4
          I__K__L__N __N__Q__W_ _K__G__R__ C__L__H__C __W__T__T_ _C__M__E__      146
           E7 orf start ->                              E7 cds ->

      541 acttgttacc cTAAaggata tagtactaga cctgcagcct cctgaccctg tagggttaca
          __L__V__T_ _L__K__D__ I__V__L__D __L__Q__P_ _P__D__P__ V__G__L__H      24
          D__L__L__P __$_                                                      150
                       <- E6 end

      601 ttgctatgag caattagaag acagctcaga agatgaggtg gacaaggtgg acaaacaaga
          __C__Y__E_ _Q__L__E__ D__S__S__E __D__E__V_ _D__K__V__ D__K__Q__D      44

      661 cgcacaacct ttaacacaac attaccaaat actgacctgt tgctgtggat gTGAcagcaa
          _A__Q__P_ _L__T__Q__ H__Y__Q__I __L__T__C_ _C__C__G__ C__D__S__N      64
                                                    E1 orf start ->
```

```
 721 cgtccgactg gttgtggagt gcacagacgg agacatcaga caactacaag acctttgct
     __V__R__L_ _V__V__E__ C__T__D__G __D__I__R_ _Q__L__Q__ D__L__L__L      84

 781 gggcacacta aatattgtgt gtcccatctg cgcaccaaaa ccaTAAcaag gATGgcggac
                                                          _M__A__D_      3
     __G__T__L_ _N__I__V__ C__P__I__C __A__P__K_ _P__$_                  98
                                                  E1 cds ->
                                                 <- E7 end

 841 gattcaggta cagaaaatga ggggtcgggg tgtacaggat ggtttatggt agaagccata
     _D__S__G__ T__E__N__E __G__S__G_ _C__T__G__ W__F__M__V __E__A__I_      23

 901 gtagagcaca ctacaggtac acaaatatca gaagatgagg aagaggaggt ggaggacagt
     _V__E__H__ T__T__G__T __Q__I__S_ _E__D__E__ E__E__E__V __E__D__S_      43

 961 gggtatgaca tggtggactt tattgatgac aggcatatta cacaaaattc tgtggaagca
     _G__Y__D__ M__V__D__F __I__D__D_ _R__H__I__ T__Q__N__S __V__E__A_      63

1021 caggcattgt ttaataggca ggaggcggat gctcattatg cgactgtgca ggacctaaaa
     _Q__A__L__ F__N__R__Q __E__A__D_ _A__H__Y__ A__T__V__Q __D__L__K_      83

1081 cgaaagtatt taggcagtcc atatgtaagt cctataagca atgtagctaa tgcagtagaa
     _R__K__Y__ L__G__S__P __Y__V__S_ _P__I__S__ N__V__A__N __A__V__E_     103

1141 agtgagataa gtccacggtt agacgccatt aaacttacaa cacagccaaa aaaggtaaag
     _S__E__I__ S__P__R__L __D__A__I_ _K__L__T__ T__Q__P__K __K__V__K_     123

1201 cgacggctgt ttgaaacacg ggaattaacg gacagtggat atggctattc tgaagtggaa
     _R__R__L__ F__E__T__R __E__L__T_ _D__S__G__ Y__G__Y__S __E__V__E_     143

1261 gctgcaacgc aggtagagaa acatggcgac ccggaaaatg ggggagatgg tcaggaaagg
     _A__A__T__ Q__V__E__K __H__G__D_ _P__E__N__ G__G__D__G __Q__E__R_     163

1321 gacacaggga gggacataga gggtgagggg gtggaacata gagaggcgga agcagtagac
     _D__T__G__ R__D__I__E __G__E__G_ _V__E__H__ R__E__A__E __A__V__D_     183

1381 gacagcaccc gagagcatgc agacacatca ggaatattag aattactaaa atgtaaggat
     _D__S__T__ R__E__H__A __D__T__S_ _G__I__L__ E__L__L__K __C__K__D_     203

1441 atacgatcta cattacatgg taagtttaaa gactgctttg ggctgtcatt tgttgattta
     _I__R__S__ T__L__H__G __K__F__K_ _D__C__F__ G__L__S__F __V__D__L_     223

1501 attaggccat ttaaaagtga tagaaccaca tgtgccgatt gggtggttgc aggatttggt
     _I__R__P__ F__K__S__D __R__T__T_ _C__A__D__ W__V__V__A __G__F__G_     243

1561 atacatcata gcatagcaga tgcatttcaa aagttaattg agccattaag tttatatgca
     _I__H__H__ S__I__A__D __A__F__Q_ _K__L__I__ E__P__L__S __L__Y__A_     263

1621 catatacaat ggcttacaaa tgcatgggga atggtactat tagtattaat aaggtttaaa
     _H__I__Q__ W__L__T__N __A__W__G_ _M__V__L__ L__V__L__I __R__F__K_     283

1681 gtaaataaga gcagatgtac cgtggcacgt acattaggta cgttattaaa tatacctgaa
     _V__N__K__ S__R__C__T __V__A__R_ _T__L__G__ T__L__L__N __I__P__E_     303

1741 aatcacatgt taattgagcc tcctaaaata caaagtggcg tacgagccct gtattggttt
     _N__H__M__ L__I__E__P __P__K__I_ _Q__S__G__ V__R__A__L __Y__W__F_     323

1801 aggacaggca tttcaaatgc aagtacagtt ataggggagg cgccggaatg gataacgcgc
     _R__T__G__ I__S__N__A __S__T__V_ _I__G__E__ A__P__E__W __I__T__R_     343

1861 cagaccgtta ttgaacatag tttggctgac agtcaattta aattaactga aatggtgcag
     _Q__T__V__ I__E__H__S __L__A__D_ _S__Q__F__ K__L__T__E __M__V__Q_     363

1921 tgggcatatg ataatgatat ttgtgaagaa agtgagatag catttgaata tgcacagcgt
     _W__A__Y__ D__N__D__I __C__E__E_ _S__E__I__ A__F__E__Y __A__Q__R_     383

1981 ggagactttg actccaatgc aagggccttt ttaaatagta atatgcaggc taaatatgta
     _G__D__F__ D__S__N__A __R__A__F_ _L__N__S__ N__M__Q__A __K__Y__V_     403

2041 aaagattgtg caattatgtg cagacattat aaacatgcag aaatgaaaaa gatgtctatt
     _K__D__C__ A__I__M__C __R__H__Y_ _K__H__A__ E__M__K__K __M__S__I_     423
```

```
2101 aaacaatgga ttaagtatag gggtactaaa gttgacagtg taggtaactg gaagccaatt
     _K__Q__W__ I__K__Y__R __G__T__K_ _V__D__S__ V__G__N__W __K__P__I_      443

2161 gtgcagtttc taagacatca aaacatagaa tttattccat ttttaagcaa actaaaatta
     _V__Q__F__ L__R__H__Q __N__I__E_ _F__I__P__ F__L__S__K __L__K__L_      463

2221 tggctgcacg gaacgcccaa aaaaaattgt atagccattg tagggccacc tgacactggg
     _W__L__H__ G__T__P__K __K__N__C_ _I__A__I__ V__G__P__P __D__T__G_      483

2281 aagtcgtgct tttgcatgag tttaattaag tttttggggg gaacagttat tagttatgtt
     _K__S__C__ F__C__M__S __L__I__K_ _F__L__G__ G__T__V__I __S__Y__V_      503

2341 aattcctgca gccatttctg gctacagcca ctaacggatg caaaagtggc attattggat
     _N__S__C__ S__H__F__W __L__Q__P_ _L__T__D__ A__K__V__A __L__L__D_      523

2401 gatgccacac aaccatgttg gacatatatg gatacatata tgagaaacct attagatggt
     _D__A__T__ Q__P__C__W __T__Y__M_ _D__T__Y__ M__R__N__L __L__D__G_      543

2461 aatcctatga gcatagatag aaaacataga gcattaacat taattaagtg tccaccgcta
     _N__P__M__ S__I__D__R __K__H__R_ _A__L__T__ L__I__K__C __P__P__L_      563

2521 ctggttacat caaatataga cattagcaaa gaggagaaat acaaatattt acatagtaga
     _L__V__T__ S__N__I__D __I__S__K_ _E__E__K__ Y__K__Y__L __H__S__R_      583

2581 gttaccacat ttacatttcc aaatccattc ccctttgaca gaaatgggaa tgcagtatat
     _V__T__T__ F__T__F__P __N__P__F_ _P__F__D__ R__N__G__N __A__V__Y_      603

2641 gaactatcag atgcaaactg gaaatgtttc tttgaaagac tgtcgtccag ccTAGacatt
     _E__L__S__ D__A__N__W __K__C__F_ _F__E__R__ L__S__S__S __L__D__I_      623
                                                E2 orf start ->

2701 gaggattcag aggacgagga agATGgaagc aatagccaag cgtttagatg cgtgccagga
     _E__D__S__ E__D__E__E __D__G__S_ _N__S__Q__ A__F__R__C __V__P__G_      643
                           _M__E__A __I__A__K_ _R__L__D__ A__C__Q__D      13
                  E2 cds ->

2761 tcagttgtta gaactttaTG Aagaaaacag tattgatata cacaaacaca ttatgcattg
     _S__V__V__ R__T__L__$ _                                            649
     __Q__L__L_ _E__L__Y__ E__E__N__S __I__D__I_ _H__K__H__ I__M__H__W      33
                      <- E1 end

2821 gaaatgcata cgattggaaa gtgtattact acacaaagca aaacaaatgg gcctgagcca
     __K__C__I_ _R__L__E__ S__V__L__L __H__K__A_ _K__Q__M__ G__L__S__H      53

2881 catcgggcta caagtagtac caccattaac tgtgtcagag actaaaggac ataatgctat
     __I__G__L_ _Q__V__V__ P__P__L__T __V__S__E_ _T__K__G__ H__N__A__I      73

2941 tgaaatgcaa atgcatttag aatccttagc aaaaactcag tatggtgtgg aaccttggac
     __E__M__Q_ _M__H__L__ E__S__L__A __K__T__Q_ _Y__G__V__ E__P__W__T      93

3001 attacaggac accagttatg aaatgtggct aacaccaccc aaacggtgct ttaaaaaaca
     __L__Q__D_ _T__S__Y__ E__M__W__L __T__P__P_ _K__R__C__ F__K__K__Q      113

3061 gggaaatact gtggaggtaa aatttgatgg ctgtgaagac aatgtaatgg agtatgtggt
     __G__N__T_ _V__E__V__ K__F__D__G __C__E__D_ _N__V__M__ E__Y__V__V      133

3121 atggacacat atatacctgc aggacaacga ctcatgggta aaagtaacta gttccgtaga
     __W__T__H_ _I__Y__L__ Q__D__N__D __S__W__V_ _K__V__T__ S__S__V__D      153

3181 tgccaagggc atatattata catgtggaca atttaaaaca tattatgTAA attttAATAA
     __A__K__G_ _I__Y__Y__ T__C__G__Q __F__K__T_ _Y__Y__V__ N__F__N__K      173
                                              E4 orf start ->   -> signal

3241 Agaggcacaa aagtATGgta gtaccaatca ttgggaagta tgttatggca gcacagttat
     __E__A__Q_ _K__Y__G__ S__T__N__H __W__E__V_ _C__Y__G__ S__T__V__I      193
                  _M__V_ _V__P__I__ I__G__K__Y __V__M__A_ _A__Q__L__      15
           E4 cds ->

3301 atgttctcct gcatctgtat ctagcactgt acgagaagta tccattgctg aacctactac
     __C__S__P_ _A__S__V__ S__S__T__V __R__E__V_ _S__I__A__ E__P__T__T      213
     Y__V__L__L __H__L__Y_ _L__A__L__ Y__E__K__Y __P__L__L_ _N__L__L__      35
```

```
3361 atacaccccc gcacagacca ccgcccctac agtgtccgcc tgcaccacgg aagacggcgt
     __Y__T__P_ _A__Q__T__ T__A__P__T __V__S__A_ _C__T__T__ E__D__G__V      233
     H__T__P__P __H__R__P_ _P__P__L__ Q__C__P__P __A__P__R_ _K__T__A__      55

3421 gtcggcgccg cctaggaagc gagcacgtgg accgtccact aacaacaccc tgtgtgtggc
     __S__A__P_ _P__R__K__ R__A__R__G __P__S__T_ _N__N__T__ L__C__V__A      253
     C__R__R__R __L__G__S_ _E__H__V__ D__R__P__L __T__T__P_ _C__V__W__      75

3481 caacatcaga tccgtggaca gtacaatcaa caacatcgtc actgacaatt acaacaagca
     __N__I__R_ _S__V__D__ S__T__I__N __N__I__V_ _T__D__N__ Y__N__K__H      273
     P__T__S__D __P__W__T_ _V__Q__S__ T__T__S__S __L__T__I_ _T__T__S__      95

3541 ccaaagaagg aacaactgtc acagtgcagc tacgcctaTA Gtgcaactgc aaggtgattc
     __Q__R__R_ _N__N__C__ H__S__A__A __T__P__I_ _V__Q__L__ Q__G__D__S      293
     T__K__E__G __T__T__V_ _T__V__Q__ L__R__L__$ _                         108
                                                  <- E4 end

3601 caattgttta aaatgtttta gatatagact gaatgacaaa tataaacatt tgtttgaatt
     __N__C__L_ _K__C__F__ R__Y__R__L __N__D__K_ _Y__K__H__ L__F__E__L      313

3661 agcatcttca acgtggcatt gggcctcacc tgaggcacca ataaaaatg caattgtaac
     __A__S__S_ _T__W__H__ W__A__S__P __E__A__P_ _H__K__N__ A__I__V__T      333

3721 attaacatat agcagtgagg aacaacgtca gcaattttta aacagtgtaa aaataccacc
     __L__T__Y_ _S__S__E__ E__Q__R__Q __Q__F__L_ _N__S__V__ K__I__P__P      353

3781 caccattagg cataaggtgg ggtttatgtc attacattta ttgTAAccat tacacctgta
     __T__I__R_ _H__K__V__ G__F__M__S __L__H__L_ _L__$_                    367
                                                  <- E2 end

3841 tatatgtata tgtgtacaTA Acatacgtgt ATGgaggtag tgcctgtaca aattgctgca
                                     _M__E__V__ V__P__V__Q __I__A__A_      10
              E5A orf start ->
                              E5A cds ->

3901 gcaacaacta caacattgat attgcctgtt gttattgcat ttgcagtatg tattcttagt
     _A__T__T__ T__T__L__I __L__P__V_ _V__I__A__ F__A__V__C __I__L__S_      30

3961 attgtactta taatattaat atcTGAtttt gtagtatata catctgtgct ggtactaaca
     _I__V__L__ I__I__L__I __S__D__F_ _V__V__Y__ T__S__V__L __V__L__T_      50
                   E5B orf start ->

4021 cttcttttat atttgctttt gtggctttta ttaacaaccc ctttgcaatt ctttttacta
     _L__L__L__ Y__L__L__L __W__L__L_ _L__T__T__ P__L__Q__F __F__L__L_      70

4081 acactgtgtg tgtgctattt tcctgccttt tatatacaca tatacattgt gcaaacgcaa
     _T__L__C__ V__C__Y__F __P__A__F_ _Y__I__H__ I__Y__I__V __Q__T__Q_      90

4141 caaTAATGgt gatgttaacc tgtcacttaa atgatggtga tacatggttg tttctgtggt
     _Q__$_                                                                91
          _M__V __M__L__T_ _C__H__L__ N__D__G__D __T__W__L_ _F__L__W__      18
  E5B cds ->
        <- E5A end

4201 tgtttactgc atttgttgta gctgtacttg gattgttgtt actacattac aggctgtac
     L__F__T__A __F__V__V_ _A__V__L__ G__L__L__L __L__H__Y_ _R__A__V__      38


❄    ❄    ❄    ❄    Bases 4261 to 7440 not shown.    ❄    ❄    ❄    ❄


7441 atgtatgttt ttgtgcAATA AAcaattatt atgtgtgtcc tgttacaccc agtgactaag
               signal ->   <-

7501 ttgtgtttag cacgcgccgt ttgtgttgcc ttcatattat attatagata tttgtaatat

7561 acctatacta tgttaccccc ccccacttgc aACCGTTTTC GGTtgccctt acatacactt
                                 ▲         -> E2 bind

7621 acctcaaatt tgttataacg tgtttggtac taatcccata tgttgtgtgc caaggtacat
```

```
7681 atggcctgc caagtatct gccaacaaca cacctggcca gggcgcggta ttgcatgact
                 ▲   ▲                        ▲

7741 aatgtacAAT AAAcctgtcg gtttgtacaa tgttgtggat tgcagccaaa ggttaaaagc
       ▲
     signal ->

7801 atttttggct tctagctgaa cattttgta ccttagtat attatgcaca atacccacaa

7861 aatgagtaac ctaaggtcac acacctgcaA CCGGTTTCGG Ttacccacac cctacatatt
                                  -> E2 bind

7921 tccttcttat a
```

E6 nuc

```
                 11344
                 38836
                 76016
HPV11            TACGG   102-554
HPV11P           CCTAA   102-554
HPV11VC          CCTAA   102-554
```

E6 aa

```
                 00011
                 12912
                 29302
HPV11            SITPG
HPV11P           *L**E
HPV11VC          *L**E
```

E7 nuc

```
                 6
                 6
                 2
HPV11            G   530-826
HPV11P           T   530-826
HPV11VC          T   530-826
```

E7 aa

```
                 4
                 5
HPV11            A
HPV11P           S
HPV11VC          S
```

E1 nuc

```
                 11122
                 37715
                 18878
                 23400
HPV11            CCGCA   832-2781
HPV11P           -GC--   832-2781
HPV11VC          GGCTG   832-2781
```

E1 aa

```
                 13^345
                 61148
                 18873
HPV11            QRRLR
HPV11P           -AA--
HPV11VC          EAA**
```

E2 nuc

```
                2233333
                8844667
                8838242
                4867657
HPV11           CTGCAAA   2723-3826
HPV11P          ---T---   2723-3826
HPV11VC         TCATCGC   2723-3826
6c                ATCG-   3137-3826
```

E2 aa

```
                0022333
                5535003
                4685285
HPV11           ILRIRKT
HPV11P          ---*---
HPV11VC         *****R*
6c                ***R-
```

E4 nuc

```
                33
                44
                38
                67
HPV11           GC   3255-3581
HPV11P          -T   3255-3581
HPV11VC         AT   3255-3581
6c              AT   3255-3581
```

E4 aa

```
                67
                18
HPV11           GS
HPV11P          -L
HPV11VC         EL
6c              EL
```

noncoding

```
                3
                8
                3
                2
HPV11           A   3827-3870
HPV11P          -   3827-3870
HPV11VC         -   3827-3870
6c              G   3827-3870
```

E5A nuc

```
                 33
                 99
                 59
                 21
HPV11            AG   3871-4146
HPV11P           --   3871-4146
HPV11VC          TC   3871-4146
6c               -C   3871-4146
```

E5A aa

```
                 24
                 81
HPV11            IV
HPV11P           --
HPV11VC          FL
6c               -L
```

E5B nuc

```
                 4
                 2
                 5
                 3
HPV11            G   4146-4370
HPV11P           -   4146-4370
HPV11VC          A   4146-4370
6c               -   4146-4370
```

E5B aa

```
                 3
                 6
HPV11            R
HPV11P           -
HPV11VC          *
6c               -
```

LCR

```
                   ii          insertion      insertion      dii  insertion
                 77777777777777777777777777777777777777777777777777
                 45555666666666666666666666666666666666667777777777
                 70488224466889999999999999999999999999911144444444
                 99755676878354444444444444447777777777909955555555
HPV11            CTT..AATTGTTC.........................TTA..........   7277-7931
B11-7                 --------.........................---GC........   7625-7874
B11-15D2              ------CT.........................---GC........   7625-7874
G11-79                ----A---.........................-C-GC........   7625-7874
HPV11P           ---..--------.........................---GC........   7277-7931
HPV11VC          T.CCC--------.........................---GC........   7277-7931
J11-2                 --------.........................---GC........   7625-7874
NY11-9                ---C----.........................-C-GC........   7625-7874
NY11-14               --------.........................---GC........   7625-7874
PPH11A           ---..--------.........................---GC........   7277-7931
S11-27                -----C--.........................-C-GC........   7625-7874
G11-30                C-A-----CAACACACCTGCCAACATAT..TGCC-GCGGTTGTGT   7625-7874
SN11-1                CGA-----................TATATAATG--.GCGGTTGTGT   7625-7874
```

LCR cont'd.

insertion d

```
                77777777777777
                77777777778888
                44444446770011
                55555559384879
HPV11           .......ATGTGTA   7277-7931
B11-7           .........---C---   7625-7874
B11-15D2        .......-------   7625-7874
G11-79          .......-------   7625-7874
HPV11P          .......-------   7277-7931
HPV11VC         .......-------   7277-7931
J11-2           .......------C   7625-7874
NY11-9          .......-------   7625-7874
NY11-14         .......G---.C-   7625-7874
PPH11A          .......-------   7277-7931
S11-27          .......-------   7625-7874
G11-30          TATCTCC-GT----   7625-7874
SN11-1          TATCTCC--TC---   7625-7874
```

```
LOCUS       HPV16R       7906 bp ds-DNA   circular  VRL       15-APR-1991
DEFINITION  Human papillomavirus type 16 (HPV16), complete genome.
ACCESSION   <not in GenBank>
KEYWORDS    circular; complete genome.
SOURCE      Papilloma virus type 16 DNA recovered from a human invasive
            cervical carcinoma.
  ORGANISM  Human papillomavirus type 16
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
COMMENT     Naturally occurring variants of HPV16R have been observed throughout
            the genome.
            Reference: Bavin,P.J. Walker,P.G. and Emery,V.C.
                J. Med. Virol. 39, 267-72 (1993)
                Sequences constructed by editing reference sequence
                according to figures in article, with corrections from authors.
            Reference: Chan,S.Y., Ho,L., Ong,C.K., Chow,V., Drescher,B., et al.,
                J. Virol. 66, 2057-66 (1992)
                Variant Accession Nos. M83776-M83804, M83840
            Reference: Eschle,D., Durst,M., ter Meulen,J., Luande,J., Eberhardt,H.C.,
                Pawlita,M., and Gissmann,L.
                J. Gen. Virol. 73, 1829-32 (1992)
                Sequences constructed by editing reference sequence
                according to figures in article.
            Reference: Fujinaga,Y., Okazawa,K., Nishikawa,A., Yamakawa,Y., Fukushima,M.,
                Kato,I., and Fujinaga,K.
                Virus Genes 9, 85-92 (1994)
                Sequences constructed by editing reference sequence according
                to figures in article.
            Reference: Haegert,D.G., Younghusband,B.H., Galutira,D.F., and Egger,K.N.
                Unpublished (1995)
                Variant Accession Nos. U33065-U33069, U33119
            Reference: Haegert,D.G., Galutira,D.F., and Younghusband,B.H.
                Unpublished (1995)
                Variant Accession Nos. U14511-U14516
            Reference: Ho,L., Chan,S.Y., Chow,V., Chong,T., Tay,S.K., Villa,L.L.,
                and Bernard,H.U., J. Clin. Microbiol. 29, 1765-72 (1991).
                Variant Accession Nos. M83776-M83778 M83834-M83841
            Reference: Ho,L., Chan,S.Y., Burk,R.D., Das,B.C., Fujinaga,K., et al.,
                J. Virol. 67, 6413-23 (1993)
                Variant Accession Nos. L22661-L22674
            Reference: Icenogle,J.P., Sathya,P., Miller,D.L., Tucker,R.A., and Rawls,W.E.
                Virology 184, 101-7 (1991)
                Sequences constructed by editing reference sequence according to figures
                in article.
            Reference: Icenogle,J.P., Laga,M., Miller,D., Manoka,A.T., Tucker,R.A.,
                and Reeves,W.C.
                J. Infect. Dis. 1992, 166 (1210-6)
                Sequences constructed by editing reference sequence according to figures
                in article.
            Reference: Icenogle,J.P., Clancy,K.A., and Lin,S.Y.
                Virology 214, 664-9 (1995)
                Variant Accession Nos. K02718 U37217 U37410.
            Reference: Kirnbauer,R., Taub,J., Greenstone,H., Roden,R., et al.
                J. Virol. 67, 6929-36 (1993)
                Sequences constructed by editing reference sequence according
                to figures in article.
            Reference: Meneguzzi,G.
                Pharmaceutical Composition Useful in the Prevention or Treatment of
                Papillomavirus-Induced Tumours, Patent: WO 9010459-A 21 20-SEP-1990
                Variant Accession No. A00446.
            Reference: Pushko,P., Sasagawa,T., Cuzick,J., and Crawford,L.
                J. Gen. Virol. 75, 911-6 (1994)
                Sequences constructed by editing reference sequence (HPV16R) according
                to figures in article, with corrections to sequence numbering
```

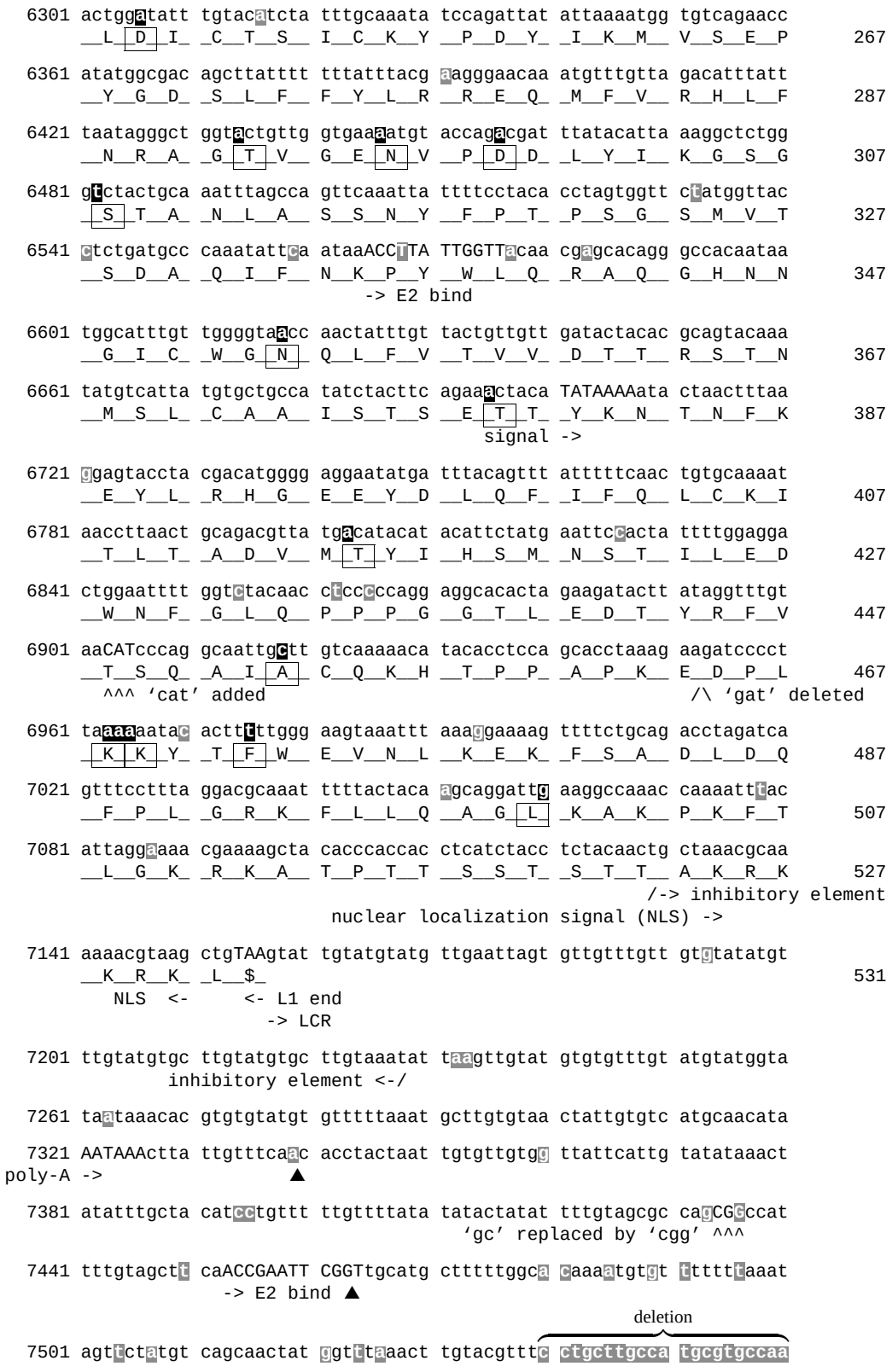
```
            from authors.
          Reference: Smits,H.L., Traanberg,K.F., Krul,M.R., Prussia,P.R., et al.
            J. Gen. Virol. 75, 2457-62 (1994)
            Sequences constructed by editing reference sequence (HPV16R) according
            to figures in article, with corrections to sequence numbering
            from authors.
          Reference: Xi,L.F., Demers,G.W., Koutsky,L.A., Kiviat,N.B., et al.
            J. Infect. Dis. 172, 747-55 (1995)
            Sequences constructed by editing reference sequence (HPV16R) according
            to figures in article.
          Reference: Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.-C.M.,
            Hildesheim,A., Rush,B.B., and Jenison,S.A., J. Virol. 69, 7743-53 (1995)
            Variant Accession Nos. U34078-34193
FEATURES  Location/Qualifiers
     CDS             83..559
                     /note="E6 ORF from 65 to 559"
                     /product="transforming protein"
                     /gene="E6"
                     /note="putative"
                     /codon_start=1
                     /translation="MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQL
                     LRREVYDFAFRDLCIVYRDGNPYAVCDKCLKFYSKISEYRHYCYSLYGTTLEQQYNKP
                     LCDLLIRCINCQKPLCPEEKQRHLDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL"
     CDS             562..858
                     /note="E7 ORF from 544 to 858"
                     /product="transforming protein"
                     /gene="E7"
                     /note="putative"
                     /codon_start=1
                     /translation="MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQ
                     AEPDRAHYNIVTFCCKCDSTLRLCVQSTHVDIRTLEDLLMGTLGIVCPICSQKP"
     CDS             3850..4101
                     /note="E5 ORF from 3808 to 4101"
                     /gene="E5"
                     /note="putative"
                     /codon_start=1
                     /translation="MTNLDTASTTLLACFLLCFCVLLCVCLLIRPLLLSVSTYTSLII
                     LVLLLWITAASAFRCFIVYIIFVYIPLFLIHTHARFLIT"
     CDS             4237..5658
                     /note="L2 ORF from 4135 to 5658"
                     /product="minor capsid protein"
                     /gene="L2"
                     /note="putative"
                     /codon_start=1
                     /translation="MRHKRSAKRTKRASATQLYKTCKQAGTCPPDIIPKVEGKTIADQ
                     ILQYGSMGVFFGGLGIGTGSGTGGRTGYIPLGTRPPTATDTLAPVRPPLTVDPVGPSD
                     PSIVSLVEETSFIDAGAPTSVPSIPPDVSGFSITTSTDTTPAILDINNTVTTVTTHNN
                     PTFTDPSVLQPPTPAETGGHFTLSSSTISTHNYEEIPMDTFIVSTNPNTVTSSTPIPG
                     SRPVARLGLYSRTTQQVKVVDPAFVTTPTKLITYDNPAYEGIDVDNTLYFSSNDNSIN
                     IAPDPDFLDIVALHRPALTSRRTGIRYSRIGNKQTLRTRSGKSIGAKVHYYYDLSTID
                     PAEEIELQTITPSTYTTTSHAASPTSINNGLYDIYADDFITDTSTTPVPSVPSTSLSG
                     YIPANTTIPFGGAYNIPLVSGPDIPINITDQAPSLIPIVPGSPQYTIIADAGDFYLHP
                     SYYMLRKRRKRLPYFFSDVSLAA"
     CDS             5639..7156
                     /note="L1 ORF from 5528 to 7156"
                     /product="major capsid protein"
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /translation="MSLWLPSEATVYLPPVPVSKVVSTDEYVARTNIYYHAGTSRLLA
                     VGHPYFPIKKPNNNKILVPKVSGLQYRVFRIHLPDPNKFGFPDTSFYNPDTQRLVWAC
                     VGVEVGRGQPLGVGISGHPLLNKLDDTENASAYAANAGVDNRECISMDYKQTQLCLIG
                     CKPPIGEHWGKGSPCTNVAVNPGDCPPLELINTVIQDGDMVDTGFGAMDFTTLQANKS
                     EVPLDICTSICKYPDYIKMVSEPYGDSLFFYLRREQMFVRHLFNRAGTVGENVPDDLY
```

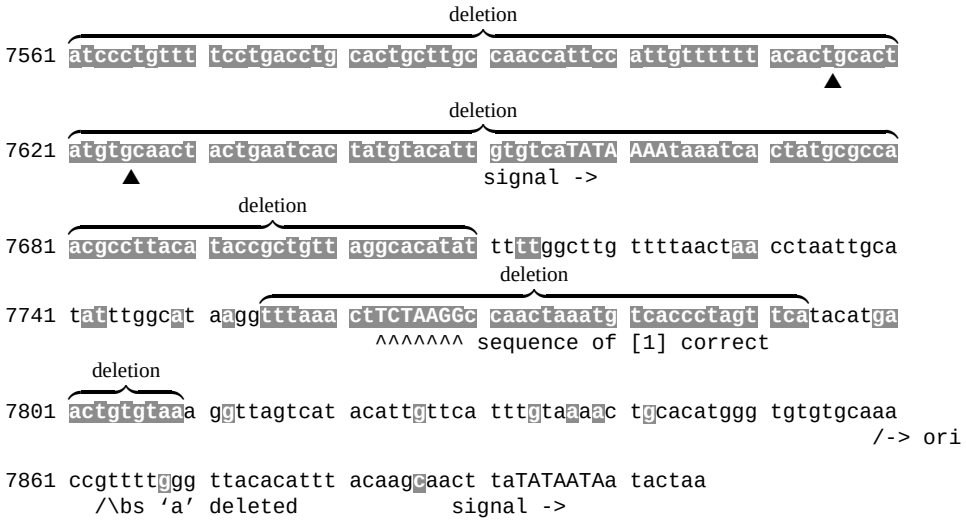
```
                    IKGSGSTANLASSNYFPTPSGSMVTSDAQIFNKPYWLQRAQGHNNGICWGNQLFVTVV
                    DTTRSTNMSLCAAISTSETTYKNTNFKEYLRHGEEYDLQFIFQLCKITLTADVMTYIH
                    SMNSTILEDWNFGLQPPPGGTLEDTYRFVTSQAIACQKHTPPAPKEDPLKKYTFWEVN
                    LKEKFSADLDQFPLGRKFLLQAGLKAKPKFTLGKRKATPTTSSTSTTAKRKKRKL"
BASE COUNT     2599 a   1377 c   1511 g   2419 t

        1 actacaataa ttcatgTATA AAAptaaggg cgtaACCGAA ATCGGTtgaA CCGAAACCGG
                    signal ->                   -> E2 bind      -> E2 bind

       61 TTAGTATAAA Agcagacatt ttatgcacca aaagagaact gcaATGtttc aggacccaca
                                  _M__H__Q __K__R__T_ _A__M__F__ Q__D__P__Q      13
    ori <-/                                    E6 cds ->
       signal ->                              |-> mRNA start site from
            <- LCR                               P97 promoter
  E6 orf start->

      121 ggagcgACCC AGAAAGTTac cacagttatg cacagagctg caaacaacta tacatgatat
          __E__R__P_ _R__K__L__ P__Q__L__C __T__E__L_ _Q__T__T__ I__H__D__I      33
                -> E2 bind

      181 aattattagaa tgtgtgtact gcaagcaaca gttactgcga cgtgagGTat atgactttgc
          _I__L__E_ _C__V__Y__ C__K__Q__Q __L__L__R_ _R__E__V__ Y__D__F__A      53
                                                    5 sj /\ (E6*transcripts)

      241 ttttcgggat ttatgcatag tatatagaga tgggaatcca tatgctgtat gtgattaaatg
          __F__R__D_ _L__C__I__ V__Y__R__D__G__N__P_ _Y__A__V__ C__D__K__C      73

      301 tttaaagttt tattctaaaa ttagtgagta tagacattat tgttatagtt tgtatggaac
          __L__K__F__ _Y__S__K__ I__S__E__Y __R__H__Y_ _C__Y__S__ L__Y__G__T      93

      361 aacattagaa cagcaataca acaaaccgtt gtgtgatttg ttaattAGgt gtattaactg
          __T__L__E_ _Q__Q__Y__ N__K__P__L __C__D__L_ _L__I__R__ C__I__N__C      113
                                                             /\ 3 sj (E6*I)

      421 tcaaaagcca ctgtgtcctg aagaaaagca aagacatctg gacaaaaagc aaagattcca
          __Q__K__P_ _L__C__P__ E__E__K__Q __R__H__L_ _D__K__K__ Q__R__F__H      133

      481 taatataagg ggtcggtgga ccggtcgatg tatgtcttgt tgcAGatcat caagaacacg
          __N__I__R_ _G__R__W__ T__G__R__C __M__S__C_ _C__R__S__ S__R__T__R      153
                                                        /\ 3 sj (E6*II)

      541 TAGagaaacc cagctgTAAt cATGcatgga gatacaccta cattgcatga atatatgtta
                                 _M__H__G_ _D__T__P__ T__L__H__E __Y__M__L_      13
          __R__E__T_ _Q__L__$_                                                  158
E7 orf start ->      E6 end <-   -> E7 cds

      601 gatttgcaac cagagacaac tgatctctac tgttatgagc aattaaatga cagctcagag
          _D__L__Q__ P__E__T__T __D__L__Y_ _C__Y__E__ Q__L__N__D __S__S__E_      33

      661 gaggaggatg aaatagatgg tccagctgga caagcagaac cggacagagc ccattacaat
          _E__E__D__ E__I__D__G __P__A__G_ _Q__A__E__ P__D__R__A __H__Y__N_      53

      721 attgtaacct ttgttgcaa gtgtgactct acgcttcggt tgtgcgtaca aagcacacac
          _I__V__T__ F__C__C__K __C__D__S_ _T__L__R__ L__C__V__Q __S__T__H_      73

      781 gtagacattc gtacattgga agacctgtta atgggcacac taggaatggt gtgccccatc
          _V__D__I__ R__T__L__E __D__L__L_ _M__G__T__ L__G__I__V __C__P__I_      93

      841 tgttctcaga aaccaTAAtc taccATGgct gatcctgcag GTaccaatgg ggaagagggt
          _C__S__Q__ K__P__$_       _M__A_ _D__P__A__ G__T__N__G __E__E__G_      98/12
                E1 orf start ->     -> E1 cds  5 sj /\
                          <- E7 end
```

❄ ❄ ❄ ❄ Bases 900 to 3780 not shown. ❄ ❄ ❄ ❄

```
3781 acgtgaccaa tttttgtctc aagtTAAaat accaaaaact attacagtgt ctactggatt
     __R__D__Q_ _F__L__S__ Q__V__K__I __P__K__T_ _I__T__V__ S__T__G__F      362
                   E5 orf start ->
                   NH2 terminus unknown

3841 tatgtctatA TGAcaaaTct tgatactgca tccacaacat tactggcgtg ctttttgctt
               _ M__T__N__L __D__T__A__ _S__T__T__ L__L__A__C __F__L__L_      17
     __M__S__I_ _$_                                                         365
              -> E5 cds
                 <- E2 end

3901 tgcttttgtg tgcttttgtg tgtctgccta ttaatacgtc cgctgctttt gtctgtgtct
     _C__F__C__ V__L__L__C __V__C__L_ _L__I__R__ P__L__L__L __S__V__S_      37
          ^ additional ‘t’

3961 acatacacat cattaataat attggtatta ctattgtgga taacagcagc ctctgcgttt
     _T__Y__T__ S__L__I__ I__L__V__L_ _L__L__W__ I__T__A__A __S__A__F_      57

4021 aggtgtttta ttgtatatat tatattgtgt tatataccat tatttttaat acatacacat
     _R__C__F__ I__V__Y__ I__I__F__V_ _Y__I__P__ L__F__L__I __H__T__H_      77

4081 gcacgcttTt taattacaTA Atgtatatgt acataatgta attgttacat aTAAttgttg
     _A__R__F__ L__I__T__$ _                                                83
                           <- E5 end                L2 orf start ->

4141 tataccataa cttactattt tttctttttt attttcatat ataatttttt tttttgtttg

4201 tttgtttgtt ttttAATAAA ctgttattac ttaacaATGc gacacaaacg ttctgcaaaa
                                          _M__ R__H__K__R __S__A__K_      8
early poly-A signal ->              L2 cds ->

4261 cgcacaaaac gtgcatcggc tacccaactt TATAAAAcat gcaaacaggc aggtacatgt
     _R__T__K__ R__A__S__A __T__Q__L_ _Y__K__T__ C__K__Q__A __G__T__C_      28
                               signal ->

4321 ccacctgaca ttatacctaa ggttgaaggc aaaactattg ctgaTcaaat attacaatat
     _P__P__D__ I__I__P__K __V__E__G_ _K__T__I__ A__D__Q__I __L__Q__Y_      48
                                                 ^ ‘a’ replaced by ‘t’

4381 ggaagtatgg gtgtattttt tggtggttta ggaattggaa cagggtcggg tacaggcgga
     _G__S__M__ G__V__F__F __G__G__L_ _G__I__G__ T__G__S__G __T__G__G_      68

4441 cgcactgggt attattccatt gggaacaagg cctcccacag ctacagatac acttgctcct
     _R__T__G__ Y__I__P__L __G__T__R_ _P__P__T__ A__T__D__T __L__A__P_      88

4501 gtaagacccc cttaacagt agatcctgtg ggcccttctg atcctctat agtttcttta
     _V__R__P__ P__L__T__V __D__P__V_ _G__P__S__ D__P__S__I __V__S__L_      108

4561 gtggaagaaa ctagttttat tgatgctggt gcaccaacat ctgtaccttc cattccccca
     _V__E__E__ T__S__F__I __D__A__G_ _A__P__T__ S__V__P__S __I__P__P_      128

4621 gatgtatcag gatttagtat tactacttca actgatacca cacctgctat attagatatt
     _D__V__S__ G__F__S__I __T__T__S_ _T__D__T__ T__P__A__I __L__D__I_      148

4681 aataatactg ttactactgt tactacacat aataatccca ctttcactga cccatctgta
     _N__N__T__ V__T__T__V __T__T__H_ _N__N__P__ T__F__T__D __P__S__V_      168

4741 ttgcagcctc caacacctgc agaaactgga gggcatttta cactttcatc atccactatt
     _L__Q__P__ P__T__P__A __E__T__G_ _G__H__F__ T__L__S__S __S__T__I_      188

4801 agtacacata attatgaaga aattcctatg gatacattta ttgttagcac aaacctaac
     _S__T__H__ N__Y__E__E __I__P__M_ _D__T__F__ I__V__S__T __N__P__N_      208

4861 acagtaacta gtagcacacc cataccaggg tctcgcccag tggcacgcct aggattatat
     _T__V__T__ S__S__T__P __I__P__G_ _S__R__P__ V__A__R__L __G__L__Y_      228

4921 agtcgcacaa cacaacaggt taaagttgta gaccctgctt ttgtaaccac tcccactaaa
     _S__R__T__ T__Q__Q__V __K__V__V_ _D__P__A__ F__V__T__T __P__T__K_      248

4981 cttattacat atgataatcc tgcatatgaa ggtatagatg tggataatac attatatttt
     _L__I__T__ Y__D__N__P __A__Y__E_ _G__I__D__ V__D__N__T __L__Y__F_      268
```

```
 5041 tctagtaatg ataatagtat taatatagct ccagatcctg acttttgga tatagttgct
      _S_S_N__ D__N__S__I __N__I__A_ _P__D__P__ D__F__L__D __I__V__A_      288

 5101 ttacataggc cagcattaac ctctaggcgt actggcatta ggtagagtag aattggtaat
      _L__H__R__ P__A__L__T __S__R__R_ _T__G__I__ R__Y__S__R __I__G__N_      308

 5161 aaacaaacac tacgtactcg tagtggaaaa tctataggtg ctaaggtaca ttattattat
      _K__Q__T__ L__R__T__R __S__G__K_ _S__I__G__ A__K__V__H __Y__Y__Y_      328

 5221 gatttaagta ctattgatcc tgcagaagaa atagaattac aaactataac accttctaca
      _D__L__S__ T__I__D__P __A__E__E_ _I__E__L__ Q__T__I__T __P__S__T_      348

 5281 tatactacca cttcacatgc agcctcacct acttctatta ataatggatt atatgatatt
      _Y__T__T__ T__S__H__A __A__S__P_ _T__S__I__ N__N__G__L __Y__D__I_      368

 5341 tatgcagatg actttattac agatacttct acaaccccgg taccatctgt accctctaca
      _Y__A__D__ D__F__I__T __D__T__S_ _T__T__P__ V__P__S__V __P__S__T_      388

 5401 tcttttatcag gttatattcc tgcaaatata acaattcctt ttggtggtgc atacaatatt
      _S__L__S__ G__Y__I__P __A__N__T_ _T__I__P__ F__G__G__A __Y__N__I_      408

 5461 cctttagtat caggtcctga tataccaatt aatataactg accaagctcc ttcattaatt
      _P__L__V__ S__G__P__D __I__P__I_ _N__I__T__ D__Q__A__P __S__L__I_      428

 5521 cctaTAGttc cagggtctcc acaatatata attattgctg atgcaggtga cttttattta
      _P__I__V__ P__G__S__P __Q__Y__T_ _I__I__A__ D__A__G__D __F__Y__L_      448
L1 orf start ->

 5581 catcctagtt attacatgtt acgaaaacga cgtaaacgtt taccatattt tttttcAGAT
      _H__P__S__ Y__Y__M__L __R__K__R_ _R__K__R__ L__P__Y__F __F__S__D_      468
                                                                 _M         1
                                                               /\ 3 sj
                                                          L1 cds ->

 5641 Gtctctttgg ctgccTAGtg aggccactgt ctacttgcct cctgtcccag tatctaaggt
      _V__S__L__ A__A__$_                                               473
      __S__L__W_ _L__P__S__ E__A__T__V __Y__L__P_ _P__V__P__ V__S__K__V      21
                        <- L2 end

 5701 tgtaagcacg gatgaatatg ttgcacgcac aaacatatat tatcatgcag gaacatccag
      __V__S__T_ _D__E__Y__ V__A__R__T __N__I__Y_ _Y__H__A__ G__T__S__R      67

 5761 actacttgca gttggacatc cctattttcc tattaaaaaa cctaacaata acaaaatatt
      __L__L__A_ _V__G__H__ P__Y__F__P __I__K__K_ _P__N__N__ N__K__I__L      87

 5821 agttcctaaa gtatcaggat tacaatacag ggtatttaga atacatttac ctgaccccaa
      __V__P__K_ _V__S__G__ L__Q__Y__R __V__F__R_ _I__H__L__ P__D__P__N      107

 5881 taagtttggt tttcctgaca cctcatttta taatccagat acacagcggc tggtttgggc
      _K__F__G_ _F__P__D__ T__S__F__Y __N__P__D_ _T__Q__R__ L__V__W__A      127

 5941 ctgtgtaggt gttgaggtag gtcgtggtca gccattaggt gtgggcatta gtggccatcc
      __C__V__G_ _V__E__V__ G__R__G__Q __P__L__G_ _V__G__I__ S__G__H__P      147

 6001 tttattaaat aaattggatg acacagaaaa tgctagtgct tatgcagcaa atgcaggtgt
      __L__L__N_ _K__L__D__ D__T__E__N __A__S__A_ _Y__A__A__ N__A__G__V      167

 6061 ggataatagg gaatgtatat ctatggatta caaacaaaca caattgtgtt taattggttg
      __D__N__R_ _E__C__I__ S__M__D__Y __K__Q__T_ _Q__L__C__ L__I__G__C      187

 6121 caaaccacct ataggggaac actggggcaa aggatcccca tgtacaatg ttgcagtaaa
      __K__P__P_ _I__G__E__ H__W__G__K __G__S__P_ _C__T__N__ V__A__V__N      207

 6181 tccaggtgat tgtccACCAT TAGAGTTaat aaacacagtt attcaggatg gtgatatggt
      _P__G__D_ _C__P__P__ L__E__L__I __N__T__V_ _I__Q__D__ G__D__M__V      227
                     -> E2 bind

 6241 tGatactggc tttggtgcta tggactttac tacattacag gctaacaaaa gtgaagttcc
      __D__T__G_ _F__G__A__ M__D__F__T __T__L__Q_ _A__N__K__ S__E__V__P      247
       ^ wild-type has 'g', original clone had 'c'
```


6301 actggatatt tgtacatcta tttgcaaata tccagattat attaaaatgg tgtcagaacc
__L__D__I__ _C__T__S__ I__C__K__Y __P__D__Y_ _I__K__M__ V__S__E__P 267
6361 atatggcgac agcttatttt tttatttacg aagggaacaa atgtttgtta gacatttatt
__Y__G__D_ _S__L__F__ F__Y__L__R __R__E__Q_ _M__F__V__ R__H__L__F 287
6421 taatagggct ggtactgttg gtgaaatatgt accagacgat ttatacatta aaggctctgg
__N__R__A_ _G__T__V__ G__E__N__V __P__D__D_ _L__Y__I__ K__G__S__G 307
6481 gtctactgca aatttagcca gttcaaatta ttttcctaca cctagtggtt ctatggttac
_S__T__A_ _N__L__A__ S__S__N__Y __F__P__T_ _P__S__G__ S__M__V__T 327
6541 ctctgatgcc caaatattca ataaACCTTA TTGGTTacaa cgagcacagg gccacaataa
__S__D__A_ _Q__I__F__ N__K__P__Y __W__L__Q_ _R__A__Q__ G__H__N__N 347
-> E2 bind
6601 tggcatttgt tggggtaacc aactatttgt tactgttgtt gatactacac gcagtacaaa
__G__I__C_ _W__G__N__ Q__L__F__V __T__V__V_ _D__T__T__ R__S__T__N 367
6661 tatgtcatta tgtgctgcca tatctacttc agaaactaca TATAAAAata ctaactttaa
__M__S__L_ _C__A__A__ I__S__T__S __E__T__T_ _Y__K__N__ T__N__F__K 387
signal ->
6721 ggagtaccta cgacatgggg aggaatatga tttacagttt atttttcaac tgtgcaaaat
__E__Y__L_ _R__H__G__ E__E__Y__D __L__Q__F_ _I__F__Q__ L__C__K__I 407
6781 aaccttaact gcagacgtta tgacatacat acattctatg aattcgacta ttttggagga
__T__L__T_ _A__D__V__ M__T__Y__I __H__S__M_ _N__S__T__ I__L__E__D 427
6841 ctggaatttt ggtgtacaac ctccgccagg aggcacacta gaagatactt ataggtttgt
__W__N__F_ _G__L__Q__ P__P__P__G __G__T__L_ _E__D__T__ Y__R__F__V 447
6901 aaCATcccag gcaattgctt gtcaaaaaca tacacctcca gcacctaaag aagatcccct
__T__S__Q_ _A__I__A__ C__Q__K__H __T__P__P_ _A__P__K__ E__D__P__L 467
^^^ 'cat' added
/\ 'gat' deleted
6961 taaaaaatac acttttttggg aagtaaattt aaaggaaaag ttttctgcag acctagatca
__K__K__Y_ _T__F__W__ E__V__N__L __K__E__K_ _F__S__A__ D__L__D__Q 487
7021 gtttccttta ggacgcaaat ttttactaca agcaggattg aaggccaaac caaaattcac
__F__P__L_ _G__R__K__ F__L__L__Q __A__G__L__ _K__A__K__ P__K__F__T 507
7081 attaggaaaa cgaaaagcta cacccaccac ctcatctacc tctacaactg ctaaacgcaa
__L__G__K_ _R__K__A__ T__P__T__T __S__S__T_ _S__T__T__ A__K__R__K 527
/-> inhibitory element
nuclear localization signal (NLS) ->
7141 aaaacgtaag ctgTAAgtat tgtatgtatg ttgaattagt gttgtttgtt gtgtatatgt
__K__R__K_ _L__$_ 531
NLS <- <- L1 end
-> LCR
7201 ttgtatgtgc ttgtatgtgc ttgtaaatat taagttgtat gtgtgtttgt atgtatggta
inhibitory element <-/
7261 tatataaacac gtgtgtatgt gtttttaaat gcttgtgtaa ctattgtgtc atgcaacata
7321 AATAAActta ttgtttcaac acctactaat tgtgttgtgg ttattcattg tatataaact
poly-A ->
▲
7381 atatttgcta catccgtgttt ttgttttata tatactatat tttgtagcgc cagCGccat
'gc' replaced by 'cgg' ^^^
7441 tttgtagctt caACCGAATT CGGTtgcatg ctttttggca caaaatgtgt ttttttaaat
-> E2 bind ▲
deletion
7501 agttctatgt cagcaactat ggttttaaact tgtacgtttc ctgcttgcca tgcgtgccaa



noncoding

HPV16R	TC	1-82
HDS117A.04/09/92	--	1-82
HDS117A.08/12/92	--	1-82
HDS84C.03/26/90	--	1-82
HDS84C.07/23/90	--	1-82
HDS138.05/01/91	--	1-82
HDS106C.05/08/90	--	1-82
HDS106A.02/13/90	-T	1-82
HDS106A.05/08/90	-T	1-82
HDS84A.03/26/90	C-	1-82
HDS84A.07/23/90	C-	1-82
HDS84B.03/26/90	C-	1-82
HDS84D.07/23/90	C-	1-82
HDS106B.02/13/90	C-	1-82
HDS117B.04/09/92	C-	1-82

E6 nuc

```
                 11111111122222233334455
                 03344778856688913560413
                 91235682367969505043562
HPV16R           TAGCGGTATCGGTATTCTAAACA   83-559
OR.5110          -----------------------   104-559
OR.4724          -----------------------   104-559
NM.T455          -----------------------   104-559
OR.4997          ----------AA-----G-----   104-559
OR.6311          ---------T-------G-----   104-559
O356             ----------------G-----   104-559
OR.8329          C---------------G-----   104-559
OR.8987          C---------------G-----   104-559
GALU1601                 --------G---     236-488
OR.9237          -G--------------G-----   104-559
NM.T197          -G--------------G-----   104-559
O425             -G--------------G-----   104-559
OR.0198          -G--------------G-----   104-559
NM.T446          -G--------------G-----   104-559
O489             -G----------------------   104-559
OR.5428          ------G----------------   104-559
OR.4716          ------G----------------   104-559
OR.2087          ------G----------------   104-559
OR.7574          ------G----------------   104-559
GALU1607                 G-----------     226-497
O204             -------G---------------   104-559
GALU1603                 ------G-----     226-496
OR.6170          C----------------------   104-559
OR.7145          C-TGT-------AG--T--G---   104-559
OR.3473          C-TGT-------AG--T--G---   104-559
OR.3759          C-TGT-------AG--T--G---   104-559
H84              C-TGT-------AG--T--G---   104-559
O424             --CGTT------AG--T------   104-559
OR.7632          --CGT-------AG--T------   104-559
OR.1905          --CGT-------AG--T------   104-559
OR.6106          --CGT-------AG--T------   104-559
OR.7587          --CGT-------AG--T------   104-559
GALU1627                 ---AG--T----     226-498
O561             ---GT-------AG--T------   104-559
O162             -G-GT-------AGG-TG-----   104-559
OR.3136          ----T-------AG--TG-----   104-559
GALU1629                 ---AG--TGT--     226-470
O512             ----T-------AG--TG---GG   104-559
OR.8863          ----T-------AG--TG----G   104-559
NM.9999          ----T-------AG--TG----G   104-559
NM.4094          ----T-------AG--TG----G   104-559
OR.4072          ----T-------AG--TG----G   104-559
OR.7754          ----T-------AG--TG----G   104-559
OR.4541          ----T-------AG--TG----G   104-559
OR.8160          ----T-------AG--TG----G   104-559
NM.T529          ----T-------AG--TG----G   104-559
OR.7908          ----T-------AG--TG----G   104-559
OR.7875          ----T-------AG--TG----G   104-559
OR.4451          ----T-------AG--TG----G   104-559
OR.1783          ----T---G---AG--TG----G   104-559
GALU1649                 ---AG--TG---     226-499
OR.8392          ----T---G---AG--TG----G   104-559
O462             ----T---G---AG--TG----G   104-559
OR.5691          ----T---G---AG--TG----G   104-559
O191             ----T---G---AG--TG--C-G   104-559
```

E6 aa

```
                 0000000000000000000001111
                 0112233335666677899024 5
                 9771122448238916504715 0
HPV16R           FRRQQDDIICRDAVDFHLTLESS
OR.5110          ------------------------
OR.4724          ------------------------
NM.T455          ------------------------
OR.4997          ----------KN-----V-----
OR.6311          ---------*-------V-----
O356             -----------------V-----
OR.8329          *----------------V-----
OR.8987          *----------------V-----
GALU1601                  --------V---
OR.9237          -GG--------------V-----
NM.T197          -GG--------------V-----
O425             -GG--------------V-----
OR.0198          -GG--------------V-----
NM.T446          -GG--------------V-----
O489             -GG--------------------
OR.5428          -----EE----------------
OR.4716          -----EE----------------
OR.2087          -----EE----------------
OR.7574          -----EE----------------
GALU1607                 W-----------
O204             -------VV--------------
GALU1603                 ------L-----
OR.6170          *------------------------
OR.7145          *IIDD-------**--Y--*---
OR.3473          *IIDD-------**--Y--*---
OR.3759          *IIDD-------**--Y--*---
H84              *IIDD-------**--Y--*---
O424             -TTDDYY-----**--Y------
OR.7632          -TTDD-------**--Y------
OR.1905          -TTDD-------**--Y------
OR.6106          -TTDD-------**--Y------
OR.7587          -TTDD-------**--Y------
GALU1627                 ---**--Y----
O561             ---DD-------**--Y------
O162             -GGDD-------**E-YV-----
OR.3136          ---HH-------**--YV-----
GALU1629                 ---**--YV*--
O512             ---HH-------**--YV---C*
OR.8863          ---HH-------**--YV----*
NM.9999          ---HH-------**--YV----*
NM.4094          ---HH-------**--YV----*
OR.4072          ---HH-------**--YV----*
OR.7754          ---HH-------**--YV----*
OR.4541          ---HH-------**--YV----*
OR.8160          ---HH-------**--YV----*
NM.T529          ---HH-------**--YV----*
OR.7908          ---HH-------**--YV----*
OR.7875          ---HH-------**--YV----*
OR.4451          ---HH-------**--YV----*
OR.1783          ---HH--RR---**--YV----*
GALU1649                 ---**--YV---
OR.8392          ---HH--RR---**--YV----*
O462             ---HH--RR---**--YV----*
OR.5691          ---HH--RR---**--YV----*
O191             ---HH--RR---**--YV--D-*
```

E7 nuc

```
                  56666667777788
                  71446670389924
                  56573681290586
HPV16R            CAAAGGTCTTCTTT   562-858
ALABAMA-B24        ----------      582-801
ALASKA-C32         ----------      582-801
ALASKA-C51         ----------      582-801
CASKI              ----------      582-801
MICHIGAN-B8        ----------      582-801
PANAMA-136         ----------      582-801
PANAMA-156         ----------      582-801
PANAMA-206         ----------      582-801
PANAMA-349         ----------      582-801
PANAMA-A12         ----------      582-801
PANAMA-A13         ----------      582-801
PANAMA-B4          ----------      582-801
PANAMA-D6          ----------      582-801
PANAMA-F15         ----------      582-801
PANAMA-F38         ----------      582-801
PANAMA-F43         ----------      582-801
TB7=A32           --------------   562-848
GB11=WV_4043      --------------   562-855
TB17=A62          -------------    562-837
GB6=WV_3471       --------------   562-855
GB4=WV_3097       --------------   562-855
GB1=WV_2965       --------------   562-855
GB13=WV_75        --------------   562-855
HPK_IB            --------------   562-855
HPK_II            --------------   562-855
GB14=WV_1024      --------------   562-855
GB3=WV_3096       --------------   562-855
GB5=WV_3270       --------------   562-855
B20               --------------   562-850
GB10=WV_4042      ------------     562-826
GB12=WV_4044      ------------     562-826
PANAMA-F30         ----A------     582-801
SIHA               -C---------     582-801
MISSOURI-B11       ------T----     582-801
PANAMA-C20         -----C-----     582-801
O559              ---------C--C-   562-858
BT6               --------------   562-858
BT11              --------------   562-858
DT4               --------------   562-858
DT24              -T------------   562-858
BT23              --------------   562-858
DT42              --------------   562-858
DC255             --------------   562-858
DC212             --------------   562-858
DC141             --------------   562-858
DC207             --------------   562-858
DC269             --------------   562-858
F94_27            --C-----------   562-858
F94_143           --------------   562-858
F94_73            --------------   562-858
TB6=A27           ----------T--    562-837
F94_42            ---G----------   562-858
F94_85            ---G----------   562-858
F94_135           ---G----------   562-858
F94_161           ---G---------C   562-858
F94_44            ---G--------CC   562-858
F94_76            ---G---------C   562-858
F94_80            ---G---------C   562-858
```

E7 nuc cont'd

```
                 56666667777788
                 71446670389924
                 56573681290586
HPV16R           CAAAGGTCTTCTTT   562-858
F94_89           ---G---------C   562-858
F94_113          ---G---------C   562-858
F94_179          ----A--------C   562-858
PANAMA-200        --------C-G     582-801
TB14=A54         ---------C-G-    562-837
TB4=A21          ---------C-G--   562-848
B25              ---------C-G--   562-848
TB8=A36          ---------C-G--   562-848
B29              ---------C-G--   562-848
B61              ---------C-G--   562-848
TB9=A42          ---------C-G-    562-840
TB11=A46         ---------C-G-    562-837
B23              ---------C-G--   562-848
TB16=A61         ---------C-G-    562-837
BT7              ---------C-G--   562-858
BT9              ---------C-G--   562-858
BT12             ---------C-G--   562-858
BT10             ---------C-G--   562-858
BT20             ---------C-G--   562-858
BT22             ---------C-G--   562-858
O561             ---------C-G--   562-858
TB5=A24          ---G-----C-G--   562-848
B22              ---G-----C-G--   562-848
TB15=A56         ---G-----C-G-    562-837
B30              ---G-----C-G--   562-848
B33              ---G-----C-G--   562-848
TB12=A50         ---G-----C-G-    562-837
TB1=A6           ---G-----C-G     562-801
TB3=A16          ---G-----C-G     562-801
PANAMA-C10        --G-----C-G     582-801
84               ---G-----C-G--   562-858
1268             ---G-----C-G--   562-858
H84              ---G-----C-G--   562-858
GEORGIA-B9        -------CC-G     582-801
PANAMA-148        -------CC-G     582-801
PANAMA-B20        -------CC-G     582-801
ALABAMA-B23       -------CC-G     582-801
BT8              --------CC-G--   562-858
BT15             --------CC-G--   562-858
O462             --------CC-G--   562-858
F94_2            A-------CC-G--   562-858
```

E7 aa	01223334577789 59894597767895
HPV16R	TTLNEEDPFIRTIS
ALABAMA-B24	----------
ALASKA-C32	----------
ALASKA-C51	----------
CASKI	----------
MICHIGAN-B8	----------
PANAMA-136	----------
PANAMA-156	----------
PANAMA-206	----------
PANAMA-349	----------
PANAMA-A12	----------
PANAMA-A13	----------
PANAMA-B4	----------
PANAMA-D6	----------
PANAMA-F15	----------
PANAMA-F38	----------
PANAMA-F43	----------
TB7=A32	----------
GB11=WV_4043	----------
TB17=A62	----------
GB6=WV_3471	----------
GB4=WV_3097	----------
GB1=WV_2965	----------
GB13=WV_75	----------
HPK_IB	----------
HPK_II	----------
GB14=WV_1024	----------
GB3=WV_3096	----------
GB5=WV_3270	----------
B20	----------
GB10=WV_4042	----------
GB12=WV_4044	----------
PANAMA-F30	----*------
SIHA	-F---------
MISSOURI-B11	------L----
PANAMA-C20	-----*-----
O559	---------*--*-
BT6	----------
BT11	----------
DT4	----------
DT24	-S----------
BT23	----------
DT42	----------
DC255	----------
DC212	----------
DC141	----------
DC207	----------
DC269	----------
F94_27	--F----------
F94_143	----------
F94_73	----------
TB6=A27	----------C--
F94_42	---S----------
F94_85	---S----------
F94_135	---S----------
F94_161	---S---------*
F94_44	---S--------**
F94_76	---S---------*
F94_80	---S---------*

E7 aa cont'd

```
                  01223334577789
                  59894597767895
HPV16R            TTLNEEDPFIRTIS
F94_89            ---S---------*
F94_113           ---S---------*
F94_179           ----*--------*
PANAMA-200         --------*-*
TB14=A54          ---------*-*-
TB4=A21           ---------*-*--
B25               ---------*-*--
TB8=A36           ---------*-*--
B29               ---------*-*--
B61               ---------*-*--
TB9=A42           ---------*-*-
TB11=A46          ---------*-*-
B23               ---------*-*--
TB16=A61          ---------*-*-
BT7               ---------*-*--
BT9               ---------*-*--
BT12              ---------*-*--
BT10              ---------*-*--
BT20              ---------*-*--
BT22              ---------*-*--
O561              ---------*-*--
TB5=A24           ---S-----*-*--
B22               ---S-----*-*--
TB15=A56          ---S-----*-*-
B30               ---S-----*-*--
B33               ---S-----*-*--
TB12=A50          ---S-----*-*-
TB1=A6            ---S-----*-*
TB3=A16           ---S-----*-*
PANAMA-C10         --S-----*-*
84                ---S-----*-*--
1268              ---S-----*-*--
H84               ---S-----*-*--
GEORGIA-B9         -------**-*
PANAMA-148         -------**-*
PANAMA-B20         -------**-*
ALABAMA-B23        -------**-*
BT8               --------**-*--
BT15              --------**-*--
O462              --------**-*--
F94_2             K-------**-*--
```

E5 nuc

```
                    33333333333334444444444
                    88888899999990000000000
                    56677803778891134445578
                    87824344092813750260979
HPV16R              TTGCACTATATTCCGATATTAAT   3850-4101
DT24                ------------------------  3850-4101
BT11                ------------------------  3850-4101
DT4                 ------------------------  3850-4101
BT6                 -A----------------------  3850-4101
CASKI               ---A-----C-------G-----   3850-4101
SB10                ----G--G-C-------G---T-   3850-4101
SB13                ---------CC-----CG---T-   3850-4101
SB17                ---------C-------GC--T-   3850-4101
SB7                 ---------C-------G---T-   3850-4101
DC207               ---------C-------G-----   3850-4101
DC212               ---------C-------G-----   3850-4101
DT42                ---------C-------G-----   3850-4101
A00446              ---------C-------G-----   3850-4101
DC255               ---------C-------G-----   3850-4101
DC141               ---------C-------G-----   3850-4101
SB2                 ---------C-----G-G-----   3850-4101
BT23                ---------C-------G--N--   3850-4101
DC269               -----T---C-------G-----   3850-4101
SB5                 ---A----CC-------G-----   3850-4101
SIHA                ------A--C--G----G-----   3850-4101
SB16                C--------C--G-A--G----C   3850-4101
BT8                 C--------C--G-A--G-----   3850-4101
BT15                C--------C--G-A--G-----   3850-4101
BT7                 C--------C--G----G--C--   3850-4101
SB21A               C--------C--G----G--C-C   3850-4101
BT22                CNA------C--T----T--N--   3850-4101
BT10                C-A------C--T----T-G---   3850-4101
BT9                 C-A------C--T----T-----   3850-4101
BT20                C-A------C--T----T-----   3850-4101
BT12                C-A------C--T----T-----   3850-4101
TB8                 C-A------C--T----T----G   3850-4101
TB4                 C-A------C--T----T----C   3850-4101
TB16                CCA------C--T----T----C   3850-4101
TB13                C-A------G-CTT---T----C   3850-4101
```

E5 aa

```
                  00000112444445566666778
                  36789299145785624567060
HPV16R            NTASTLFISILLLSAVIIFVPTF
DT24              ------------------------
BT11              ------------------------
DT4               ------------------------
BT6               -*----------------------
CASKI             ---Y-----L-------V-----
SB10              ----A--V-L-------V---*-
SB13              ---------L*-----TV---*-
SB17              ---------L-------VS--*-
SB7               ---------L-------V---*-
DC207             ---------L-------V-----
DC212             ---------L-------V-----
DT42              ---------L-------V-----
A00446            ---------L-------V-----
DC255             ---------L-------V-----
DC141             ---------L-------V-----
SB2               ---------L-----*-V-----
BT23              ---------L-------V--*--
DC269             -----*---L-------V-----
SB5               ---Y----PL-------V-----
SIHA              ------I--L--V----V-----
SB16              *--------L--V-*--V----*
BT8               *--------L--V-*--V-----
BT15              *--------L--V-*--V-----
BT7               *--------L--V----V--*--
SB21A             *--------L--V----V--*-*
BT22              **T------L--*----L--*--
BT10              *-T------L--*----L-*---
BT9               *-T------L--*----L-----
BT20              *-T------L--*----L-----
BT12              *-T------L--*----L-----
TB8               *-T------L--*----L----L
TB4               *-T------L--*----L----*
TB16              **T------L--*----L----*
TB13              *-T------V-**F---L----*
```

L2 nuc

```
                 4444444444444444444444444444455555555555555555555
                 2334444444555556666678888999999000001111222222222223
                 8010112567114590144212581345663344734452333458899l
                 1817098219285790414504471840392415392526126791605O
HPV16R           TGAGAAGTGATATTATTTTCATCAAGAAGATATGATGCAACTGAATTAAT   4237-5658
BT15             ------                                              4237-4422
BT12             ------                                              4237-4422
BT20             ------                                              4237-4422
BT8              ------                                              4237-4422
DC269            ------                                              4237-4422
DC141            ------                                              4237-4422
DC255            ------                                              4237-4422
BT23             ------                                              4237-4422
DT4              ------                                              4237-4422
BT6              ------                                              4237-4422
BT11             ------                                              4237-4422
DT24             ------                                              4237-4422
DT42             ------                                              4237-4422
DC212            ------                                              4237-4422
DC207            ------                                              4237-4422
BT7              ------                                              4237-4422
BT22             ------                                              4237-4422
BT10             ------                                              4237-4422
BT9              ------                                              4237-4422
T49              ----------------------------------------------------4264-5632
T3               ----------------------------------------------------4264-5632
T45              ----------------------------------------------------4264-5632
T17              ----------------------------------------------------4264-5632
S83              ----------------------------------------C-----------4264-5632
S23              -------------------T-----A-------A-----T-----------4264-5632
OR.2087          -------------------T-----A-------A-----C-----------4273-5658
OR.7574          -------------------T-----A------CA-----C-----------4273-5658
OR.4716          -------------------T-----A------CA-----C-----------4273-5658
OR.5428          ------------------------A------C------C-----------4273-5658
OR.4724          ------------------------A----G-C------C-----------4273-5658
114B             ------------------------A-------------C-----------4237-5658
OR.5110          ------------------------A-------------C-----------4273-5658
OR.6170          ----G-------------------A-------------C-----------4273-5658
S108             ------------------------A-------------T-----------4264-5632
OR.4997          ------------------------A-------------T-----------4273-5658
OR.6311          --G---------------------A-------------T-----------4273-5658
OR.8329          ------------------------A--A--------T-T-----------4273-5658
OR.8987          ------------------------A--A--------T-T-----------4273-5658
114K             ------------------------A---------A---T-C---------4237-5658
S29              -----C-------C-----------A-------------C-----------4264-5632
NM.T455          -----C-------------T-----A-----G-------C-----------4273-5658
S93              --------A----------------A------C------C-----C----4264-5632
S27              ------------------------A------C------G-----------4264-5632
OR.9237          ----------------C--------A------C------G-----------4273-5658
NM.T446          ----------------C--------A------C------G-----------4273-5658
NM.T197          ----------------C--------A------C------C--A-------4273-5658
OR.0198          ----------A-----C--------A------C------G--A-------4273-5658
OR.1905          CA----A-A------C--AT-A--TA------C---A-----A----G-C 4273-5658
OR.6106          CA----A-A------C--ATC---TA------C---A-----A----G-C 4273-5658
OR.7587          CA----A-A------C-CAT----TA------C---A-----A----G-C 4273-5658
OR.7632          CA----A-A------C-CAT----TA------C---A-----AC---G-C 4273-5658
OR.3759          C--A--T-A--GG--C--AT--TG-A------C---A-----A-G--C-C 4273-5658
OR.3473          C-----T----GGG-C--AT--TG-A------C---A-----A-G--C-C 4273-5658
1194             C-----T----GGG-C--AT--TG-A------C---A-----A-G--C-C 4237-5658
OR.7145          C-----T----GGG-C--AT--TG-A------C---A-----A-G--C-C 4273-5658
OR.3136          C-----TCA------C--AT---G-A-G-G-TC-T-A-C---A---A--C 4273-5658
NM.T529          C-----TCA----GCC--AT---G-AGG-G-TC---A-----A---A-CA 4273-5658
NM.9999          C-----TCAT----CC--AT---G-AGG-G-TC---A-----A---A-CA 4273-5658
OR.8160          C-----TCAT----CC--AT---G-AGG-G-TC---A-----A---A-CA 4273-5658
OR.4541          C-----TCA-----CC--AT---G-AGG-G-TC---A---G-A---A-CA 4273-5658
NM.4094          C-----TCA-----CC--AT---G-AGG-G-TC---A---G-A---A-CA 4273-5658
```

L2 aa

```
                                                  ^         ^
                 00000000000011111111112222222222222333333333333333
                 12255667789900222336990123334466677000033333445555
                 54578142512437126563266754683566909123602247190238
HPV16R           AQAGLGSYLTPTPSTSIITFNINPLQKVVTLLSSPIRYILTTDELYTTSP
BT15             ------
BT12             ------
BT20             ------
BT8              ------
DC269            ------
DC141            ------
DC255            ------
BT23             ------
DT4              ------
BT6              ------
BT11             ------
DT24             ------
DT42             ------
DC212            ------
DC207            ------
BT7              ------
BT22             ------
BT10             ------
BT9              ------
T49              --------------------------------------------------
T3               --------------------------------------------------
T45              --------------------------------------------------
T17              --------------------------------------------------
S83              ---------------------------------------F----------
S23              -------------------*-----*-------N-----F----------
OR.2087          -------------------*-----*-------N-----F----------
OR.7574          -------------------*-----*------PN-----F----------
OR.4716          -------------------*-----*------PN-----F----------
OR.5428          -------------------------*------P------F----------
OR.4724          -------------------------*----VVP------F----------
114B             -------------------------*-------------F----------
OR.5110          -------------------------*-------------F----------
OR.6170          ----*--------------------*-------------F----------
S108             -------------------------*-------------F----------
OR.4997          -------------------------*-------------F----------
OR.6311          --*----------------------*-------------F----------
OR.8329          -------------------------*--I--------*-F----------
OR.8987          -------------------------*--I--------*-F----------
114K             -------------------------*---------*---F-*--------
S29              -----*-------*-----------*-------------F----------
NM.T455          -----*-------------*-----*-----*-------F----------
S93              --------*----------------*------P------F-----H----
S27              -------------------------*------P------*----------
OR.9237          ----------------*--------*------P------*----------
NM.T446          ----------------*--------*------P------*----------
NM.T197          ----------------*--------*------P------F--N-------
OR.0198          ----------*-----*--------*------P------*--N-------
OR.1905          **----*-*------P--**-*--**------P---*-----N----A-*
OR.6106          **----*-*------P--**H---**------P---*-----N----A-*
OR.7587          **----*-*------P-***----**------P---*-----N----A-*
OR.7632          **----*-*------P-***----**------P---*-----ND---A-*
OR.3759          *--*--*-*--**--P--**--**-*------P---*-----N-*--P-*
OR.3473          *-----*----***-P--**--**-*------P---*-----N-*--P-*
1194             *-----*----***-P--**--**-*------P---*-----N-*--P-*
OR.7145          *-----*----***-P--**--**-*------P---*-----N-*--P-*
OR.3136          *-----***------P--**---*-*-*-AFFP-*-*-L---N---*--*
NM.T529          *-----***----**P--**---*-***-AFFP---*-----N---*-**
NM.9999          *-----****----*P--**---*-***-AFFP---*-----N---*-**
OR.8160          *-----****----*P--**---*-***-AFFP---*-----N---*-**
OR.4541          *-----***-----*P--**---*-***-AFFP---*---SSN---*-**
NM.4094          *-----***-----*P--**---*-***-AFFP---*---SSN---*-**
```

L2 nuc cont'd.

```
                  5555555555555555555
                  3333333344444555555
                  4667888802789900136
                  6899167935575606824
HPV16R            ATCGTTCGTATCTAGGAAC   4237-5658
T49               -------------------   4264-5632
T3                -------------------   4264-5632
T45               -------------------   4264-5632
T17               -------------------   4264-5632
S83               -------------------   4264-5632
S23               --T-------------C--   4264-5632
OR.2087           --T-------------C--   4273-5658
OR.7574           --T-------------C--   4273-5658
OR.4716           --T-------------C--   4273-5658
OR.5428           ----------------C--   4273-5658
OR.4724           -------------------   4273-5658
114B              ---------C---------   4237-5658
OR.5110           -------------------   4273-5658
OR.6170           -------------------   4273-5658
S108              -------------------   4264-5632
OR.4997           -------------------   4273-5658
OR.6311           -------------------   4273-5658
OR.8329           -------------------   4273-5658
OR.8987           -------------------   4273-5658
114K              -------------------   4237-5658
S29               -------------------   4264-5632
NM.T455           C------------------   4273-5658
S93               -------------------   4264-5632
S27               ----A--------------   4264-5632
OR.9237           ------T------------   4273-5658
NM.T446           -------------------   4273-5658
NM.T197           -------------------   4273-5658
OR.0198           -------------------   4273-5658
OR.1905           -------AC---C--A--G   4273-5658
OR.6106           ---A---AC---C--A--G   4273-5658
OR.7587           ---A---AC---C--A--G   4273-5658
OR.7632           ---A---AC---C--A--G   4273-5658
OR.3759           --TA---AC--TC--A--G   4273-5658
OR.3473           --TA---AC--TC--A--G   4273-5658
1194              --TA---AC--TC--A--G   4237-5658
OR.7145           --TA---AC--TCC-A--G   4273-5658
OR.3136           -GTA-G-AC--TC--A-GG   4273-5658
NM.T529           -GTA-G-AC-ATC--A--G   4273-5658
NM.9999           -GTA-G-AC-ATC--A--G   4273-5658
OR.8160           -GTA-G-AC-ATC--A--G   4273-5658
OR.4541           -GTA-G-AC-ATC--A--G   4273-5658
NM.4094           -GTA-G-AC-ATC-AA--G   4273-5658
```

L2 aa cont'd.

```
                 3333333333344444444
                 7778888889112222234
                 0881244597370024823
HPV16R           ASSPVSSVSNGPIIDAIPA
T49              -------------------
T3               -------------------
T45              -------------------
T17              -------------------
S83              -------------------
S23              -FF-------------L--
OR.2087          -FF-------------L--
OR.7574          -FF-------------L--
OR.4716          -FF-------------L--
OR.5428          ----------------L--
OR.4724          -------------------
114B             ---------H---------
OR.5110          -------------------
OR.6170          -------------------
S108             -------------------
OR.4997          -------------------
OR.6311          -------------------
OR.8329          -------------------
OR.8987          -------------------
114K             -------------------
S29              -------------------
NM.T455          *------------------
S93              -------------------
S27              ----E--------------
OR.9237          -----FF------------
NM.T446          -------------------
NM.T197          -------------------
OR.0198          -------------------
OR.1905          -------I*---TT-T--G
OR.6106          ---*---I*---TT-T--G
OR.7587          ---*---I*---TT-T--G
OR.7632          ---*---I*---TT-T--G
OR.3759          -FF*---I*--*TT-T--G
OR.3473          -FF*---I*--*TT-T--G
1194             -FF*---I*--*TT-T--G
OR.7145          -FF*---I*--*TT-T--G
OR.3136          -VV*-AAI*--*TT-T-*G
NM.T529          -VV*-AAI*-**TT-T--G
NM.9999          -VV*-AAI*-**TT-T--G
OR.8160          -VV*-AAI*-**TT-T--G
OR.4541          -VV*-AAI*-**TT-T--G
NM.4094          -VV*-AAI*-**TTNT--G
```

L1 nuc

```
                5555566666666666666666666666666666666666667777
                6889901111122233344445555556678888899999990000
                9671671568901401934583456781920256616667795678
                8491203450098766146622198738513642583450541087
HPV16R          GCATTAAACATAGTAAAAAATTCCTAAAAGACCTCCAAACTGAGTA   5639-7156
ALASKA-C51                      -----------                     6348-6615
MICHIGAN-B8                     -----------                     6348-6615
PANAMA-156                      -----------                     6348-6615
PANAMA-206                      -----------                     6348-6615
PANAMA-B4                       -----------                     6348-6615
PANAMA-D6                       -----------                     6348-6615
BT6                             ----------                      6371-6577
BT11                            ----------                      6371-6577
DT4                             ----------                      6371-6577
DT24                            ----------                      6371-6577
114K            ------------------------------------------------5639-7156
NM.T455         ------C-----------------------------------------5666-7149
S23             ---------C--------------------------------A--   5663-7130
OR.5110         ------------------------------------------------5666-7149
OR.6170         ------------------------------------------------5666-7149
S29             ------------------------------------------------5663-7130
S83             ------------------------------------------------5663-7130
OR.4724         -----------------G----------------------------5666-7149
OR.4997         -----------------G----------------------------5666-7149
OR.8329         -----------------G----------------------------5666-7149
OR.8987         -----------------G----------------------------5666-7149
S27             -----------------G----------------------------5663-7130
114B            ------------A----G----------------------------5639-7156
OR.6311         ----C------------G----------------------------5666-7149
NM.T446         -----------------G---------------C------------5666-7149
OR.2087         -----------------G-------------------------A--5666-7149
OR.5428         -----------------G-------------------------A--5666-7149
OR.4716         -----------------G-------------------------A--5666-7149
S93             ----------------GG----------------------------5663-7130
NM.T197         -----------C-----G---------------C------------5666-7149
OR.9237         -----------------G---------------C----C-------5666-7149
OR.0198         -----------------G---------------C----------C-5666-7149
OR.7574         ---------C-------G-------------------------A--5666-7149
S99             -------C---------G---------------C------------5663-7130
T45             ----------------GG-G-------------------------5663-7130
ALASKA-C32                      -G---------                     6348-6615
CASKI                           -G---------                     6348-6615
PANAMA-A13                      -G---------                     6348-6615
PANAMA-C20                      -G---------                     6348-6615
PANAMA-F30                      -G---------                     6348-6615
PANAMA-F38                      -G---------                     6348-6615
BT23                            -G--------                      6371-6577
DT42                            -G--------                      6371-6577
DC255                           -G--------                      6371-6577
DC212                           -G--------                      6371-6577
DC141                           -G--------                      6371-6577
DC207                           -G--------                      6371-6577
DC269                           -G--------                      6371-6577
BT8                             -G--------                      6371-6577
BT15                            -G--------                      6371-6577
BT9                             -G--------                      6371-6577
BT12                            -G--------                      6371-6577
BT10                            -G--------                      6371-6577
BT22                            -G--------                      6371-6577
ALABAMA-B24                     -G---C-----                     6348-6615
MISSOURI-B11                    -G--------T                     6348-6615
PANAMA-349                      -G--------T                     6348-6615
SIHA                            -G---------                     6348-6615
T49             --G-------C---G-----------G-------T----C-----   5663-7130
```

L1 nuc cont'd

```
                5555566666666666666666666666666666666666667777
                6889901111122233344445555556678888899999990000
                9671671568901401934583456781920256616667795678
                8491203450098766146622198738513642583450541087
HPV16R          GCATTAAACATAGTAAAAAATTCCTAAAAGACCTCCAAACTGAGTA   5639-7156
T3              --G-------C---G------------G-------T----C-----   5663-7130
T17             --G-------C---G------------G-------T----C-----   5663-7130
BT7                           -G--C-----                        6371-6577
BT20                          -GC-------                        6371-6577
GEORGIA-B9                    -G-----T---                       6348-6615
PANAMA-200                    -G-----T---                       6348-6615
PANAMA-B20                    -G-----T---                       6348-6615
ALABAMA-B23                   -G-----TC--                       6348-6615
649                         -G-G--C--T----                      6270-6619
77                          -G-G--C--T----                      6270-6619
OR.1905         AT-C----A----C-G-G-----TA----A--T------T-A-T--   5666-7149
OR.6106         AT-C-G--A----C-G-G-----TAG---A--T----T-T-A-T-G   5666-7149
OR.7632         AT-C----A----C-G-G-----TAG---A-TT------T-A-T--   5666-7149
OR.7587         AT-C----A----C-G-G-----TAG---A--T------T-A-T--   5666-7149
NM.4094         AT-C----AC---C---G-----T----CAT-T-T----T-A-T--   5666-7149
NM.9999         AT-C----AC---C---G-----T----CAT-T-T----T-A-T--   5666-7149
OR.4541         AT-C----AC---C---G-----T----CAT-T-T----T-A-T--   5666-7149
OR.8160         AT-C----AC---C---G-----T----CAT-T-T----T-A-T--   5666-7149
NM.T529         AT-C----AC---C---G----AT----CAT-T-T----T-A-T--   5666-7149
OR.3136         AT-C----A----C-G-G-----T----CA--T-T----T-A-T--   5666-7149
OR.3759         AT-C----A----C-G-G--C--T----CA--T-T----T-A-T--   5666-7149
OR.3473         AT-C----AC---C-G-G--C--T----CA--T-T----T-A-T--   5666-7149
OR.7145         AT-C----AC---C-G-G--C--T----CA--T-T----T-A-T--   5666-7149
1194            AT-C----AC---C-G-G--C--T----CA--T-T----T-AGT--   5639-7156
84                          -G-G--C--T----CA--T-T-C--T-A         6270-6619,6658-7007
1268                                     CA--T-T-.--T-A         6658-7007
```

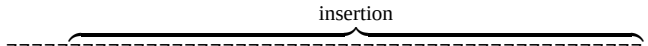
L1 aa

```
                  0000111111111222222222333333333344444444444444444
                  2789045778899022567789001112568900024444457788
                  0611849261414336160328170357319668972234621403
HPV16R            KHNYGRIGTNDIVTDTRTNDSSTFPLRNTKTSLPPAKKKYFKQLFG
ALASKA-C51                        -----------
MICHIGAN-B8                       -----------
PANAMA-156                        -----------
PANAMA-206                        -----------
PANAMA-B4                         -----------
PANAMA-D6                         -----------
BT6                               ----------
BT11                              ----------
DT4                               ----------
DT24                              ----------
114K              ----------------------------------------------
NM.T455           ------L---------------------------------------
S23               ---------T-----------------------------------*--
OR.5110           ----------------------------------------------
OR.6170           ----------------------------------------------
S29               ----------------------------------------------
S83               ----------------------------------------------
OR.4724           ----------------A---------------------------
OR.4997           ----------------A---------------------------
OR.8329           ----------------A---------------------------
OR.8987           ----------------A---------------------------
S27               ----------------A---------------------------
114B              ------------I----A---------------------------
OR.6311           ----*-----------A---------------------------
NM.T446           ----------------A---------------*------------
OR.2087           ----------------A------------------------*--
OR.5428           ----------------A------------------------*--
OR.4716           ----------------A------------------------*--
S93               ----------------*A---------------------------
NM.T197           -----------L-----A---------------*------------
OR.9237           ----------------A---------------*----Q-------
OR.0198           ----------------A---------------*----------*-
OR.7574           ---------T-------A------------------------*--
S99               -------*---------A---------------*------------
T45               ----------------*A-G-------------------------
ALASKA-C32                        -A---------
CASKI                             -A---------
PANAMA-A13                        -A---------
PANAMA-C20                        -A---------
PANAMA-F30                        -A---------
PANAMA-F38                        -A---------
BT23                              -A--------
DT42                              -A--------
DC255                             -A--------
DC212                             -A--------
DC141                             -A--------
DC207                             -A--------
DC269                             -A--------
BT8                               -A--------
BT15                              -A--------
BT9                               -A--------
BT12                              -A--------
BT10                              -A--------
BT22                              -A--------
ALABAMA-B24                       -A---*-----
MISSOURI-B11                      -A--------*
PANAMA-349                        -A--------*
SIHA                              -A---------
T49               --D-------*---G------------S-------V----S-----
```

L1 aa cont'd

```
                    00001111111112222222223333333333444444444444444
                    27890457788990225677890011125689000244444577 88
                    06118492614143361603281703573196689722346214 03
HPV16R              KHNYGRIGTNDIVTDTRTNDSSTFPLRNTKTSLPPAKKKYFKQLFG
T3                  --D-------*---G------------S-------V----S-----
T17                 --D-------*---G------------S-------V----S-----
BT7                                 -A--P-----
BT20                                -AH-------
GEORGIA-B9                          -A-----*---
PANAMA-200                          -A-----*---
PANAMA-B20                          -A-----*---
ALABAMA-B23                         -A-----**--
649                               -*-A--P--*----
77                                -*-A--P--*----
OR.1905             *Y-*----N----*-*-A-----**----*--*------*-*-F--
OR.6106             *Y-*-*--N----*-*-A-----***---*--*---NN-*-*-F-*
OR.7632             *Y-*----N----*-*-A-----***---*-**------*-*-F--
OR.7587             *Y-*----N----*-*-A-----***---*--*------*-*-F--
NM.4094             *Y-*----NT---*---A-----*----P*S-*-*----*-*-F--
NM.9999             *Y-*----NT---*---A-----*----P*S-*-*----*-*-F--
OR.4541             *Y-*----NT---*---A-----*----P*S-*-*----*-*-F--
OR.8160             *Y-*----NT---*---A-----*----P*S-*-*----*-*-F--
NM.T529             *Y-*----NT---*---A----**----P*S-*-*----*-*-F--
OR.3136             *Y-*----N----*-*-A-----*----P*--*-*----*-*-F--
OR.3759             *Y-*----N----*-*-A--P--*----P*--*-*----*-*-F--
OR.3473             *Y-*----NT---*-*-A--P--*----P*--*-*----*-*-F--
OR.7145             *Y-*----NT---*-*-A--P--*----P*--*-*----*-*-F--
1194                *Y-*----NT---*-*-A--P--*----P*--*-*----*-**F--
84                                -*-A--P--*----P*--*-*-TT-*-*
1268                                            P*--*-*-XX-*-*
```

LCR




```
                  7777777777777777777777777777777777777777777777777
                  1222333333333333333333333333333333333333333333333
                  9336344444444444444444444444444444444444444444444
                  3233900000000000000000000000000000000000000000000
HPV16R            GAAAA.....................................................  7157-7906
BT6               -----.....................................................  7174-7853
BT11              -----.....................................................  7174-7853
DT24              T----.....................................................  7174-7853
DT42              T----.....................................................  7174-7853
DC255             T----.....................................................  7174-7853
DC212             T----.....................................................  7174-7853
DC141             T----.....................................................  7174-7853
DC269             T----.....................................................  7174-7853
DC207             T----.....................................................  7174-7853
BT7               T-C--.....................................................  7174-7853
BT8               T-C--.....................................................  7174-7853
BT15              T-C--.....................................................  7174-7853
BAVINC0           T----.....................................................  7157-7906
BT20              TG---.....................................................  7174-7853
BT10              TG---.....................................................  7174-7853
BT12              TG---.....................................................  7174-7853
BT9               TG---.....................................................  7174-7853
BT23              T----.....................................................  7174-7853
DT4               T----.....................................................  7174-7853
BT22              NNN--.....................................................  7174-7853
BB1                    ACAAACTGTATTTTTTTAAATAGTTCTATGTCAGCAACTATAGTT          7340-7843
BAVINC3           T-CGT.....................................................  7157-7906
```

LCR cont'd.

```
                  7777777777777777777777777777777777777777777777777
                  3333333333333333333333333333333444444444455555555
                  4444444444444444444444444446993356888899002224444
                  0000000000000000000000000000045360401591647146012 3
HPV16R            ........................GCCGGT.ACAGTTTAGTACCTG   7157-7906
OR.6170                                     ---------------          7480-7843
BB8A                                        -----------....          7480-7843
BT6               ......................------.---------------      7174-7853
HDS81B.10/04/89                          -.---------------          7447-7832
HDS81B.11/07/90                          -.---------------          7447-7832
HDS81B.06/19/91                          -.---------------          7447-7832
HDS81B.02/24/92                          -.---------------          7447-7832
HDS69.10/16/89                           -.---------------          7447-7832
BT11              ......................------.---------------      7174-7853
OR.5110                                     ---------------          7480-7843
NM.T455                                     ---------------          7480-7843
GALUC64                                     ---------------          7480-7843
DT24              ......................------.---------------      7174-7853
HDS85.06/12/90                           -.--------A------          7447-7832
CASKI                                       --------A------          7480-7843
OR.4724                                     --------A------          7480-7843
OR.4997                                     --------A------          7480-7843
OR.6311                                     --------A------          7480-7843
OR.8329                                     --------A------          7480-7843
OR.8987                                     --------A------          7480-7843
NM.T197                                     --------A------          7480-7843
OR.9237                                     --------A------          7480-7843
NM.T446                                     --------A------          7480-7843
OR.0198                                     --------A------          7480-7843
SIHA                                        --------A------          7480-7843
DT42              ......................------.--------A------      7174-7853
```

LCR cont'd.

```
                                 insertion                d i                deletion
                 7777777777777777777777777777777777777777777777777777
                 3333333333333333333333333333333444444444455555555
                 4444444444444444444444444444699335688889900222444
                 0000000000000000000000000000453604015916471460123
HPV16R           ............................GCCGGT.ACAGTTTAGTACCTG   7157-7906
DC255            ............................------.--------A------   7174-7853
DC212            ............................------.--------A------   7174-7853
DC141            ............................------.--------A------   7174-7853
DC269            ............................------.--------A------   7174-7853
DC207            ............................C-----.--------A------   7174-7853
BT7              ............................------.--------A------   7174-7853
BT8              ............................------.--------A------   7174-7853
BT15             ............................------.--------A------   7174-7853
SB14                                                ----C---A------   7480-7843
J-5=SEQH                                            --------A------   7480-7843
OR.7574                                             --------A------   7480-7843
SB1                                                 --------A------   7480-7843
SB7                                                 --------A------   7480-7843
OR.2087                                             --------A------   7480-7843
OR.5428                                             --------A------   7480-7843
OR.4716                                             --------A------   7480-7843
SB13                                                --------A------   7480-7843
BAVINC0          ............................---.-CT--------AC-----   7157-7906
SA-9=SEQA                                           -----C--A------   7480-7843
SH-2a=SEQK                                          --------A------   7480-7843
AP-4=SEQI                                           --------A------   7480-7843
GB13                                                --------A------   7480-7843
IND-4=SEQG                                          --------A------   7480-7843
GB10                                                --------A------   7480-7843
GB21A                                               --------A------   7480-7843
BB24                                                ---------------   7480-7843
SS1A                                                TA-------------   7480-7843
IND-5=SEQJ                                          --------A------   7480-7843
BT20             ............................------.--------A------   7174-7853
BT10             ............................------.--------A------   7174-7853
BT12             ............................------.--------A------   7174-7853
BT9              ............................------.--------A------   7174-7853
SB5AA                                               ---------------   7480-7843
BT23             ............................------.---------------   7174-7853
DT4              ............................------.---------------   7174-7853
BB3                                                 --------A------   7480-7843
BT22             ............................------.--------A------   7174-7853
IND-7=SEQF                                          --------A------   7480-7843
GALUC153                                            --------A------   7480-7843
GALUC136                                            --------A-G----   7480-7843
GALUC73                                             -------CA------   7480-7843
SL-6a=SEQD                                          --------A------   7480-7843
TS6A                                                ---A----A------   7480-7843
TB13                                                ---A----A------   7480-7843
TS6B                                                ---A----A------   7480-7843
OR.1905                                             ---A----A------   7480-7843
TB4                                                 ---A----A------   7480-7843
OR.7587                                             ---A----A------   7480-7843
TB8                                                 ---A----A------   7480-7843
GALUC91                                             ---A----A------   7480-7843
OR.6106                                             ---A----A------   7480-7843
OR.7632                                             ---A----A------   7480-7843
TB16                                                ---A----A------   7480-7843
TB15                                                ---A----A------   7480-7843
OR.3136                                             --CA----A------   7480-7843
SB21                                                --CA----A------   7480-7843
BB11                                                --CA----A------   7480-7843
AM-6=SEQB                                           --CA----A------   7480-7843
```

LCR cont'd.

```
                                 insertion                d i                deletion
                        ___________________________________________________________
                        7777777777777777777777777777777777777777777777777777
                        3333333333333333333333333333333344444444445555555555
                        4444444444444444444444444444469933568888990022244444
                        0000000000000000000000000000045360401591647146012 3
HPV16R                  ............................GCCGGT.ACAGTTTAGTACCTG   7157-7906
OR.5691                                                  --CA----A------   7480-7843
OR.1783                                                  --CA----A------   7480-7843
AN-12=SEQE                                               --CA----A------   7480-7843
OR.4541                                                  --CA----A------   7480-7843
BB1                     TAAACTTGTACGTTTCCTGCTTGCCATG------.--CA----A------   7340-7843
HDS81A.10/04/89                                        -.--CA----A------   7447-7832
HDS81A.11/07/90                                        -.--CA----A------   7447-7832
HDS81A.06/19/91                                        -.--CA----A------   7447-7832
HDS81A.02/24/92                                        -.--CA----A------   7447-7832
HDS109.07/25/90                                        -.--CA----A------   7447-7832
HDS109.11/30/90                                        -.--CA----A------   7447-7832
NM.T529                                                  --CA----A------   7480-7843
NM.9999                                                  --CA----A------   7480-7843
OR.7908                                                  --CA----A------   7480-7843
OR.8863                                                  --CA----A------   7480-7843
OR.7754                                                  --CA----A------   7480-7843
OR.4072                                                  --CA----A------   7480-7843
OR.4451                                                  --CA----A------   7480-7843
OR.8160                                                  --CA----A------   7480-7843
NM.4094                                                  --CA----A------   7480-7843
SB16                                                     --CA----A------   7480-7843
OR.7875                                                  --CA----A------   7480-7843
BB19                                                     --CA---GA------   7480-7843
BB14                                                     --CA---GA------   7480-7843
P-1=SEQC                                                 --CA---GA------   7480-7843
BAVINC3                 ............................-TT-.-T--CA----A------   7157-7906
TS5                                                      --CA----A------   7480-7843
TS5C                                                     --CA----A------   7480-7843
TS5A                                                     --CA----A------   7480-7843
TS5B                                                     --CA----A------   7480-7843
TS3A                                                     --CA--C-A------   7480-7843
SN-4=SEQL                                                --CA----A------   7480-7843
TS3C                                                     --CA----A------   7480-7843
TS3B                                                     --CA----A------   7480-7843
TB1                                                      --CA----A------   7480-7843
TS3                                                      --CA----A------   7480-7843
OR.3473                                                  --CA----A------   7480-7843
OR.3759                                                  --CA----A------   7480-7843
OR.7145                                                  --CA----A------   7480-7843
AN-10a=SEQN                                              --CA----A------   7480-7843
SB21A                                                    --CA----A------   7480-7843
AN-10=SEQM                                               --------A------   7480-7843
GALUC106                                                 --CA----A------   7480-7843
```

LCR cont'd.

deletion

```
                 7777777777777777777777777777777777777777777777777
                 5555555555555555555555555555555555555555555555555
                 4444445555555556666666666777777777788888888899999
                 4567890123456789012345678901234567890123456789012
HPV16R           CTTGCCATGCGTGCCAAATCCCTGTTTTCCTGACCTGCACTGCTTGCCAA  7157-7906
OR.6170          ---------------------------------------------------  7480-7843
BB8A             ...................................................  7480-7843
BT6              ---------------------------------------------------  7174-7853
HDS81B.10/04/89  ---------------------------------------------------  7447-7832
HDS81B.11/07/90  ---------------------------------------------------  7447-7832
HDS81B.06/19/91  ---------------------------------------------------  7447-7832
HDS81B.02/24/92  ---------------------------------------------------  7447-7832
HDS69.10/16/89   ---------------------------------------------------  7447-7832
BT11             ---------------------------------------------------  7174-7853
OR.5110          ---------------------------------------------------  7480-7843
NM.T455          ---------------------------------------------------  7480-7843
GALUC64          ---------------------------------------------------  7480-7843
DT24             ---------------------------------------------------  7174-7853
HDS85.06/12/90   ---------------------------------------------------  7447-7832
CASKI            ---------------------------------------------------  7480-7843
OR.4724          ---------------------------------------------------  7480-7843
OR.4997          ---------------------------------------------------  7480-7843
OR.6311          ---------------------------------------------------  7480-7843
OR.8329          ---------------------------------------------------  7480-7843
OR.8987          ---------------------------------------------------  7480-7843
NM.T197          ---------------------------------------------------  7480-7843
OR.9237          ---------------------------------------------------  7480-7843
NM.T446          ---------------------------------------------------  7480-7843
OR.0198          ---------------------------------------------------  7480-7843
SIHA             ---------------------------------------------------  7480-7843
DT42             ---------------------------------------------------  7174-7853
DC255            ---------------------------------------------------  7174-7853
DC212            ---------------------------------------------------  7174-7853
DC141            ---------------------------------------------------  7174-7853
DC269            ---------------------------------------------------  7174-7853
DC207            ---------------------------------------------------  7174-7853
BT7              ---------------------------------------------------  7174-7853
BT8              ---------------------------------------------------  7174-7853
BT15             ---------------------------------------------------  7174-7853
SB14             ---------------------------------------------------  7480-7843
J-5=SEQH         ---------------------------------------------------  7480-7843
OR.7574          ---------------------------------------------------  7480-7843
SB1              ---------------------------------------------------  7480-7843
SB7              ---------------------------------------------------  7480-7843
OR.2087          ---------------------------------------------------  7480-7843
OR.5428          ---------------------------------------------------  7480-7843
OR.4716          ---------------------------------------------------  7480-7843
SB13             ---------------------------------------------------  7480-7843
BAVINC0          ---------------------------------------------------  7157-7906
SA-9=SEQA        ---------------------------------------------------  7480-7843
SH-2a=SEQK       ---------------------------------------------------  7480-7843
AP-4=SEQI        ---------------------------------------------------  7480-7843
GB13             ---------------------------------------------------  7480-7843
IND-4=SEQG       ---------------------------------------------------  7480-7843
GB10             --------A------------------------------------------  7480-7843
GB21A            --------A------------------------------------------  7480-7843
BB24             ---------------------------------------------------  7480-7843
SS1A             --------------------------------T------------------  7480-7843
IND-5=SEQJ       ---------------------------------------------------  7480-7843
BT20             ---------------------------------------------------  7174-7853
BT10             ---------------------------------------------------  7174-7853
BT12             ---------------------------------------------------  7174-7853
BT9              ---------------------------------------------------  7174-7853
```

LCR cont'd.

deletion

```
                  7777777777777777777777777777777777777777777777777
                  5555555555555555555555555555555555555555555555555
                  4444445555555556666666666777777777788888888899999
                  4567890123456789012345678901234567890123456789012
HPV16R            CTTGCCATGCGTGCCAAATCCCTGTTTTCCTGACCTGCACTGCTTGCCAA  7157-7906
SB5AA             ---------------------------------------------------  7480-7843
BT23              ---------------------------------------------------  7174-7853
DT4               ---------------------------------------------------  7174-7853
BB3               ----------------------A----------------------------  7480-7843
BT22              ---------------------------------------------------  7174-7853
IND-7=SEQF        ---------------------------------------------------  7480-7843
GALUC153          ---------------------------------------------------  7480-7843
GALUC136          ---------------------------------------------------  7480-7843
GALUC73           ---------------------------------------------------  7480-7843
SL-6a=SEQD        ---------------------------------------------------  7480-7843
TS6A              ---------------G-----------------------------------  7480-7843
TB13              ---------------------------------------------------  7480-7843
TS6B              ---------------------------------------------------  7480-7843
OR.1905           ---------------------------------------------------  7480-7843
TB4               ---------------------------------------------------  7480-7843
OR.7587           ---------------------------------------------------  7480-7843
TB8               ---------------------------------------------------  7480-7843
GALUC91           ---------------------------------------------------  7480-7843
OR.6106           ---------------------------------------------------  7480-7843
OR.7632           ---------------------------------------------------  7480-7843
TB16              ---------------------------------------------------  7480-7843
TB15              ---------------------------------------------------  7480-7843
OR.3136           ---------------------------------------------------  7480-7843
SB21              ---------------------------------------------------  7480-7843
BB11              ---------------------------------------------------  7480-7843
AM-6=SEQB         ---------------------------------------------------  7480-7843
OR.5691           ---------------------------------------------------  7480-7843
OR.1783           ---------------------------------------------------  7480-7843
AN-12=SEQE        ---------------------------------------------------  7480-7843
OR.4541           ---------------------------------------------------  7480-7843
BB1               ---------------------------------------------------  7340-7843
HDS81A.10/04/89   ---------------------------------------------------  7447-7832
HDS81A.11/07/90   ---------------------------------------------------  7447-7832
HDS81A.06/19/91   ---------------------------------------------------  7447-7832
HDS81A.02/24/92   ---------------------------------------------------  7447-7832
HDS109.07/25/90   ---------------------------------------------------  7447-7832
HDS109.11/30/90   ---------------------------------------------------  7447-7832
NM.T529           ---------------------------------------------------  7480-7843
NM.9999           ---------------------------------------------------  7480-7843
OR.7908           ---------------------------------------------------  7480-7843
OR.8863           ---------------------------------------------------  7480-7843
OR.7754           ---------------------------------------------------  7480-7843
OR.4072           ---------------------------------------------------  7480-7843
OR.4451           ---------------------------------------------------  7480-7843
OR.8160           ---------------------------------------------------  7480-7843
NM.4094           ---------------------------------------------------  7480-7843
SB16              ---------------------------------------------------  7480-7843
OR.7875           ---------------------------------------------------  7480-7843
BB19              ---------------------------------------------------  7480-7843
BB14              ---------------------------------------------------  7480-7843
P-1=SEQC          ---------------------------------------------------  7480-7843
BAVINC3           ---------------------------------------------------  7157-7906
TS5               ---------------------------------------------------  7480-7843
TS5C              ---------------------------------------------------  7480-7843
TS5A              ---------------------------------------------------  7480-7843
TS5B              ---------------------------------------------------  7480-7843
TS3A              ---------------------------------------------------  7480-7843
SN-4=SEQL         ---------------------------------------------------  7480-7843
TS3C              --------------------T------------------------------  7480-7843
```

LCR cont'd.

deletion

```
                 7777777777777777777777777777777777777777777777777
                 5555555555555555555555555555555555555555555555555
                 4444445555555555666666666677777777778888888889999
                 4567890123456789012345678901234567890123456789012
HPV16R           CTTGCCATGCGTGCCAAATCCCTGTTTTCCTGACCTGCACTGCTTGCCAA  7157-7906
TS3B             ---------------------------------------------------  7480-7843
TB1              ---------------------------------------------------  7480-7843
TS3              ----------------------A----------------------------  7480-7843
OR.3473          ---------------------------------------------------  7480-7843
OR.3759          ---------------------------------------------------  7480-7843
OR.7145          ---------------------------------------------------  7480-7843
AN-10a=SEQN      ---------------------------------------------------  7480-7843
SB21A            ---------------------------------------------------  7480-7843
AN-10=SEQM       ---------------------------------------------------  7480-7843
GALUC106         ---------------------------------------------------  7480-7843
```

LCR cont'd.

deletion

```
                                  i         i
                  7777777777777777777777777777777777777777777777777
                  5555556666666666666666666666666666666666666666666
                  9999990000000000111111111112222222222333333333344
                  4567890123456789012345667890123456678901234567890 1
HPV16R            CCATTCCATTGTTTTTTACACT.GCACTATGTG.CAACTACTGAATCACT  7157-7906
OR.6170           ----------------------.----------.----------------  7480-7843
BB8A              ......................................................  7480-7843
BT6               ----------------------.----------.----------------  7174-7853
HDS81B.10/04/89   ----------------------.----------.----------------  7447-7832
HDS81B.11/07/90   ----------------------.----------.----------------  7447-7832
HDS81B.06/19/91   ----------------------.----------.----------------  7447-7832
HDS81B.02/24/92   ----------------------.----------.----------------  7447-7832
HDS69.10/16/89    ----------------------.----------.----------------  7447-7832
BT11              ----------------------.----------.----------------  7174-7853
OR.5110           ----------------------.----------.----------------  7480-7843
NM.T455           ----------------------.----------.----------------  7480-7843
GALUC64           ----------------------.----------.----------------  7480-7843
DT24              ----------------------.----------.----------------  7174-7853
HDS85.06/12/90    ----------------------.----------.----------------  7447-7832
CASKI             ----------------------.----------.----------------  7480-7843
OR.4724           ----------------------.----------.----------------  7480-7843
OR.4997           ----------------------.----------.----------------  7480-7843
OR.6311           ----------------------.----------.----------------  7480-7843
OR.8329           ----------------------.----------.----------------  7480-7843
OR.8987           ----------------------.----------.----------------  7480-7843
NM.T197           ----------------------.----------.----------------  7480-7843
OR.9237           ----------------------.----------.----------------  7480-7843
NM.T446           ----------------------.----------.----------------  7480-7843
OR.0198           ----------------------.----------.----------------  7480-7843
SIHA              ----------------------.----------.----------------  7480-7843
DT42              ----------------------.----------.----------------  7174-7853
DC255             ----------------------.----------.----------------  7174-7853
DC212             ----------------------.----------.----------------  7174-7853
DC141             ----------------------.----------.----------------  7174-7853
DC269             ----------------------.----------.----------------  7174-7853
DC207             ----------------------.----------.----------------  7174-7853
BT7               ----------------------.----------.----------------  7174-7853
BT8               ----------------------.----------.----------------  7174-7853
BT15              ----------------------.----------.----------------  7174-7853
SB14              ----------------------.----------.----------------  7480-7843
J-5=SEQH          ----------------------.----------.----------------  7480-7843
OR.7574           ----------------------.----------.----------------  7480-7843
SB1               ----------------------.----------.----------------  7480-7843
SB7               ----------------------.----------.----------------  7480-7843
OR.2087           ----------------------.----------.----------------  7480-7843
OR.5428           ----------------------.----------.----------------  7480-7843
OR.4716           ----------------------.----------.----------------  7480-7843
SB13              ----------------------.----------.----------------  7480-7843
BAVINC0           ----------------------.----------.----------------  7157-7906
SA-9=SEQA         ----------------------.----------.----------------  7480-7843
SH-2a=SEQK        ----------------------.----------.----------------  7480-7843
AP-4=SEQI         ----------------------.----------.----------------  7480-7843
GB13              ----------------------.----------.----------------  7480-7843
IND-4=SEQG        ----------------------.----------.----------------  7480-7843
GB10              ----------------------.----------.----------------  7480-7843
GB21A             ----------------------.---------A.----------------  7480-7843
BB24              ----------------------.----------.----------------  7480-7843
SS1A              ----------------------.----------.------------G--  7480-7843
IND-5=SEQJ        ----------------------.----------.----------------  7480-7843
BT20              ----------------------.----------.----------------  7174-7853
BT10              ----------------------.----------.----------------  7174-7853
BT12              ----------------------.----------.----------------  7174-7853
BT9               ----------------------.----------.----------------  7174-7853
SB5AA             ----------------------.----------.----------------  7480-7843
BT23              ----------------------.----------.----------------  7174-7853
DT4               ----------------------.----------.----------------  7174-7853
BB3               ----------------------.----------.----------------  7480-7843
BT22              ----------------------.----------.----------------  7174-7853
IND-7=SEQF        ----------------------.----------.----------------  7480-7843
```

LCR cont'd.

```
                                        deletion
                                        i          i
                 77777777777777777777777777777777777777777777777777
                 55555566666666666666666666666666666666666666666666
                 99999900000000001111111111122222222222333333333344
                 45678901234567890123456678901234566789012345678901
HPV16R           CCATTCCATTGTTTTTTACACT.GCACTATGTG.CAACTACTGAATCACT   7157-7906
GALUC153         ----------------------.----------.----------------   7480-7843
GALUC136         ----------------------.----------.----------------   7480-7843
GALUC73          ----------------------.----------.----------------   7480-7843
SL-6a=SEQD       ----------------------.----------.----------------   7480-7843
TS6A             ----------------------.----------.----------------   7480-7843
TB13             ----------------------.----------.----------------   7480-7843
TS6B             ----------------------.----------.----------------   7480-7843
OR.1905          ----------------------.----------.----------------   7480-7843
TB4              ----------------------.----------.----------------   7480-7843
OR.7587          ----------------------.----------.----------------   7480-7843
TB8              ----------------------.----------.----------------   7480-7843
GALUC91          ----------------------.----------.----------------   7480-7843
OR.6106          ----------------------.----------.----------------   7480-7843
OR.7632          ----------------------.----------.----------------   7480-7843
TB16             ----------------------.----------.----------------   7480-7843
TB15             ----------------------.----------.----------------   7480-7843
OR.3136          ----------------------.----------.----------------   7480-7843
SB21             ----------------------G----------G----------------   7480-7843
BB11             ----------------------.----------.----------------   7480-7843
AM-6=SEQB        ----------------------.----------.----------------   7480-7843
OR.5691          ----------------------.----------.----------------   7480-7843
OR.1783          ----------------------.----------.----------------   7480-7843
AN-12=SEQE       ----------------------.----------.----------------   7480-7843
OR.4541          ----------------------.----------.----------------   7480-7843
BB1              ----------------------.----------.----------------   7340-7843
HDS81A.10/04/89  ----------------------.----------.----------------   7447-7832
HDS81A.11/07/90  ----------------------.----------.----------------   7447-7832
HDS81A.06/19/91  ----------------------.----------.----------------   7447-7832
HDS81A.02/24/92  ----------------------.----------.----------------   7447-7832
HDS109.07/25/90  ----------------------.----------.----------------   7447-7832
HDS109.11/30/90  ----------------------.----------.----------------   7447-7832
NM.T529          ----------------------.----------.----------------   7480-7843
NM.9999          ----------------------.----------.----------------   7480-7843
OR.7908          ----------------------.----------.----------------   7480-7843
OR.8863          ----------------------.----------.----------------   7480-7843
OR.7754          ----------------------.----------.----------------   7480-7843
OR.4072          ----------------------.----------.----------------   7480-7843
OR.4451          ----------------------.----------.----------------   7480-7843
OR.8160          ----------------------.----------.----------------   7480-7843
NM.4094          ----------------------.----------.----------------   7480-7843
SB16             ----------------------.----------.----------------   7480-7843
OR.7875          ----------------------.----------.----------------   7480-7843
BB19             ----------------------.----------.----------------   7480-7843
BB14             ----------------------.----------.----------------   7480-7843
P-1=SEQC         ----------------------.----------.----------------   7480-7843
BAVINC3          ----------------------.----------.----------------   7157-7906
TS5              ----------------------.----------.----------------   7480-7843
TS5C             ----------------------.----------.----------------   7480-7843
TS5A             ----------------------.----------.----------------   7480-7843
TS5B             ----------------------.--------C-.----------------   7480-7843
TS3A             ---C------------------.----------.----------------   7480-7843
SN-4=SEQL        ----------------------.----------.----------------   7480-7843
TS3C             ----------------------.----------.----------------   7480-7843
TS3B             ----------------------.----------.----------------   7480-7843
TB1              ----------------------.----------.----------------   7480-7843
TS3              ----------------------.----------.----------------   7480-7843
OR.3473          ----------------------.----------.----------------   7480-7843
OR.3759          ----------------------.----------.----------------   7480-7843
OR.7145          ----------------------.----------.----------------   7480-7843
AN-10a=SEQN      ----------------------.----------.----------------   7480-7843
SB21A            ----------------------.----------.----------------   7480-7843
AN-10=SEQM       ----------------------.----------.----------------   7480-7843
GALUC106         ----------------------.----------.----------------   7480-7843
```

LCR cont'd.

deletion

```
                  77777777777777777777777777777777777777777777777777
                  66666666666666666666666666666666666666666666666666
                  44444444555555555566666666667777777777888888888899
                  23456789012345678901234567890123456789012345678901
HPV16R            ATGTACATTGTGTCATATAAAATAAATCACTATGCGCCAACGCCTTACAT   7157-7906
OR.6170           ----------------------------------------------------   7480-7843
BB8A              ....................................................   7480-7843
BT6               ----------------------------------------------------   7174-7853
HDS81B.10/04/89   ----------------------------------------------------   7447-7832
HDS81B.11/07/90   ----------------------------------------------------   7447-7832
HDS81B.06/19/91   ----------------------------------------------------   7447-7832
HDS81B.02/24/92   ----------------------------------------------------   7447-7832
HDS69.10/16/89    ----------------------------------------------------   7447-7832
BT11              ----------------------------------------------------   7174-7853
OR.5110           ----------------------------------------------------   7480-7843
NM.T455           ----------------------------------------------------   7480-7843
GALUC64           --T-------------------------------------------------   7480-7843
DT24              ----------------------------------------------------   7174-7853
HDS85.06/12/90    ----------------------------------------------------   7447-7832
CASKI             ----------------------------------------------------   7480-7843
OR.4724           ----------------------------------------------------   7480-7843
OR.4997           ----------------------------------------------------   7480-7843
OR.6311           ----------------------------------------------------   7480-7843
OR.8329           ----------------------------------------------------   7480-7843
OR.8987           ----------------------------------------------------   7480-7843
NM.T197           ----------------------------------------------------   7480-7843
OR.9237           ----------------------------------------------------   7480-7843
NM.T446           ----------------------------------------------------   7480-7843
OR.0198           ----------------------------------------------------   7480-7843
SIHA              ----------------------------------------------------   7480-7843
DT42              ----------------------------------------------------   7174-7853
DC255             ----------------------------------------------------   7174-7853
DC212             ----------------------------------------------------   7174-7853
DC141             ----------------------------------------------------   7174-7853
DC269             ----------------------------------------------------   7174-7853
DC207             ----------------------------------------------------   7174-7853
BT7               ----------------------------------------------------   7174-7853
BT8               ----------------------------------------------------   7174-7853
BT15              ----------------------------------------------------   7174-7853
SB14              ----------------------------------------------------   7480-7843
J-5=SEQH          ----------------------------------------------------   7480-7843
OR.7574           ----------------------------------------------------   7480-7843
SB1               ----------------------------------------------------   7480-7843
SB7               ----------------------------------------------------   7480-7843
OR.2087           ----------------------------------------------------   7480-7843
OR.5428           ----------------------------------------------------   7480-7843
OR.4716           ----------------------------------------------------   7480-7843
SB13              ----------------------------------------------------   7480-7843
BAVINC0           ----------------------------------------------------   7157-7906
SA-9=SEQA         ----------------------------------------------------   7480-7843
SH-2a=SEQK        ----------------------------------------------------   7480-7843
AP-4=SEQI         ----------------------------------------------------   7480-7843
GB13              ----------------------------------------------------   7480-7843
IND-4=SEQG        ----------------------------------------------------   7480-7843
GB10              ----------------------------------------------------   7480-7843
GB21A             ----------------------------------------------------   7480-7843
BB24              ----------------------------------------------------   7480-7843
SS1A              ----------------------------------------------------   7480-7843
IND-5=SEQJ        ----------------------------------------------------   7480-7843
BT20              ----------------------------------------------------   7174-7853
BT10              ----------------------------------------------------   7174-7853
BT12              ----------------------------------------------------   7174-7853
BT9               ----------------------------------------------------   7174-7853
SB5AA             ------------------------G---------------------------   7480-7843
BT23              ----------------------------------------------------   7174-7853
DT4               ----------------------------------------------------   7174-7853
BB3               ----------------------------------------------------   7480-7843
BT22              ----------------------------------------------------   7174-7853
IND-7=SEQF        ----------------------------------------------------   7480-7843
```

LCR cont'd.

deletion

```
                 7777777777777777777777777777777777777777777777777
                 6666666666666666666666666666666666666666666666666
                 4444444455555555566666666667777777777888888888999
                 2345678901234567890123456789012345678901234567890
HPV16R           ATGTACATTGTGTCATATAAAATAAATCACTATGCGCCAACGCCTTACAT   7157-7906
GALUC153         ---------------------------------------------------   7480-7843
GALUC136         ---------------------------------------------------   7480-7843
GALUC73          ---------------------------------------------------   7480-7843
SL-6a=SEQD       ---------------------------------------------------   7480-7843
TS6A             -----------------------------------------------A--   7480-7843
TB13             ---------------------------------------------------   7480-7843
TS6B             ------G----------------------------------------A--   7480-7843
OR.1905          ---------------------------------------------A--   7480-7843
TB4              ---------------------------------------------A--   7480-7843
OR.7587          ---------------------------------------------A--   7480-7843
TB8              ---------------------------------------------A--   7480-7843
GALUC91          ---------------------------------------------A--   7480-7843
OR.6106          ---------------------------------------------A--   7480-7843
OR.7632          ---------------------------------------------A--   7480-7843
TB16             --------------------------T------------------A--   7480-7843
TB15             --------------------------T------------------A--   7480-7843
OR.3136          --------------------------T------------------A--   7480-7843
SB21             --------------------------T---------------------   7480-7843
BB11             --------------------------T------------------A--   7480-7843
AM-6=SEQB        --------------------------T------------------A--   7480-7843
OR.5691          --------------------------T------------------A--   7480-7843
OR.1783          --------------------------T------------------A--   7480-7843
AN-12=SEQE       --------------------------T------------------A--   7480-7843
OR.4541          --------------------------T------------------A--   7480-7843
BB1              --------------------------T------------------A--   7340-7843
HDS81A.10/04/89  --------------------------T------------------A--   7447-7832
HDS81A.11/07/90  --------------------------T------------------A--   7447-7832
HDS81A.06/19/91  --------------------------T------------------A--   7447-7832
HDS81A.02/24/92  --------------------------T------------------A--   7447-7832
HDS109.07/25/90  --------------------------T------------------A--   7447-7832
HDS109.11/30/90  --------------------------T------------------A--   7447-7832
NM.T529          --------------------------T------------------A--   7480-7843
NM.9999          --------------------------T------------------A--   7480-7843
OR.7908          --------------------------T------------------A--   7480-7843
OR.8863          --------------------------T------------------A--   7480-7843
OR.7754          --------------------------T------------------A--   7480-7843
OR.4072          --------------------------T------------------A--   7480-7843
OR.4451          --------------------------T------------------A--   7480-7843
OR.8160          --------------------------T------------------A--   7480-7843
NM.4094          --------------------------T------------------A--   7480-7843
SB16             --------------------------T------------------A--   7480-7843
OR.7875          --------------------------T------------------A--   7480-7843
BB19             --------------------------T------------------A--   7480-7843
BB14             --------------------------T------------------A--   7480-7843
P-1=SEQC         --------------------------T------------------A--   7480-7843
OR.8392                                                        A--   7689-7843
BAVINC3          ------G------------------T------------------A--   7157-7906
TS5              -G---------------------------------------------A--   7480-7843
TS5C             -G---T-------------------T------------------A--   7480-7843
TS5A             -G-------------------------T------------------A--   7480-7843
TS5B             -G-------------------------T------------------A--   7480-7843
TS3A             --------------------------T------------------A--   7480-7843
SN-4=SEQL        --------------------------T------------------A--   7480-7843
TS3C             --------------------------T------------------A--   7480-7843
TS3B             --------------------------T------------------A--   7480-7843
TB1              --------------------------T------------------A--   7480-7843
TS3              --------------------------T------------------A--   7480-7843
OR.3473          --------------------------T------------------A--   7480-7843
OR.3759          --------------------------T------------------A--   7480-7843
OR.7145          --------------------------T---------------------   7480-7843
AN-10a=SEQN      --------------------------T---------------------   7480-7843
SB21A            --------------------------T---------------------   7480-7843
AN-10=SEQM       --------------------------T---------------------   7480-7843
GALUC106         --------------------------T------------------A--   7480-7843
```

LCR cont'd.

```
                7777777777777777777777777777777777777777777777777777
                6666666677777777777777777777777777777777777777777777
                9999999900000000001112344455555566666666677777777
                2345678901234567890349023925678901234567890123456
HPV16R          ACCGCTGTTAGGCACATATTTAAATAATTTAAACTTCTAAGGCCAACTAA  7157-7906
OR.6170         ---------------------------------------------------  7480-7843
BB8A            .................----------------------------------  7480-7843
BT6             ---------------------------------------------------  7174-7853
HDS117A.04/09/92                            ----------------------  7750-7906
HDS117A.08/12/92                            ----------------------  7750-7906
HDS84C.03/26/90                             ----------------------  7750-7906
HDS84C.07/23/90                             ----------------------  7750-7906
HDS106C.05/08/90                            ----------------------  7750-7906
HDS138.05/01/91                             ----------------------  7750-7906
HDS81B.10/04/89 ---------------------------------------------------  7447-7832
HDS81B.11/07/90 ---------------------------------------------------  7447-7832
HDS81B.06/19/91 ---------------------------------------------------  7447-7832
HDS81B.02/24/92 ---------------------------------------------------  7447-7832
HDS69.10/16/89  ---------------------------------------------------  7447-7832
BT11            ---------------------------------------------------  7174-7853
OR.5110         ---------------------------------------------------  7480-7843
NM.T455         ---------------------------------------------------  7480-7843
HDS84D.07/23/90                             ----------------------  7750-7906
GALUC64         ---------------------------------------------------  7480-7843
DT24            ---------------------------------------------------  7174-7853
HDS85.06/12/90  ---------------------------------------------------  7447-7832
CASKI           ---------------------------------------------------  7480-7843
OR.4724         ---------------------------------------------------  7480-7843
OR.4997         ---------------------------------------------------  7480-7843
OR.6311         ---------------------------------------------------  7480-7843
OR.8329         ---------------------------------------------------  7480-7843
OR.8987         ---------------------------------------------------  7480-7843
NM.T197         ---------------------------------------------------  7480-7843
OR.9237         ---------------------------------------------------  7480-7843
NM.T446         ---------------------------------------------------  7480-7843
OR.0198         ---------------------------------------------------  7480-7843
SIHA            ----------------------------.....................  7480-7843
DT42            ---------------------------------------------------  7174-7853
DC255           ---------------------------------------------------  7174-7853
DC212           ---------------------------------------------------  7174-7853
DC141           ---------------------------------------------------  7174-7853
DC269           ---------------------------------------------------  7174-7853
DC207           ---------------------------------------------------  7174-7853
BT7             ---------------------------------------------------  7174-7853
BT8             ---------------------------------------------------  7174-7853
BT15            ---------------------------------------------------  7174-7853
HDS117B.04/09/92                            ----------------------  7750-7906
HDS84A.03/26/90                             ----------------------  7750-7906
SB14            ---------------------------------------------------  7480-7843
J-5=SEQH        ----------------------C--------C-------------------  7480-7843
OR.7574         ----------------------C---------------------------  7480-7843
SB1             ----------------------C---------------------------  7480-7843
SB7             ----------------------C---------------------------  7480-7843
OR.2087         ----------------------C---------------------------  7480-7843
OR.5428         ----------------------C---------------------------  7480-7843
OR.4716         ----------------------C---------------------------  7480-7843
SB13            ---------------------------------------------------  7480-7843
HDS106A.02/13/90                            ----------------------  7750-7906
HDS106A.05/08/90                            ----------------------  7750-7906
BAVINC0         ----------------------.---------------------------  7157-7906
SA-9=SEQA       ---------------------------------------------------  7480-7843
SH-2a=SEQK      ------------------------------------------T--------  7480-7843
AP-4=SEQI       ---------------------------------------------------  7480-7843
GB13            -------------------G-------------------------------  7480-7843
IND-4=SEQG      ------------------------------------G------------  7480-7843
GB10            ---------------------------------------------------  7480-7843
GB21A           ----------------------------------------------G----  7480-7843
BB24            --------------------------C-----------------------  7480-7843
SS1A            ---------------------------------------------------  7480-7843
HDS84B.03/26/90                             ----------------------  7750-7906
IND-5=SEQJ      ---------------------------------------------------  7480-7843
BT20            ---------------------------------------------------  7174-7853
BT10            ---------------------------------------------------  7174-7853
BT12            ---------------------------------------------------  7174-7853
BT9             ---------------------------------------------------  7174-7853
HDS84A.07/23/90                             ----------------------  7750-7906
SB5AA           ---------------------------------------------------  7480-7843
```

LCR cont'd.

```
                                deletion                d                deletion
                    ┌──────────────────────┐            ┌┐   ┌──────────────────────────────┐
                    77777777777777777777777777777777777777777777777777777
                    66666666777777777777777777777777777777777777777777777
                    99999999000000000011123444555555666666666677777777
                    23456789012345678903490239256789012345678901234567
HPV16R              ACCGCTGTTAGGCACATATTTAAATAATTTAAACTTCTAAGGCCAACTAA   7157-7906
HDS106B.02/13/90                                -----------------------   7750-7906
BT23                ----------------------------------------------------   7174-7853
DT4                 ----------------------------------------------------   7174-7853
BB3                 ----------------------------------------------------   7480-7843
BT22                ----------------------------------------------------   7174-7853
IND-7=SEQF          --------------------G-----------------------------   7480-7843
GALUC153            ----------------------C---------------------------   7480-7843
GALUC136            ----------------------------------------------------   7480-7843
GALUC73             ----------------------------------------------------   7480-7843
SL-6a=SEQD          --------------------A---------------T-------------   7480-7843
TS6A                --------------------A---------------T-------------   7480-7843
TB13                --------------------A---------------T-------------   7480-7843
TS6B                --------------------A---------------T-------------   7480-7843
OR.1905             --------------------A---------------T-------------   7480-7843
TB4                 --------------------A---------------T-------------   7480-7843
OR.7587             -----------------------------------T-------------   7480-7843
TB8                 -----------------------------------T-------------   7480-7843
GALUC91             -----------------------------------T-------------   7480-7843
OR.6106             -----------------------------------T-------------   7480-7843
OR.7632             -----------------------------------T-------------   7480-7843
TB16                -----------------------------------T-------------   7480-7843
TB15                -----------------------------------T-------------   7480-7843
OR.3136             ----------------------C--------------T-------------   7480-7843
SB21                -----------------------------------T-------------   7480-7843
BB11                ---------------------T--G-----------T-------------   7480-7843
AM-6=SEQB           ---------------------T--G-----------T-------------   7480-7843
OR.5691             ----------------------C--------------T-------------   7480-7843
OR.1783             ----------------------C--------------T-------------   7480-7843
AN-12=SEQE          ----------------------C--------------T-------------   7480-7843
OR.4541             ----------------------C--G-C---------T-------------   7480-7843
BB1                 ----------------------C--G-----------T-------------   7340-7843
HDS81A.10/04/89     ----------------------C--G-----------T-------------   7447-7832
HDS81A.11/07/90     ----------------------C--G-----------T-------------   7447-7832
HDS81A.06/19/91     ----------------------C--G-----------T-------------   7447-7832
HDS81A.02/24/92     ----------------------C--G-----------T-------------   7447-7832
HDS109.07/25/90     ----------------------C--G-----------T-------------   7447-7832
HDS109.11/30/90     ----------------------C--G-----------T-------------   7447-7832
NM.T529             ----------------------C--G-----------T-------------   7480-7843
NM.9999             ----------------------C--G-----------T-------------   7480-7843
OR.7908             ----------------------C--G-----------T-------------   7480-7843
OR.8863             ----------------------C--G-----------T-------------   7480-7843
OR.7754             ----------------------C--G-----------T-------------   7480-7843
OR.4072             ----------------------C--G-----------T-------------   7480-7843
OR.4451             ----------------------C--G-----------T-------------   7480-7843
OR.8160             ----------------------C--G-----------T-------------   7480-7843
NM.4094             ----------------------C--G-----------T-------------   7480-7843
SB16                ----------------------C--G-----------T-------------   7480-7843
OR.7875             ----------------------C--G-----------T-------------   7480-7843
BB19                ----------------------C--G-----------T-------------   7480-7843
BB14                ----------------------C--------------T-------------   7480-7843
P-1=SEQC            ----------------------T--------------T-------------   7480-7843
OR.8392             ----------------------T--------------T-------------   7689-7843
BAVINC3             ----------------------C-G------------T-------------   7157-7906
TS5                 -----------------------------------T-------------   7480-7843
TS5C                -----------------------------------T-------------   7480-7843
TS5A                -----------------------------------T-------------   7480-7843
TS5B                -----------------------------------T-------------   7480-7843
TS3A                -----------------------------------T-------------   7480-7843
SN-4=SEQL           -----------------------------------T-------------   7480-7843
TS3C                -----------------------------------T-------------   7480-7843
TS3B                ------------------------G----------T-------------   7480-7843
TB1                 -----------------------------------T-------------   7480-7843
TS3                 -----------------------------------T-------------   7480-7843
OR.3473             -----------------------------------T-------------   7480-7843
OR.3759             -----------------------------------T-------------   7480-7843
OR.7145             -----------------------------------T-------------   7480-7843
AN-10a=SEQN         -----------------------------------T-------------   7480-7843
SB21A               -----------------------------------T-------------   7480-7843
AN-10=SEQM          -----------------------------------T-------------   7480-7843
GALUC106            ----------------------C--------------T-------------   7480-7843
```

LCR cont'd.

```
                        deletion          deletion
                        ┌──────┐          ┌──────┐
                        7777777777777777777777777777777777
                        7777777777777777788888888888888888
                        7788888888889999900000000012333468
                        8901234567890123901234567892647928 6
HPV16R                  ATGTCACCCTAGTTCAGAACTGTGTAAGGGAAGGC   7157-7906
OR.6170                 -----------------------------------   7480-7843
BB8A                    -----------------------------------   7480-7843
BT6                     -----------------------------------   7174-7853
HDS117A.04/09/92        -----------------------------------   7750-7906
HDS117A.08/12/92        -----------------------------------   7750-7906
HDS84C.03/26/90         -----------------------------------   7750-7906
HDS84C.07/23/90         -----------------------------------   7750-7906
HDS106C.05/08/90        -----------------------------------   7750-7906
HDS138.05/01/91         -----------------------------------   7750-7906
HDS81B.10/04/89         ----------------------------          7447-7832
HDS81B.11/07/90         ----------------------------          7447-7832
HDS81B.06/19/91         ----------------------------          7447-7832
HDS81B.02/24/92         ----------------------------          7447-7832
HDS69.10/16/89          ----------------------------          7447-7832
BT11                    -----------------------------------   7174-7853
OR.5110                 -----------------------------------   7480-7843
NM.T455                 -----------------------------------   7480-7843
HDS84D.07/23/90         -----------------------------------   7750-7906
GALUC64                 -----------------------------------   7480-7843
DT24                    -----------------------------------   7174-7853
HDS85.06/12/90          ----------------------------          7447-7832
CASKI                   -----------------------------------   7480-7843
OR.4724                 -----------------------------------   7480-7843
OR.4997                 -----------------------------------   7480-7843
OR.6311                 -----------------------------------   7480-7843
OR.8329                 -----------------------------------   7480-7843
OR.8987                 -----------------------------------   7480-7843
NM.T197                 -----------------------------------   7480-7843
OR.9237                 -----------------------------------   7480-7843
NM.T446                 -----------------------------------   7480-7843
OR.0198                 -----------------------------------   7480-7843
SIHA                    ................----------------      7480-7843
DT42                    -----------------------------------   7174-7853
DC255                   -----------------------------------   7174-7853
DC212                   -----------------------------------   7174-7853
DC141                   -----------------------------------   7174-7853
DC269                   -----------------------------------   7174-7853
DC207                   -----------------------------------   7174-7853
BT7                     -----------------------------------   7174-7853
BT8                     -----------------------------------   7174-7853
BT15                    -----------------------------------   7174-7853
HDS117B.04/09/92        -----------------------------------A-  7750-7906
HDS84A.03/26/90         -----------------------------------A-  7750-7906
SB14                    ------------------A------------A      7480-7843
J-5=SEQH                ---C--------------------------A      7480-7843
OR.7574                 ---C--------------------------A      7480-7843
SB1                     ---C--------------------------A      7480-7843
SB7                     -----------------------------A      7480-7843
OR.2087                 -----------------------------A      7480-7843
OR.5428                 -----------------------------A      7480-7843
OR.4716                 -----------------------------A      7480-7843
SB13                    -----------------------------A      7480-7843
HDS106A.02/13/90        -----------------------------A--    7750-7906
HDS106A.05/08/90        -----------------------------A--    7750-7906
BAVINC0                 -----------------------------------   7157-7906
SA-9=SEQA               -----------------------------------   7480-7843
SH-2a=SEQK              -----------------------------------   7480-7843
AP-4=SEQI               ----T------------------------------   7480-7843
GB13                    -----------------------------------   7480-7843
IND-4=SEQG              -----------------------------------   7480-7843
GB10                    -----------------------------------   7480-7843
GB21A                   -----------------------------------   7480-7843
BB24                    G----------------------------------   7480-7843
SS1A                    -----------------------------------   7480-7843
HDS84B.03/26/90         --------------T------------------A-   7750-7906
IND-5=SEQJ              --------------T------------------     7480-7843
BT20                    -----------------------------------   7174-7853
BT10                    -----------------------------------   7174-7853
BT12                    -----------------------------------   7174-7853
BT9                     -----------------------------------   7174-7853
HDS84A.07/23/90         -----------------------------------A-  7750-7906
```

LCR cont'd.

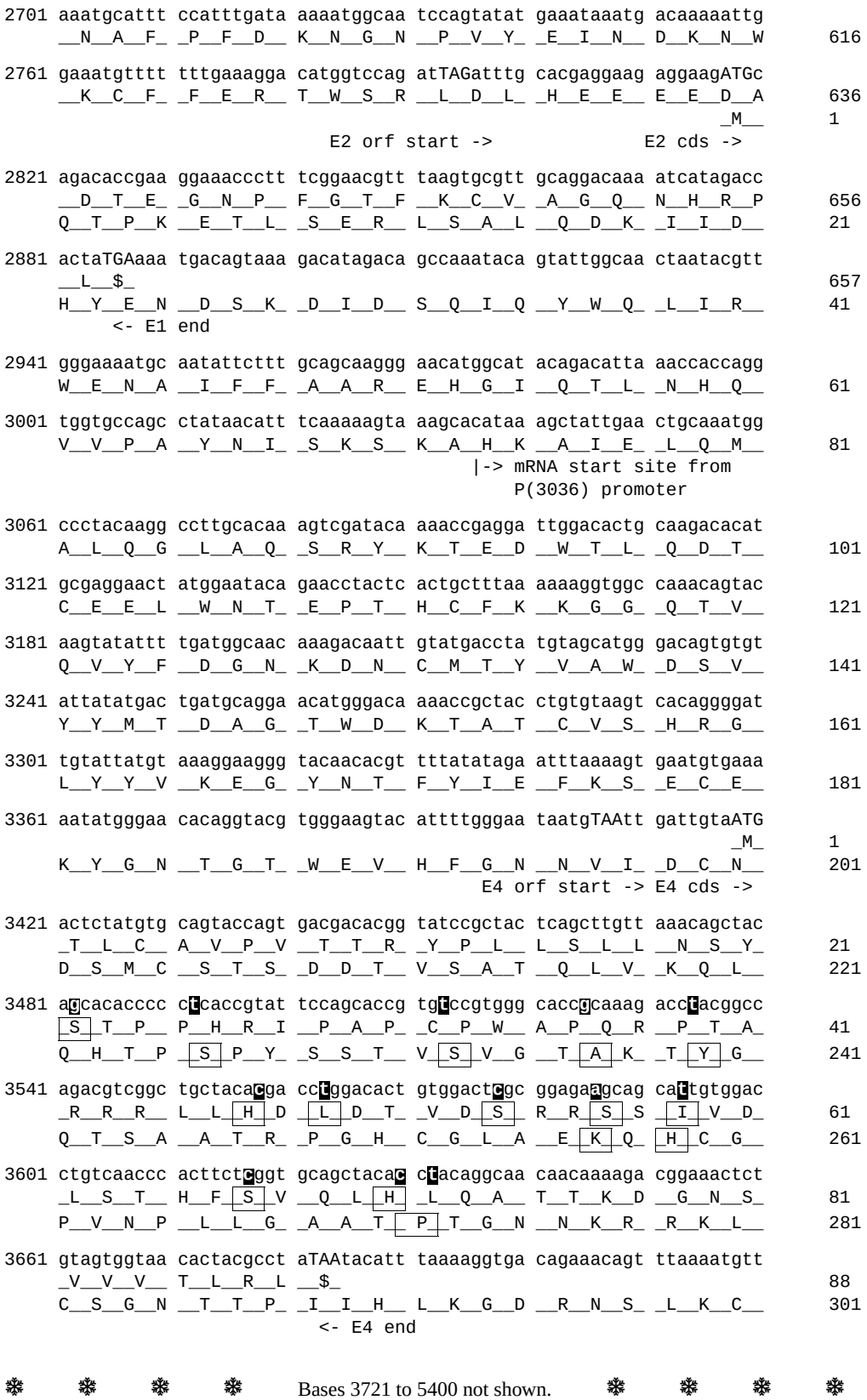
```
                        deletion          deletion
                        ^^^^^^^^^^^^^^^^^^^^^^^^^d
                        7777777777777777777777777777777777777
                        7777777777777777788888888888888888888
                        7788888888889999900000000012333468
                        8901234567890123901234567892647928 6
HPV16R                  ATGTCACCCTAGTTCAGAACTGTGTAAGGGAAGGC   7157-7906
SB5AA                   -----------------------------------   7480-7843
HDS106B.02/13/90        -----------------------------------A-  7750-7906
BT23                    -----------------------------------   7174-7853
DT4                     -----------------------------------   7174-7853
BB3                     -----------------------------------   7480-7843
BT22                    -----------------------------------   7174-7853
IND-7=SEQF              -----------------------------------   7480-7843
GALUC153                ---C-------------------------------   7480-7843
GALUC136                -----------------------------------   7480-7843
GALUC73                 -----------------------------------   7480-7843
SL-6a=SEQD              -----------------------------------   7480-7843
TS6A                    --------T--------------------T---     7480-7843
TB13                    --------T--------------------T---     7480-7843
TS6B                    --------T--------------------T---     7480-7843
OR.1905                 --------T--------------------T---     7480-7843
TB4                     --------T--------------------T---     7480-7843
OR.7587                 --------T--------------------T---     7480-7843
TB8                     --------T--------------------T---     7480-7843
GALUC91                 --------T------------------------     7480-7843
OR.6106                 --------T--------------------T---     7480-7843
OR.7632                 --------T--------------------T---     7480-7843
TB16                    --------T--------------------T---     7480-7843
TB15                    --------T-------............-T---     7480-7843
OR.3136                 --------T-------------------AT---     7480-7843
SB21                    --------T------------------------     7480-7843
BB11                    --------T------------------------     7480-7843
AM-6=SEQB               --------T------------------------     7480-7843
OR.5691                 --------T------------------------     7480-7843
OR.1783                 --------T------------------------     7480-7843
AN-12=SEQE              --------T------------------------     7480-7843
OR.4541                 --------T------------------------     7480-7843
BB1                     --------T------------------------     7340-7843
HDS81A.10/04/89         --------T--------------------         7447-7832
HDS81A.11/07/90         --------T--------------------         7447-7832
HDS81A.06/19/91         --------T--------------------         7447-7832
HDS81A.02/24/92         --------T--------------------         7447-7832
HDS109.07/25/90         --------T--------------------         7447-7832
HDS109.11/30/90         --------T--------------------         7447-7832
NM.T529                 --------T------------------------     7480-7843
NM.9999                 --------T------------------------     7480-7843
OR.7908                 --------T------------------------     7480-7843
OR.8863                 --------T------------------------     7480-7843
OR.7754                 --------T------------------------     7480-7843
OR.4072                 --------T------------------------     7480-7843
OR.4451                 --------T------------------------     7480-7843
OR.8160                 --------T------------------------     7480-7843
NM.4094                 --------T------------------------     7480-7843
SB16                    --------T------------------------     7480-7843
OR.7875                 --------T------------------------     7480-7843
BB19                    --------T------------------------     7480-7843
BB14                    --------T------------------------     7480-7843
P-1=SEQC                --------T------------------------     7480-7843
OR.8392                 --------T------------------------     7689-7843
BAVINC3                 --------T------------------------G    7157-7906
TS5                     --------T-------------------ATT--     7480-7843
TS5C                    --------T-------------------ATT--     7480-7843
TS5A                    --------T-------------------ATT--     7480-7843
TS5B                    --------T-------------------ATT--     7480-7843
TS3A                    --------T-------------------ATTG-     7480-7843
SN-4=SEQL               --------T-------------------A-TG-     7480-7843
TS3C                    --------T-------------------ATTG-     7480-7843
TS3B                    --------T-------------------ATTG-     7480-7843
TB1                     --------T-------------------ATTG-     7480-7843
TS3                     --------T-------------------ATTG-     7480-7843
OR.3473                 --------T-------------------ATCG-     7480-7843
OR.3759                 --------T-------------------ATCG-     7480-7843
OR.7145                 --------T-------------------ATCG-     7480-7843
AN-10a=SEQN             --------T-------------------ATTG-     7480-7843
SB21A                   --------T-------------------ATTG-     7480-7843
AN-10=SEQM              --------T-------------------ATTG-     7480-7843
GALUC106                --------T--------------------T---     7480-7843
```

```
LOCUS       HPV18        7857 bp ds-DNA             VRL       11-DEC-1992
DEFINITION  Human papillomavirus type 18 (HPV-18), complete genome.
ACCESSION   X05015
SOURCE      Human papillomavirus type 18 DNA recovered from a cervical
            carcinoma of a Brazilian patient.
  ORGANISM  Human papillomavirus type 18
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
            Genetic Maps; Locus Maps of Complex Genomes:  1-1,
            Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1993)
COMMENT     The sequence of L1 for the reference isolate of HPV18 L1 has been
            called into question by Hofmann et al., J. Gen. Virol. 77:465-468,
            1996; however, the authors of this paper have not made their 18L1
            sequences available to the gene libraries or to this database.
            Naturally occurring variants of HPV18 have been observed throughout
            the genome. Further information on the variants presented here is
            given below. See also "Sources of Variants" at the
            end of the Variants section of Part I.
            Reference: Hecht,J.L., Kadish,A.S., Jing,G., and Burk,R.D.
                Int. J. Cancer 60, 369-76 (1995)
                Sequences constructed by editing reference sequence according
                to figures in article.
            Reference: Inagaki,Y., Tsunokawa,Y., Takebe,N., Nawa,H., Nakanishi,S.,
                Terada,M., and Sugimura,T., J. Virol. 62, 1640-6 (1988).
                Variant Accession Nos. M20324 M20325
            Reference: Matlashewski,G., Banks,L., Wu-Liao,J., Spence,P., Pim,D., and
                Crawford,L., J. Gen. Virol. 67, 1909-16 (1986).
                Variant Accession Nos. X04354
            Reference: Ong,C.-K., Chan,S.Y., Campo,M.S., Fujinaga,K., Mavromara-Nazos,P.,
                Labropoulou,V., Pfister,H., Tay,S.K., et al., J. Virol. 67, 6424-31 (1993)
                Variant Accession Nos. L22615-L22635
            Reference: Ong,C.-K., Nee, S. Rambaut, A., Bernard, H.-U., Harvey, P. H.,
                Unpublished (1996)
                Variant Accession Nos. U59156-U59165
            Reference: Schneider-Gadicke,A., and Schwarz,E.,
                EMBO J. 5, 2285-92 (1986).
                Variant Accession Nos. M26798
            Reference: Seedorf,K., Oltersdorf,T., Krammer,G., and Rowekamp,W.,
                EMBO J. 6, 139-44 (1987).
                Variant Accession Nos. X04773
            Reference: Stewart,A.-C.M., Eriksson,A.M., Manos,M.M., Munoz,N., Bosch,F.X.,
                Peto,J. and Wheeler,C.M., J. Virol. 70, 3127-36 (1996).
                Variant Accession Nos. U45889-U45894
            Reference: ter Meulen,J. Schweigler,A.C., Eberhardt,H.C., et al.
                Int. J. Cancer 53, 257-9 (1993).
                Sequences constructed by editing reference sequence according to
                figures in article.
FEATURES             Location/Qualifiers
     CDS             105..581
                     /note="E6 ORF from bp 87 to 581"
                     /product="transforming protein"
                     /gene="E6"
                     /note="putative"
                     /codon_start=1
                     /translation="MARFEDPTRRPYKLPDLCTELNTSLQDIEITCVYCKTVLELTEV
                     FEFAFKDLFVVYRDSIPHAACHKCIDFYSRIRELRHYSDSVYGDTLEKLTNTGLYNLL
                     IRCLRCQKPLNPAEKLRHLNEKRRFHNIAGHYRGQCHSCCNRARQERLQRRRETQV"
     CDS             590..907
                     /note="E7 ORF from bp 509 to 907"
                     /product="transforming protein"
                     /gene="E7"
                     /note="putative"
                     /codon_start=1
                     /translation="MHGPKATLQDIVLHLEPQNEIPVDLLCHEQLSDSEEENDEIDGV
```

```
                     NHQHLPARRAEPQRHTMLCMCCKCEARIELVVESSADDLRAFQQLFLNTLSFVCPWCA
                     SQQ"
     CDS             914..2887
                     /note="E1 ORF from 908 to 2887"
                     /product="replication protein"
                     /gene="E1"
                     /note="putative"
                     /codon_start=1
                     /translation="MADPEGTDGEGTGCNGWFYVQAIVDKKTGDVISDDEDENATDTG
                     SDMVDFIDTQGTFCEQAELETAQALFHAQEVHNDAQVLHVLKRKFAGGSTENSPLGER
                     LEVDTELSPRLQEISLNSGQKKAKRRLFTISDSGYGCSEVEATQIQVTTNGEHGGNVC
                     SGGSTEAIDNGGTEGNNSSVDGTSDNSNIENVNPQCTIAQLKDLLKVNNKQGAMLAVF
                     KDTYGLSFTDLVRNFKSDKTTCTDWVTAIFGVNPTIAEGFKTLIQPFILYAHIQCLDC
                     KWGVLILALLRYKCGKSRLTVAKGLSTLLHVPETCMLIQPPKLRSSVAALYWYRTGIS
                     NISEVMGDTPEWIQRLTIIQHGIDDSNFDLSEMVQWAFDNELTDESDMAFEYALLADS
                     NSNAAAFLKSNCQAKYLKDCATMCKHYRRAQKRQMNMSQWIRFRCSKIDEGGDWRPIV
                     QFLRYQQIEFITFLGALKSFLKGTPKKNCLVFCGPANTGKSYFGMSFIHFIQGAVISF
                     VNSTSHFWLEPLTDTKVAMLDDATTTCWTYFDTYMRNALDGNPISIDRKHKPLIQLKC
                     PPILLTTNIHPAKDNRWPYLESRITVFEFPNAFPFDKNGNPVYEINDKNWKCFFERTW
                     SRLDLHEEEEDADTEGNPFGTFKCVAGQNHRPL"
     CDS             2817..3914
                     /note="E2 ORF from bp 2796 to 3914"
                     /product="regulatory protein"
                     /gene="E2"
                     /note="putative"
                     /codon_start=1
                     /translation="MQTPKETLSERLSALQDKIIDHYENDSKDIDSQIQYWQLIRWEN
                     AIFFAAREHGIQTLNHQVVPAYNISKSKAHKAIELQMALQGLAQSRYKTEDWTLQDTC
                     EELWNTEPTHCFKKGGQTVQVYFDGNKDNCMTYVAWDSVYYMTDAGTWDKTATCVSHR
                     GLYYVKEGYNTFYIEFKSECEKYGNTGTWEVHFGNNVIDCNDSMCSTSDDTVSATQLV
                     KQLQHTPSPYSSTVSVGTAKTYGQTSAATRPGHCGLAEKQHCGPVNPLLGAATPTGNN
                     KRRKLCSGNTTPIIHLKGDRNSLKCLRYRLRKHSDHYRDISSTWHWTGAGNEKTGILT
                     VTYHSETQRTKFLNTVAIPDSVQILVGYMTM"
     CDS             3418..3684
                     /note="E4 ORF from bp 3409 to 3684"
                     /gene="E4"
                     /note="putative"
                     /codon_start=1
                     /translation="MTLCAVPVTTRYPLLSLLNSYSTPPHRIPAPCPWAPQRPTARRR
                     LLHDLDTVDSRRSSTVDLSTHFSVQLHLQATTKDGNSVVVTLRL"
     CDS             5430..7136
                     /note="L1 ORF from bp 5418 to 7136"
                     /product="major capsid protein"
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /translation="MCLYTRVLILHYHLLPLYGPLYHPRPLPLHSILVYMVHIIICGH
                     YIILFLRNVNVFPIFLQMALWRPSDNTVYLPPPSVARVVNTDDYVTPTSIFYHAGSSR
                     LLTVGNPYFRVPAGGGNKQDIPKVSAYQYRVFRVQLPDPNKFGLPDTSIYNPETQRLV
                     WACAGVEIGRGQPLGVGLSGHPFYNKLDDTESSHAATSNVSEDVRDNVSVDYKQTQLC
                     ILGCAPAIGEHWAKGTACKSRPLSQGDCPPLELKNTVLEDGDMVDTGYGAMDFSTLQD
                     TKCEVPLDICQSICKYPDYLQMSADPYGDSMFFCLRREQLFARHFWNRAGTMGDTVPQ
                     SLYIKGTGMPASPGSCVYSPSPSGSIVTSDSQLFNKPYWLHKAQGHNNGVCWHNQLFV
                     TVVDTTPSTNLTICASTQSPVPGQYDATKFKQYSRHVEEYDLQFIFQLCTITLTADVM
                     SYIHSMNSSILEDWNFGVPPPPTTSLVDTYRFVQSVAITCQKDAAPAENKDPYDKLKF
                     WNVDLKEKFSLDLDQYPLGRKFLVQAGLRRKPTIGPRKRSAPSATTSSKPAKRVRVRA
                     RK"
BASE COUNT     2365 a   1497 c   1680 g   2315 t
```

```
     1 attaatactt ttaacaattg tagtatataa aaaagggagt aACCGAAAAC GGTcgggACC
                                          E2 bind ->       E2 bind ->

    61 GAAAACGGTg tatataaaag atgTGAgaaa cacaccacaa tactATGgcg cgctttgagg
                                                      _M__A_ _R__F__E__      5
                       E6 orf start ->           E6 cds ->
                                                      |-> mRNA start site from
                                                          P(105) promoter

   121 atccaacacg gcgaccctac aagctacctg atctgtgcac ggaactgaac acttcactgc
       D__P__T__R __R__P__Y_ _K__L__P__ D__L__C__T __E__L__N_ _T__S__L__      25

   181 aagacataga aataacctgt gtatattgca agacagtatt ggaacttaca gaggtatttg
       Q__D__I__E __I__T__C_ _V__Y__C__ K__T__V__L __E__L__T_ _E__V__F__      45

   241 aatttgcatt taaagattta tttgtggtgt atagagacag tataccccat gctgcatgcc
       E__F__A__F __K__D__L_ _F__V__V__ Y__R__D__S __I__P__H_ _A__A__C__      65

   301 ataaatgtat agatttttat tctagaatta gagaattaag acattattca gactctgtgt
       H__K__C__I __D__F__Y_ _S__R__I__ R__E__L__R __H__Y__S_ _D__S__V__      85

   361 atggagacac attggaaaaa ctaactaaca ctgggttata caatttatta ataaggtgcc
       Y__G__D__T __L__E__K_ _L__T__N__ T__G__L__Y __N__L__L_ _I__R__C__      105

   421 tgcggtgcca gaaaccgttg aatccagcag aaaaacttag acaccttaat gaaaaacgac
       L__R__C__Q __K__P__L_ _N__P__A__ E__K__L__R __H__L__N_ _E__K__R__      125

   481 gattgcacaa catagctggg cactaTAGag gccagtgcca ttcgtgctgc aaccgagcac
       R__F__H__N __I__A__G_ _H__Y__R__ G__Q__C__H __S__C__C_ _N__R__A__      145
                         E7 orf start ->

   541 gacaggaacg actccaacga cgcagagaaa cacaagtaTA AtattaagtA TGcatggacc
                                                      _ M__H__G__P      4
       R__Q__E__R __L__Q__R_ _R__R__E__ T__Q__V__$ _                        158
                                                <- E6 end -> E7 cds

   601 taaggcaaca ttgcaagaca ttgtattgca tttagagccc caaaatgaaa ttccggttga
       __K__A__T_ _L__Q__D__ I__V__L__H __L__E__P_ _Q__N__E__ I__P__V__D      24

   661 ccttctatgt cacgagcaat taagcgactc agaggaagaa aacgatgaaa tagatggagt
       __L__L__C_ _H__E__Q__ L__S__D__S __E__E__E_ _N__D__E__ I__D__G__V      44

   721 taatcatcaa catttaccag cccgacgagc cgaaccacaa cgtcacacaa tgttgtgtat
       __N__H__Q_ _H__L__P__ A__R__R__A __E__P__Q_ _R__H__T__ M__L__C__M      64

   781 gtgttgtaag tgtgaagcca gaattgagct agtagtagaa agctcagcag acgaccttcg
       __C__C__K_ _C__E__A__ R__I__E__ L __V__V__E_ _S__S__A__ D__D__L__R      84

   841 agcattccag cagctgtttc tgaacaaccct gtcctttgtg tgtccgtggt gtgcatccca
       __A__F__Q_ _Q__L__F__ L__N__T__L __S__F__V_ _C__P__W__ C__A__S__Q      104

   901 gcagTAAgca acaATGgctg atccagaagg tacagacggg gagggcacgg gttgtaacgg
       __Q__$_       _M__A__ D__P__E__G __T__D__G_ _E__G__T__ G__C__N__G      105/16
           <- E7 end
E1 orf start ->     -> E1 cds

   961 ctggttttat gtacaagcta ttgtagacaa aaaaacagga gatgtaatat cagatgacga
       __W__F__Y_ _V__Q__A__ I__V__D__K __K__T__G_ _D__V__I__ S__D__D__E      36

  1021 ggacgaaaat gcaacagaca cagggtcgga tatggtagat tttattgata cacaaggaac
       __D__E__N_ _A__T__D__ T__G__S__D __M__V__D_ _F__I__D__ T__Q__G__T      56

  1081 attttgtgaa caggcagagc tagagacagc acaggcattg ttccatgcgc aggaggtcca
       __F__C__E_ _Q__A__E__ L__E__T__A __Q__A__L_ _F__H__A__ Q__E__V__H      76

  1141 caatgatgca caagtgttgc atgttttaaa acgaaagttt gcaggaggca gcacagaaaa
       __N__D__A_ _Q__V__L__ H__V__L__K __R__K__F_ _A__G__G__ S__T__E__N      96

  1201 cagtccatta ggggagcggc tggaggtgga tacagagtta agtccacggt tacaagaaat
       __S__P__L_ _G__E__R__ L__E__V__D __T__E__L_ _S__P__R__ L__Q__E__I      116
```

```
1261 atctttaaat agtgggcaga aaaaggcaaa aaggcggctg tttacaatat cagatagtgg
     __S__L__N_ _S__G__Q__ K__K__A__K __R__R__L_ _F__T__I__ S__D__S__G      136

1321 ctatggctgt tctgaagtgg aagcaacaca gattcaggta actacaaatg gcgaacatgg
     __Y__G__C_ _S__E__V__ E__A__T__Q __I__Q__V_ _T__T__N__ G__E__H__G      156

1381 cggcaatgta tgtagtggcg gcagtacgga ggctatagac aacgggggca cagagggcaa
     __G__N__V_ _C__S__G__ G__S__T__E __A__I__D_ _N__G__G__ T__E__G__N      176

1441 caacagcagt gtagacggta caagtgacaa tagcaatata gaaaatgtaa atccacaatg
     __N__S__S_ _V__D__G__ T__S__D__N __S__N__I_ _E__N__V__ N__P__Q__C      196

1501 taccatagca caattaaaag acttgttaaa agtaaacaat aaacaaggag ctatgttagc
     __T__I__A_ _Q__L__K__ D__L__L__K __V__N__N_ _K__Q__G__ A__M__L__A      216

1561 agtatttaaa gacacatatg ggctatcatt tacagattta gttagaaatt ttaaaagtga
     __V__F__K_ _D__T__Y__ G__L__S__F __T__D__L_ _V__R__N__ F__K__S__D      236

1621 taaaaccacg tgtacagatt gggttacagc tatatttgga gtaaacccaa caatagcaga
     __K__T__T_ _C__T__D__ W__V__T__A __I__F__G_ _V__N__P__ T__I__A__E      256

1681 aggatttaaa acactaatac agccatttat attatatgcc catattcaat gtctagactg
     __G__F__K_ _T__L__I__ Q__P__F__I __L__Y__A_ _H__I__Q__ C__L__D__C      276

1741 taaatgggga gtattaatat tagccctgtt gcgttacaaa tgtggtaaga gtagactaac
     __K__W__G_ _V__L__I__ L__A__L__L __R__Y__K_ _C__G__K__ S__R__L__T      296

1801 agttgctaaa ggtttaagta cgttgttaca cgtacctgaa actgtatgt taattcaacc
     __V__A__K_ _G__L__S__ T__L__L__H __V__P__E_ _T__C__M__ L__I__Q__P      316

1861 accaaaattg cgaagtagtg ttgcagcact atattggtat agaacaggaa tatcaaatat
     __P__K__L_ _R__S__S__ V__A__A__L __Y__W__Y_ _R__T__G__ I__S__N__I      336

1921 tagtgaagta atgggagaca cacctgagtg gatacaaaga cttactatta tacaacatgg
     __S__E__V_ _M__G__D__ T__P__E__W __I__Q__R_ _L__T__I__ I__Q__H__G      356

1981 aatagatgat agcaattttg atttgtcaga aatggtacaa tgggcatttg ataatgagct
     __I__D__D_ _S__N__F__ D__L__S__E __M__V__Q_ _W__A__F__ D__N__E__L      376

2041 gacagatgaa agcgatatgg catttgaata tgccttatta gcagacagca acagcaatgc
     __T__D__E_ _S__D__M__ A__F__E__Y __A__L__L_ _A__D__S__ N__S__N__A      396

2101 agctgccttt ttaaaaagca attgccaagc taaatattta aaagattgtg ccacaatgtg
     __A__A__F_ _L__K__S__ N__C__Q__A __K__Y__L_ _K__D__C__ A__T__M__C      416

2161 caaacattat aggcgagccc aaaaacgaca aatgaatatg tcacagtgga tacgatttag
     __K__H__Y_ _R__R__A__ Q__K__R__Q __M__N__M_ _S__Q__W__ I__R__F__R      436

2221 atgttcaaaa atagatgaag ggggagattg gagaccaata gtgcaattcc tgcgatacca
     __C__S__K_ _I__D__E__ G__G__D__W __R__P__I_ _V__Q__F__ L__R__Y__Q      456

2281 acaaatagag tttataacat ttttaggagc cttaaaatca ttttttaaaag gaacccccaa
     __Q__I__E_ _F__I__T__ F__L__G__A __L__K__S_ _F__L__K__ G__T__P__K      476

2341 aaaaaattgt ttagtatttt gtggaccagc aaatacagga aaatcatatt ttggaatgag
     __K__N__C_ _L__V__F__ C__G__P__A __N__T__G_ _K__S__Y__ F__G__M__S      496

2401 ttttatacac tttatacaag gagcagtaat atcatttgtg aattccacta gtcatttttg
     __F__I__H_ _F__I__Q__ G__A__V__I __S__F__V_ _N__S__T__ S__H__F__W      516

2461 gttggaaccg ttaacagata ctaaggtggc catgttagat gatgcaacga ccacgtgttg
     __L__E__P_ _L__T__D__ T__K__V__A __M__L__D_ _D__A__T__ T__T__C__W      536

2521 gacatacttt gatacctata tgagaaatgc gttagatggc aatccaataa gtattgatag
     __T__Y__F_ _D__T__Y__ M__R__N__A __L__D__G_ _N__P__I__ S__I__D__R      556

2581 aaagcacaaa ccattaatac aactaaaatg tcctccaata ctactaacca caaatataca
     __K__H__K_ _P__L__I__ Q__L__K__C __P__P__I_ _L__L__T__ T__N__I__H      576
                       |-> mRNA start site from P(2598) promoter

2641 tccagcaaag gataatagat ggccatattt agaaagtaga ataacagtat ttgaatttcc
     __P__A__K_ _D__N__R__ W__P__Y__L __E__S__R_ _I__T__V__ F__E__F__P      596
```



❄ ❄ ❄ ❄ Bases 3721 to 5400 not shown. ❄ ❄ ❄ ❄

```
5401 tgtaacggtc ccttTAAcct cctcttgggA TGtgcctgta tacacgggtc ctgatattac
     __V__T__V_ _P__L__T__ S__S__W__D __V__P__V_ _Y__T__G__ P__D__I__T      406
                                 _ M__C__L__Y __T__R__V_ _L__I__L__      10
          L1 orf start ->    L1 cds ->

5461 attaccatct actacctctg tatggcccat tgtatcaccc acggcccctg cctctacaca
     __L__P__S_ _T__T__S__ V__W__P__I __V__S__P_ _T__A__P__ A__S__T__Q      426
     H__Y__H__L __L__P__L_ _Y__G__P__ L__Y__H__P __R__P__L_ _P__L__H__      30

5521 gtatattggt atacatggta cacattatta tttgtggcca ttatattatt ttattcctaa
     __Y__I__G_ _I__H__G__ T__H__Y__Y __L__W__P_ _L__Y__Y__ F__I__P__K      446
     S__I__L__V __Y__M__V_ _H__I__I__ I__C__G__H __Y__I__I_ _L__F__L__      50

5581 gaaacgtaaa cgtgttccct atttttttgc agatggcttt gtggcggccT AGtgacaata
     __K__R__K_ _R__V__P__ Y__F__F__A __D__G__F_ _V__A__A__ $_             462
     R__N__V__N __V__F__P_ _I__F__L__ Q__M__A__L __W__R__P_ _S__D__N__      70
                      |-> mRNA start site from P(5600) promoter
                                                          <- L2 end

5641 ccgtatatct tccacctcct tctgtggcaa gagttgtaaa taccgatgat tatgtgactc
     T__V__Y__L __P__P__P_ _S__V__A__ R__V__V__N __T__D__D_ _Y__V__T__      90

5701 ccacaagcat attttatcat gctggcagct ctagattatt aactgttggt aatccatatt
     P__T__S__I __F__Y__H_ _A__G__S__ S__R__L__L __T__V__G_ _N__P__Y__      110

5761 ttagggttcc tgcaggtggt ggcaataagc aggatattcc taaggtttct gcataccaat
     F__R__V__P __A__G__G_ _G__N__K__ Q__D__I__P __K__V__S_ _A__Y__Q__      130

5821 atagagtatt tagggtgcag ttacctgacc caaataaatt tggtttacct gatactagta
     Y__R__V__F __R__V__Q_ _L__P__D__ P__N__K__F __G__L__P_ _D__T__S__      150

5881 tttataatcc tgaaacacaa cgtttagtgt gggcctgtgc tggagtggaa attggccgtg
     I__Y__N__P __E__T__Q_ _R__L__V__ W__A__C__A __G__V__E_ _I__G__R__      170

5941 gtcagccttt aggtgttggc cttagtgggc atccatttta taataaatta gatgacactg
     G__Q__P__L __G__V__G_ _L__S__G__ H__P__F__Y __N__K__L_ _D__D__T__      190

6001 aaagttccca tgccgccacg tctaatgttt ctgaggacgt tagggacaat gtgtctgtag
     E__S__S__H __A__A__T_ _S__N__V__ S__E__D__V __R__D__N_ _V__S__V__      210

6061 attataagca gacacagtta tgtattttgg gctgtgcccc tgctattggg gaacactggg
     D__Y__K__Q __T__Q__L_ _C__I__L__ G__C__A__P __A__I__G_ _E__H__W__      230

6121 ctaaaggcac tgcttgtaaa tcgcgtcctt tatcacaggg cgattgcccc cctttagaac
     A__K__G__T __A__C__K_ _S__R__P__ L__S__Q__G __D__C__P_ _P__L__E__      250

6181 ttaaaaacac agttttggaa gatggtgata tggtagatac tggatatggt gccatggact
     L__K__N__T __V__L__E_ _D__G__D__ M__V__D__T __G__Y__G_ _A__M__D__      270

6241 ttagtacatt gcaagatact aaatgtgagg taccattgga tatttgtcag tctatttgta
     F__S__T__L __Q__D__T_ _K__C__E__ V__P__L__D __I__C__Q_ _S__I__C__      290

6301 aatatcctga ttatttacaa atgtctgcag atccttatgg ggattccatg ttttttgct
     K__Y__P__D __Y__L__Q_ _M__S__A__ D__P__Y__G __D__S__M_ _F__F__C__      310

6361 tacggcgtga gcagcttttt gctaggcatt tttggaatag agcaggtact atgggtgaca
     L__R__R__E __Q__L__F_ _A__R__H__ F__W__N__R __A__G__T_ _M__G__D__      330

6421 ctgtgcctca atccttatat attaaaggca caggtatgcc tgcttcacct ggcagctgtg
     T__V__P__Q __S__L__Y_ _I__K__G__ T__G__M__P __A__S__P_ _G__S__C__      350

6481 tgtattctcc ctctccaagt ggctctattg ttacctctga ctcccagttg tttaataaac
     V__Y__S__P __S__P__S_ _G__S__I__ V__T__S__D __S__Q__L_ _F__N__K__      370

6541 catattggtt acataaggca cagggtcata acaatggtgt ttgctggcat aatcaattat
     P__Y__W__L __H__K__A_ _Q__G__H__ N__N__G__V __C__W__H_ _N__Q__L__      390

6601 ttgttactgt ggtagatacc actccagta ccaatttaac aatatgtgct tctacacagt
     F__V__T__V __V__D__T_ _T__P__S__ T__N__L__T __I__C__A_ _S__T__Q__      410
```

```
6661 ctcctgtacc tgggcagtat gatgctacca aatttaggca gtatagcaga catgttgagg
     S__P__V__P __G__Q__Y_ _D__A__T__ K__F__K__Q __Y__S__R_ _H__V__E__      430

6721 aatatgattt gcagtttatt tttcagttgt gtactattac tttaactgca gatgttatgt
     E__Y__D__L __Q__F__I_ _F__Q__L__ C__T__I__T __L__T__A_ _D__V__M__      450

6781 cctatattca tagtatgaat agcagtattt tagaggattg gaactttggt gttccccccc
     S__Y__I__H __S__M__N_ _S__S__I__ L__E__D__W __N__F__G_ _V__P__P__      470

6841 ccccaactac tagtttggtg gatacatatc gttttgtaca atctgttgct attacctgtc
     P__P__T__T __S__L__V_ _D__T__Y__ R__F__V__Q __S__V__A_ _I__T__C__      490

6901 aaaaggatgc tgcaccggct gaaaataagg atccctatga taagttaaag ttttggaatg
     Q__K__D__A __A__P__A_ _E__N__K__ D__P__Y__D __K__L__K_ _F__W__N__      510

6961 tggatttaaa ggaaaagttt tctttagact tagatcaata tccccttgga cgtaaatttt
     V__D__L__K __E__K__F_ _S__L__D__ L__D__Q__Y __P__L__G_ _R__K__F__      530

7021 tggttcaggc tggattgcgt cgcaagccca ccataggccc tcgcaaacgt tctgctccat
     L__V__Q__A __G__L__R_ _R__K__P__ T__I__G__P __R__K__R_ _S__A__P__      550

7081 ctgccactac gtcttctaaa cctgccaagc gtgtgcgtgt acgtgccagg aagTAAtatg
     S__A__T__T __S__S__K_ _P__A__K__ R__V__R__V __R__A__R_ _K__$_          568
                                                                <- L1 end

7141 tgtgtgtgta tatatatata catctattgt tgtgtttgta tgtcctgtgt ttgtgtttgt

7201 tgtatgattg cattgtatgg tatgtatggt tgttgttgta tgttgtatgt tactatattt

7261 gttggtatgt ggcattaaat aaaatatgtt ttgtggttct gtgtgttatg tggttgcgcc

7321 ctagtgagta acaactgtat ttgtgtttgt ggtatgggtg ttgcttgttg ggctatatat

7381 tgtcctgtat ttcaagttat aaaactgcac accttacagc atccatttta tcctacaatc

7441 ctccattttg ctgtgcaACC GATTTCGGTt gcctttggct tatgtgtgtg gttttgtgca
              E2 bind ->

7501 caatacagta cgctggcact attgcaaact ttaatgtttt gggcactgct cctacatatt

7561 ttgaacaatt ggcgcgcctc tttggcgcat attaaggcga cctggtatta gtcattttcc

7621 tgtccaggtg cgctacaaca attgcttgca taactatatc cactccctaa gtaataaaac

7681 tgcttttagg cacatatttt agttttgtttt tacttagct aattgcatac ttggcttgta

7741 caactacttt catgtccaac attctgtcta cccttaacat gaactataat atgactaagc

7801 tgtgcataca tagtttatgc aACCGAAATA GGTtgggcag cacatactat acttttc
                   E2 bind ->
```

E6 nuc

```
                1245
                6884
                9759
HPV18           ACTC   105-581
HPV18v2E67      -GCA   105-581
HPV18v3E67      -G--   105-581
HPV18v4E6       -G--   105-581
HPV18v5E67      GG-A   105-581
```

E6 aa

```
                0011
                2624
                2179
HPV18           NPFR
HPV18v2E67      -***
HPV18v3E67      -*--
HPV18v4E6       -*--
HPV18v5E67      S*-*
```

E7 nuc

	677888 457066 010645	
HPV18	CCAGAC	590-907
A4	------	629-907
A20	------	634-889
HPV18v3E671	------	590-907
4055	------	610-875
B17	--C---	618-900
4050	--C---	631-902
B6	T-----	631-890
SW756	T----A	629-891
101	-T----	618-893
161	-T----	618-893
173	-T----	618-893
C41	-T----	629-891
HPV18v1E67	-T-A--	590-907
HPV18v2E67	-T-A--	590-907
HPV18v5E67	-T-A--	590-907
HELA	-T-A--	629-891
4039	T---G-	610-902
4040	T---G-	610-902
4052	T---G-	610-902
5076	T---G-	610-902
A23	T---G-	621-900
A29	T---G-	621-900
A52	T---G-	626-900
B16	T---G-	631-900
B21	T---G-	631-900
B44	T---G-	629-895
B54	T---G-	618-895
A8	T---G-	626-890
A13	T---G-	626-890
B37	T---G-	629-890
3068	T---G-	620-875
A5	T---G-	626-870
2031	-C-	699-812

HPV-18 Variants

E7 aa	156799
	741322
HPV18	PAMENN
A4	------
A20	------
HPV18v3E671	------
4055	------
B17	--L---
4050	--L---
B6	*-----
SW756	*---KK
101	-*----
161	-*----
173	-*----
C41	-*----
HPV18v1E67	-*-K--
HPV18v2E67	-*-K--
HPV18v5E67	-*-K--
HELA	-*-K--
4039	*---SS
4040	*---SS
4052	*---SS
5076	*---SS
A23	*---SS
A29	*---SS
A52	*---SS
B16	*---SS
B21	*---SS
B44	*---SS
B54	*---SS
A8	*---SS
A13	*---SS
B37	*---SS
3068	*---SS
A5	*---SS
2031	-L-

E1 nuc

```
                1111222
                0188245
                1904690
                2473980
HPV18           ACTTCGT   914-2887
HPV18v2E67      TACGTTA   914-2500
HPV18v3E67      --        914-1737
```

E1 aa

```
                00234^55
                3991522
                3480299
HPV18           STATFDD
HPV18v2E67      *K***$$
HPV18v3E67      --
```

E2 nuc

```
                3333333333333
                4455555555666
                8912356789133
                2235483863702
HPV18           GTTGTCTCATCCT   2817-3914
2               -----------G-   3479-3632
3               -A---------G-   3479-3632
12              -----------G-   3479-3632
14              A----------G-   3479-3632
16              ---A---------   3479-3632
24              A----------G-   3479-3632
40              -------------   3479-3632
42              -------------   3479-3632
47              -----------G-   3479-3632
4               ----CA-TCG-GC   3479-3632
8               --C-CA-TCG-GC   3479-3632
23              ----CACTCG-GC   3479-3632
25              ----CA-TCG-GC   3479-3632
27              ----CA-TCGTGC   3479-3632
33              ----CA-TCG-GC   3479-3632
```

E2 aa

```
                222222222222^2
                2233444555677
                2637089479722
HPV18           QSSAYRPLKHLPP
2               -----------AA
3               -T---------AA
12              -----------AA
14              *----------AA
16              ---T---------
24              *----------AA
40              -------------
42              -------------
47              -----------AA
4               ----H*-*TQ-AA
8               --P-H*-*TQ-AA
23              ----H***TQ-AA
25              ----H*-*TQ-AA
27              ----H*-*TQ*AA
33              ----H*-*TQ-AA
```

E4 nuc

```
                3333333333333
                4455555555666
                8912356789133
                2235483863702
HPV18           GTTGTCTCATCCT   3418-3684
2               ----------G-   3479-3632
3               -A---------G-   3479-3632
12              ----------G-   3479-3632
14              A----------G-   3479-3632
16              ---A---------   3479-3632
24              A----------G-   3479-3632
40              ------------   3479-3632
42              ------------   3479-3632
47              ----------G-   3479-3632
4               ----CA-TCG-GC   3479-3632
8               --C-CA-TCG-GC   3479-3632
23              ----CACTCG-GC   3479-3632
25              ----CA-TCG-GC   3479-3632
27              ----CA-TCGTGC   3479-3632
33              ----CA-TCG-GC   3479-3632
```

E4 aa

```
                2233344555677
                2526979479712
HPV18           SPCPPHLSSISHL
2               ----------Q
3               -*---------Q
12              ----------Q
14              N----------Q
16              ---*--------
24              N----------Q
40              ----------
42              ----------
47              ----------Q
4               ----*Q-LRS-Q
8               --*-*Q-LRS-Q
23              ----*QPLRS-Q
25              ----*Q-LRS-Q
27              ----*Q-LRSLQ
33              ----*Q-LRS-Q
```

L1 nuc

```
                666666666666666
                555666677889999
                678227914471489
                691567799277368
HPV18           TGTCCAAGGCTGAAA   5430-7136
IS002           ---G-----G-----   6558-7012
IS326           ---G-GG--G-----   6558-7012
IS664           A--G-----GC-G-G   6558-7012
IS168           -ACGT--AAG-A---   6558-7012
IS172           -ACGT--AAG-A-G-   6558-7012
IS768           -ACGT--AAG-AC--   6558-7012
```

L1 aa

```
                33^^334444444555
                788991234789012
                944996300136593
HPV18           GVVPPQKELPVPKLQ
IS002           ---RR----*-----
IS326           ---RR*R--*-----
IS664           *--RR----*A-R-*
IS168           -IIRR--***-*---
IS172           -IIRR--***-*-*-
IS768           -IIRR--***-*T--
```

LCR

```
                77777777777777777777777
                44555555555555666667777
                89012223345666945578012 38
                66727890691347231801476 03
HPV18           CCAGAACTCCCGAATTTAATTACCA   7137-7857
AM18-1          -----C--------C----------   7485-7805
G18-1           T-----A------CC---T------   7485-7805
J18-1           --------------C----------   7485-7805
T18-9           ------A------CC---T------   7485-7805
SC18-3          ----C-A------CC---T------   7485-7805
G18-2           ------A------CC---TC-----   7485-7805
SC18-2          T-----A----A-CC---T--G---   7485-7805
SC18-4B         T-----A-T--A-CC---T------   7485-7805
NY18-2          ---A--A----A-CC---T------   7485-7805
A18-1           T-----A-T----CC---T------   7485-7805
NY18-11         T-----A------CC---T-C----   7485-7805
SC18-4A         --G---A-----CCC---T------   7485-7805
NY18-13         T----GA-----CCC---T-C----   7485-7805
A18-2           T-----A----A-CC---T------   7485-7805
IN18-1          T-----A------CC---T---TA-   7485-7805
C4-1            ------A------CC---T--C---   7485-7805
NY18-1          --------------C------C---   7485-7805
NY18-6          T-----A------CC---T--C---   7485-7805
C18-5           -----GA----A-CC---T--C---   7485-7805
C18-7           ------A----A-CC---T--C---   7485-7805
T18-7           ------AA---A-CC---T--CTA-   7485-7805
T18-12          ------AC---A-CCGCCT-CGTA-   7485-7805
T18-10          ------AC---A-CCGCCT-C-TA-   7485-7805
T18-3           -G-A--AC--TA-CC-C-T-C-TA-   7485-7805
T18-5           -G-A--AC---A-CCGCTT-C-TA-   7485-7805
T18-8           -G-A--AC---A-CCGCCT-C-TA-   7485-7805
T18-4           -G-A--AC---A-CCGCCT-C-TAC   7485-7805
T18-18A         -G-A--AC-A-A-CCGCCT-C-TA-   7485-7805
T18-16          -G-A--AC-T-A-CCGCCT-C-TA-   7485-7805
T18-17          -G-A--AC---A-CCGCCT-CGTA-   7485-7805
T18-18B         -G-A--AC-A-A-CCGCCT-CGTA-   7485-7805
```

HPV-33 Variants

```
LOCUS       HPV33        7909 bp ds-DNA   circular  VRL       16-FEB-1987
DEFINITION  Human papillomavirus type 33 (HPV-33), complete genome.
ACCESSION   M12732
SOURCE      Human papillomavirus type 33 DNA recovered from a human invasive
            cervical carcinoma, clone p15-5.
  ORGANISM  Human papillomavirus type 33
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
COMMENT     Naturally occurring variants of HPV33 have been studied in the L1
            region.
            Reference: Stewart,A.-C.M., Eriksson,A.M., Manos,M.M., Munoz,N.,
                Bosch,F.X., Peto,J. and Wheeler,C.M.,
                J. Virol. 70, 3127-36 (1996).
                Variant Accession Nos. U45895-U45897
FEATURES             Location/Qualifiers
     CDS             5594..7093
                     /note=" L1 ORF from bp 5516 to 7093"
                     /product="major capsid protein"
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /translation="MSVWRPSEATVYLPPVPVSKVVSTDEYVSRTSIYYYAGSSRLLA
                     VGHPYFSIKNPTNAKKLLVPKVSGLQYRVFRVRLPDPNKFGFPDTSFYNPDTQRLVWA
                     CVGLEIGRGQPLGVGISGHPLLNKFDDTETGNKYPGQPGADNRECLSMDYKQTQLCLL
                     GCKPPTGEHWGKGVACTNAAPANDCPPLELINTIIEDGDMVDTGFGCMDFKTLQANKS
                     DVPIDICGSTCKYPDYLKMTSEPYGDSLFFFLRREQMFVRHFFNRAGTLGEAVPDDLY
                     IKGSGTTASIQSSAFFPTPSGSMVTSESQLFNKPYWLQRAQGHNNGICWGNQVFVTVV
                     DTTRSTNMTLCTQVTSDSTYKNENFKEYIRHVEEYDLQFVFQLCKVTLTAEVMTYIHA
                     MNPDILEDWQFGLTPPPSASLQDTYRFVTSQAITCQKTVPPKEKEDPLGKYTFWEVDL
                     KEKFSADLDQFPLGRKFLLQAGLKAKPKLKRAAPTSTRTSSAKRKKVKK"
BASE COUNT     2544 a   1354 c   1537 g   2474 t
```

❄ ❄ ❄ ❄ Bases 1 to 5580 not shown. ❄ ❄ ❄ ❄

```
     5581 tatttttta cagATGtccg tgtggcggcc TAGtgaggcc acagtgtacc tgcctcctgt
          _Y__F__F__ T__D__V__R __V__A__A_ _$_                                467
                        _M__S__ V__W__R__P __S__E__A_ _T__V__Y__ L__P__P__V      16
                  L1 cds ->                 <- L2 end

     5641 acctgtatct aaagttgtca gcactgatga atatgtgtct cgcacaagca tttattatta
          __P__V__S_ _K__V__V__ S__T__D__E __Y__V__S_ _R__T__S__ I__Y__Y__Y      36

     5701 tgctggtagt tccagacttc ttgctgttgg ccatccatat ttttctatta aaaatcctac
          __A__G__S_ _S__R__L__ L__A__V__G __H__P__Y_ _F__S__I__ K__N__P__T      56

     5761 taacgctaaa aaattattgg tacccaaagt atcaggcttg caatataggg tttttagggt
          __N__A__K_ _K__L__L__ V__P__K__V __S__G__L_ _Q__Y__R__ V__F__R__V      76

     5821 ccgtttacca gatcctaata aatttggatt tcctgacacc tccttttata accctgatac
          __R__L__P_ _D__P__N__ K__F__G__F __P__D__T_ _S__F__Y__ N__P__D__T      96

     5881 acaacgatta gtatgggcat gtgtaggcct tgaaataggt agagggcagc cattaggcgt
          __Q__R__L_ _V__W__A__ C__V__G__L __E__I__G_ _R__G__Q__ P__L__G__V      116

     5941 tggcataagt ggtcatcctt tattaaacaa atttgatgac actgaaaccg gtaacaagta
          __G__I__S_ _G__H__P__ L__L__N__K __F__D__D_ _T__E__T__ G__N__K__Y      136

     6001 tcctggacaa ccgggtgctg ataataggga atgtttatcc atggattata aacaaacaca
          __P__G__Q_ _P__G__A__ D__N__R__E __C__L__S_ _M__D__Y__ K__Q__T__Q      156

     6061 gttatgttta cttggatgta agcctccaac aggggaacat tggggtaaag gtgttgcttg
          __L__C__L_ _L__G__C__ K__P__P__T __G__E__H_ _W__G__K__ G__V__A__C      176

     6121 tactaatgca gcacctgcca atgattgtcc acctttagaa cttataaata ctattattga
          __T__N__A_ _A__P__A__ N__D__C__P __P__L__E_ _L__I__N__ T__I__I__E      196
```

```
6181 ggatggtgat atggtggaca caggatttgg ttgcatggat tttaaaacat tgcaggctaa
     __D__G__D_ _M__V__D__ T__G__F__G __C__M__D_ _F__K__T__ L__Q__A__N      216

6241 taaaagtgat gttcctattg atatttgtgg cagtacatgc aaatatccag attatttaaa
     __K__S__D_ _V__P__I__ D__I__C__G __S__T__C_ _K__Y__P__ D__Y__L__K      236

6301 aatgactagt gagccttatg gtgatagttt atttttcttt cttcgacgtg aacaaatgtt
     __M__T__S_ _E__P__Y__ G__D__S__L __F__F__F_ _L__R__R__ E__Q__M__F      256

6361 tgtaagacac ttttttaata gggctggtac attaggagag gctgttcccg atgacctgta
     __V__R__H_ _F__F__N__ R__A__G__T __L__G__E_ _A__V__P__ D__D__L__Y      276

6421 cattaaaggt tcaggaacta ctgcctctat tcaaagcagt gctttttttc ccactcctag
     __I__K__G_ _S__G__T__ T__A__S__I __Q__S__S_ _A__F__F__ P__T__P__S      296

6481 tggatcaatg gttacttccg aatctcagtt atttaataag ccatattggc tacaacgtgc
     __G__S__M_ _V__T__S__ E__S__Q__L __F__N__K_ _P__Y__W__ L__Q__R__A      316

6541 acaaggtcat aataatggta tttgttgggg caatcaggta tttgttactg tggtagatac
     __Q__G__H_ _N__N__G__ I__C__W__G __N__Q__V_ _F__V__T__ V__V__D__T      336

6601 cactcgcagt actaatatga ctttatgcac acaagtaact agtgacagta caTATAAAAa
     __T__R__S_ _T__N__M__ T__L__C__T __Q__V__T_ _S__D__S__ T__Y__K__N      356
                                                      signal ->

6661 tgaaaatttt aaagaatata taagacatgt tgaagaatat gatctacagt ttgtttttca
     __E__N__F_ _K__E__Y__ I__R__H__V __E__E__Y_ _D__L__Q__ F__V__F__Q      376

6721 actatgcaaa gttaccttaa ctgcagaagt tatgacatat attcatgcta tgaatccaga
     __L__C__K_ _V__T__L__ T__A__E__V __M__T__Y_ _I__H__A__ M__N__P__D      396

6781 tattttagaa gattggcaat ttggtttaac acctcctcca tctgctagtt tacaggatac
     __I__L__E_ _D__W__Q__ F__G__L__T __P__P__P_ _S__A__S__ L__Q__D__T      416

6841 ctataggttt gttacctctc aggctattac gtgtcaaaaa acagtacctc caaaggaaaa
     __Y__R__F_ _V__T__S__ Q__A__I__T __C__Q__K_ _T__V__P__ P__K__E__K      436

6901 ggaagacccc ttaggtaaat atacattttg ggaagtggat ttaaaggaaa aattttcagc
     __E__D__P_ _L__G__K__ Y__T__F__W __E__V__D_ _L__K__E__ K__F__S__A      456

6961 agatttagat cagtttcctt tgggacgcaa gtttttatta caggcaggtc ttaaagcaaa
     __D__L__D_ _Q__F__P__ L__G__R__K __F__L__L_ _Q__A__G__ L__K__A__K      476

7021 acctaaactt aaacgtgcag cccccacatc cacccgcaca tcgtctgcaa aacgcaaaaa
     __P__K__L_ _K__R__A__ A__P__T__S __T__R__T_ _S__S__A__ K__R__K__K      496

7081 ggttaaaaaa TAAcactTTG TGTAattgtg ttatgttgtt gttttgttct gtctatgtac
     __V__K__K_ _$_                                                       499
          LCR ->  -> repeat region start
            <- L1 end
```

❄ ❄ ❄ ❄ Bases 7141 to 7909 not shown. ❄ ❄ ❄ ❄

L1 nuc

	6666	
	6679	
	3603	
	7448	
HPV33	AACG	5594-7093
IS267	---A	6539-6987
IS549	-G--	6539-6987
IS827	C-T-	6539-6987

L1 aa

	3334
	4574
	8719
HPV33	VELD
IS267	---N
IS549	-*--
IS827	*-*-

```
LOCUS       HPV39        7833 bp ds-DNA             VRL       06-MAR-1991
DEFINITION  Human papillomavirus type 39 (HPV-39), complete genome.
ACCESSION   M62849 M38185
SOURCE      Human papillomavirus type 39 DNA isolated from a penile Bowenoid
            papule biopsy.
  ORGANISM  Human papillomavirus type 39
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
COMMENT     Naturally occurring variants of HPV39 have been studied in the L1
            region.
            Reference: Stewart,A.-C.M., Eriksson,A.M., Manos,M.M., Munoz,N.,
                Bosch,F.X., Peto,J. and Wheeler,C.M.,
                J. Virol. 70, 3127-36 (1996).
                Variant Accession Nos.  U45899-U45905
FEATURES             Location/Qualifiers
     CDS             5643..7160
                     /note="L1 ORF from bp 5610 to 7160"
                     /product="major capsid protein"
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /translation="MAMWRSSDSMVYLPPPSVAKVVNTDDYVTRTGIYYYAGSSRLLT
                     VGHPYFKVGMNGGRKQDIPKVSAYQYRVFRVTLPDPNKFSIPDASLYNPETQRLVWAC
                     VGVEVGRGQPLGVGISGHPLYNRQDDTENSPFSSTTNKDSRDNVSVDYKQTQLCIIGC
                     VPAIGEHWGKGKACKPNNVSTGDCPPLELVNTPIEDGDMIDTGYGAMDFGALQETKSE
                     VPLDICQSICKYPDYLQMSADVYGDSMFFCLRREQLFARHFWNRGGMVGDAIPAQLYI
                     KGTDIRANPGSSVYCPSPSGSMVTSDSQLFNKPYWLHKAQGHNNGICWHNQLFLTVVD
                     TTRSTNFTLSTSIESSIPSTYDPSKFKEYLRHVEEYDLQFIFQLCTVTLTTDVMSYIH
                     TMNSSILDNWNFAVAPPPSASLVDTYRYLQSAAITCQKDAPAPEKKDPYDGLKFWNVD
                     LREKFSLELDQFPLGRKFLLQARVRRRPTIGPRKRPAASTSSSSATKHKRKRVSK"
BASE COUNT     2426 a   1485 c   1660 g   2262 t
```

❄ ❄ ❄ ❄ Bases 1 to 5640 not shown. ❄ ❄ ❄ ❄

```
5641 agATGgctat gtggcggtcT AGtgacagca tggtgtattt gcctccacct tctgtggcga
       _M__A__M __W__R__S_ _S__D__S__ M__V__Y__L __P__P__P_ _S__V__A__      19
L1 cds ->                  <- L2 end

5701 aggttgtcaa tactgatgat tatgttacac gcacaggcat atattattat gctggcagct
     K__V__V__N __T__D__D_ _Y__V__T__ R__T__G__I __Y__Y__Y_ _A__G__S__      39

5761 ctagattatt aacagtagga catccatatt ttaaagtggg tatgaatggt ggtcgcaagc
     S__R__L__L __T__V__G_ _H__P__Y__ F__K__V__G __M__N__G_ _G__R__K__      59

5821 aggacattcc aaaggtgtct gcatatcaat atagggtatt tcgcgtgaca ttgcccgatc
     Q__D__I__P __K__V__S_ _A__Y__Q__ Y__R__V__F __R__V__T_ _L__P__D__      79

5881 ctaataaatt cagtattcca gatgcatcct tatataatcc agaaacacaa cgtttagtat
     P__N__K__F __S__I__P_ _D__A__S__ L__Y__N__P __E__T__Q_ _R__L__V__      99

5941 gggcttgtgt aggggtggag gtgggcaggg gccagccatt gggtgttggt attagtggac
     W__A__C__V __G__V__E_ _V__G__R__ G__Q__P__L __G__V__G_ _I__S__G__     119

6001 acccattata taatagacag gatgatactg aaaactcacc attttcatca accaccaata
     H__P__L__Y __N__R__Q_ _D__D__T__ E__N__S__P __F__S__S_ _T__T__N__     139

6061 aggacagtag ggataatgtg tctgtggatt ataaacagac acagttgtgc attataggct
     K__D__S__R __D__N__V_ _S__V__D__ Y__K__Q__T __Q__L__C_ _I__I__G__     159

6121 gtgttcccgc cattggggag cactggggta agggaaaggc atgcaagccc aataatgtat
     C__V__P__A __I__G__E_ _H__W__G__ K__G__K__A __C__K__P_ _N__N__V__     179

6181 ctacgggggа ctgtcctcct ttggaactag taaacacccc tattgaggat ggtgatatga
     S__T__G__D __C__P__P_ _L__E__L__ V__N__T__P __I__E__D_ _G__D__M__     199
```

```
6241 ttgatactgg ctatggagct atggactttg gtgcattgca ggaaaccaaa agtgaggtgc
     I__D__T__G __Y__G__A_ _M__D__F__ G__A__L__Q __E__T__K_ _S__E__V__      219

6301 ctttagatat ttgtcaatcc atttgtaaat atcctgatta tttgcaaatg tctgcagatg
     P__L__D__I __C__Q__S_ _I__C__K__ Y__P__D__Y __L__Q__M_ _S__A__D__      239

6361 tgtatgGGGA CAGTATGTTC Ttctgtttac gtagggaaca actgtttgca agacattttt
     V__Y__G__D __S__M__F_ _F__C__L__ R__R__E__Q __L__F__A_ _R__H__F__      259
           -> glucocorticoid response element

6421 ggaatcgtgg tggtatggtg ggtgacgcca ttcctgccca attgtatatt aagggcacag
     W__N__R__G __G__M__V_ _G__D__A__ I__P__A__Q __L__Y__I_ _K__G__T__      279

6481 atatacgtgc aaaccccggt agttctgtat actgcccctc tcccagcggt tccatggtaa
     D__I__R__A __N__P__G_ _S__S__V__ Y__C__P__S __P__S__G_ _S__M__V__      299

6541 cctctgattc ccagttattt aataagcctt attggctaca taaggcccag ggcacaacca
     T__S__D__S __Q__L__F_ _N__K__P__ Y__W__L__H __K__A__Q_ _G__H__N__      319

6601 atggtatatg ttggcataat caattatttc ttactgttgt ggacactacc cgtagtacca
     N__G__I__C __W__H__N_ _Q__L__F__ L__T__V__V __D__T__T_ _R__S__T__      339

6661 actttacatt atctacctct atagagtctt ccataccttc tacatatgat ccttctaagt
     N__F__T__L __S__T__S_ _I__E__S__ S__I__P__S __T__Y__D_ _P__S__K__      359

6721 ttaaggaata taccaggcac gtggaggagt atgatttaca atttatattt caactgtgta
     F__K__E__Y __T__R__H_ _V__E__E__ Y__D__L__Q __F__I__F_ _Q__L__C__      379

6781 ctgtgacatt aacaactgat gttatgtctt atattcacac tatgaattcc tctatattgg
     T__V__T__L __T__T__D_ _V__M__S__ Y__I__H__T __M__N__S_ _S__I__L__      399

6841 acaattggaa ttttgctgta gctcctccac catctgccag tttggtagac acttacagat
     D__N__W__N __F__A__V_ _A__P__P__ P__S__A__S __L__V__D_ _T__Y__R__      419

6901 acctacagtc tgcagccatt acatgtcaaa aggatgctcc agcacctgaa aagaaagatc
     Y__L__Q__S __A__A__I_ _T__C__Q__ K__D__A__P __A__P__E_ _K__K__D__      439

6961 catatgacgg tctaaagttt tggaatgttg acttaaggga aaagtttagt ttggaacttg
     P__Y__D__G __L__K__F_ _W__N__V__ D__L__R__E __K__F__S_ _L__E__L__      459

7021 atcaatttcc ttttgggacgt aaatttttgt tgcaggccag ggtccgcagg cgccctacta
     D__Q__F__P __L__G__R_ _K__F__L__ L__Q__A__R __V__R__R_ _R__P__T__      479
```

❄ ❄ ❄ ❄ Bases 7081 to 7833 not shown. ❄ ❄ ❄ ❄

L1 nuc

```
                66666666667777
                55567788990000
                99933855092223
                36983534365781
HPV39           CCCTCCTTCAATCCT   5643-7160
IS270           TTTA------GATC    6585-7039
IS114           TTTA----T-GATC    6585-7039
IS073           TTT-----T-GATC    6585-7039
IS065           TTT--A-CT-GATC    6585-7039
IS015           TTT--AC-TCGATC    6585-7039
IS214           TTTATA--T-GATC    6585-7039
IS351           TTTATAC-T-GATC    6585-7039
```

L1 aa

```
                333333^444444^44
                11136800256666
                78924144121223
HPV39           GHNVTVFFLRQFFP
IS270           ****------*YY*
IS114           ****----*-*YY*
IS073           ***-----*-*YY*
IS065           ***--*-**-*YY*
IS015           ***--*SS***YY*
IS214           ****I*--*-*YY*
IS351           ****I*SS*-*YY*
```

```
LOCUS       HPV45        7858 bp ds-DNA             VRL       04-OCT-1993
DEFINITION  Human papillomavirus type 45 (HPV-45), complete genome.
ACCESSION   X74479
SOURCE      Human papillomavirus type 45 DNA.
  ORGANISM  Human papillomavirus type 45
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
  STANDARD  full staff_review
            Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im
            Neuenheimer Feld 506, W 6900 Heidelberg, FRG
COMMENT     Naturally occurring variants of HPV45 have been studied in the L1
            region.
            Reference: Stewart,A.-C.M., Eriksson,A.M., Manos,M.M., Munoz,N.,
                Bosch,F.X., Peto,J. and Wheeler,C.M.,
                J. Virol. 70, 3127-36 (1996).
                Variant Accession Nos. U45906-U45916
FEATURES             Location/Qualifiers
     CDS             5530..7149
                     /note="putative"
                     /note="ORF L1 from bp 5524 to 7149"
                     /product="major capsid protein"
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /translation="MAHNIIYGHGIIIFLKNVNVFPIFLQMALWRPSDSTVYLPPPSV
                     ARVVSTDDYVSRTSIFYHAGSSRLLTVGNPYFRVVPNGAGNKQAVPKVSAYQYRVFRV
                     ALPDPNKFGLPDSTIYNPETQRLVWACVGMEIGRGQPLGIGLSGHPFYNKLDDTESAH
                     AATAVITQDVRDNVSVDYKQTQLCILGCVPAIGEHWAKGTLCKPAQLQPGDCPPLELK
                     NTIIEDGDMVDTGYGAMDFSTLQDTKCEVPLDICQSICKYPDYLQMSADPYGDSMFFC
                     LRREQLFARHFWNRAGVMGDTVPTDLYIKGTSANMRETPGSCVYSPSPSGSIITSDSQ
                     LFNKPYWLHKAQGHNNGICWHNQLFVTVVDTTRSTNLTLCASTQNPVPSTYDPTKFKQ
                     YSRHVEEYDLQFIFQLCTITLTAEVMSYIHSMNSSILENWNFGVPPPPTTSLVDTYRF
                     VQSVAVTCQKDTTPPEKQDPYDKLKFWTVDLKEKFSSDLDQYPLGRKFLVQAGLRRRP
                     TIGPRKRPAASTSTASTASRPAKRVRIRSKK"
BASE COUNT     2409 a   1462 c   1652 g   2335 t
```

❄ ❄ ❄ ❄ Bases 1 to 5520 not shown. ❄ ❄ ❄ ❄

```
     5521 TAGGTattcA TGgcacacaa tattatttat ggccatggta ttattatttt cctaaaaaac
                  _ M__A__H__N __I__I__Y_ _G__H__G__ I__I__I__F __L__K__N_      17
      L1 orf -> L1 ->
        start   cds
     E2 bind <-

     5581 gtaaacgtat tccctatttt tttgcagatg gctttgtggc ggccTAGtga cagtacggta
          _V__N__V__ F__P__I__F __L__Q__M_ _A__L__W__ R__P__S__D __S__T__V_      37
                                                        <- L2 end

     5641 tatcttccac caccttctgt ggccagagtt gtcagcactg atgattatgt gtctcgcaca
          _Y__L__P__ P__P__S__V __A__R__V_ _V__S__T__ D__D__Y__V __S__R__T_      57

     5701 agcatatttt atcatgcagg cagttcccga ttattaactg taggcaatcc atattttagg
          _S__I__F__ Y__H__A__G __S__S__R_ _L__L__T__ V__G__N__P __Y__F__R_      77

     5761 gttgtaccta atggtgcagg taataaacag gctgttccta aggtatccgc atatcagtat
          _V__V__P__ N__G__A__G __N__K__Q_ _A__V__P__ K__V__S__A __Y__Q__Y_      97

     5821 agggtgttta gagtagcttt acccgatcct aataaatttg gattacctga ttctactata
          _R__V__F__ R__V__A__L __P__D__P_ _N__K__F__ G__L__P__D __S__T__I_      117

     5881 tataatcctg aaacacaacg tttggtttgg gcatgtgtag gtatggaaat tggtcgtggg
          _Y__N__P__ E__T__Q__R __L__V__W_ _A__C__V__ G__M__E__I __G__R__G_      137
```

```
5941 cagcctttag gtattggcct aagtggccat ccattttata ataaattgga tgatacagaa
     _Q__P__L__ G__I__G__L __S__G__H_ _P__F__Y__ N__K__L__D __D__T__E_      157

6001 agtgctcatg cagctacagc tgttattacg caggatgtta gggataatgt gtcagttgat
     _S__A__H__ A__A__T__A __V__I__T_ _Q__D__V__ R__D__N__V __S__V__D_      177

6061 tataagcaaa cacagctgtg tattttaggt tgtgtacctg ctattggtga gcactgggcc
     _Y__K__Q__ T__Q__L__C __I__L__G_ _C__V__P__ A__I__G__E __H__W__A_      197

6121 aagggcacac tttgtaaacc tgcacaattg caacctggtg actgtcctcc tttggaactt
     _K__G__T__ L__C__K__P __A__Q__L_ _Q__P__G__ D__C__P__P __L__E__L_      217

6181 aaaaacacca ttattgagga tggtgatatg gtggatacag gttatggggc aatggatttt
     _K__N__T__ I__I__E__D __G__D__M_ _V__D__T__ G__Y__G__A __M__D__F_      237

6241 agtacattgc aggatacaaa gtgcgaggtt ccattagaca tttgtcaatc catctgtaaa
     _S__T__L__ Q__D__T__K __C__E__V_ _P__L__D__ I__C__Q__S __I__C__K_      257

6301 tatccagatt atttgcaaat gtctgctgat ccctatgggg attctatgtt tttttgccta
     _Y__P__D__ Y__L__Q__M __S__A__D_ _P__Y__G__ D__S__M__F __F__C__L_      277

6361 cgccgtgaac aactgtttgc aagacatttt tggaataggg caggtgttat gggtgacaca
     _R__R__E__ Q__L__F__A __R__H__F_ _W__N__R__ A__G__V__M __G__D__T_      297

6421 gtacctacgg acctatatat taaaggcact agcgctaata tgcgtgaaac ccctggcagt
     _V__P__T__ D__L__Y__I __K__G__T_ _S__A__N__ M__R__E__T __P__G__S_      317

6481 tgtgtgtatt ccccttctcc cagtggctct attattactt ctgattctca attatttaat
     _C__V__Y__ S__P__S__P __S__G__S_ _I__I__T__ S__D__S__Q __L__F__N_      337

6541 aagccatatt ggttacataa ggcccagggc cataacaatg gtatttgttg gcataatcag
     _K__P__Y__ W__L__H__K __A__Q__G_ _H__N__N__ G__I__C__W __H__N__Q_      357

6601 ttgtttgtta ctgtagtgga cactacccgc agtactaatt taacattatg tgcctctaca
     _L__F__V__ T__V__V__D __T__T__R_ _S__T__N__ L__T__L__C __A__S__T_      377

6661 caaaatcctg tgccatgtac atatgacgcct actaagttta agcagtatag tagacatgtg
     _Q__N__P__ V__P__S__T __Y__D__P_ _T__K__F__ K__Q__Y__S __R__H__V_      397

6721 gaggaatatg atttacagtt tatttttcag ttgtgcacta ttactttaac tgcagaggtt
     _E__E__Y__ D__L__Q__F __I__F__Q_ _L__C__T__ I__T__L__T __A__E__V_      417

6781 atgtcatata tccatagtat gaatagtagt atattagaaa attggaattt tggtgtccct
     _M__S__Y__ I__H__S__M __N__S__S_ _I__L__E__ N__W__N__F __G__V__P_      437

6841 ccaccaccta ctacaagttt ggtggataca tatcgttttg tgcaatcagt tgctgttacc
     _P__P__P__ T__T__S__L __V__D__T_ _Y__R__F__ V__Q__S__V __A__V__T_      457

6901 tgtcaaaagg atactacacc tccagaaaag caggatccat atgataaatt aaagttttgg
     _C__Q__K__ D__T__T__P __P__E__K_ _Q__D__P__ Y__D__K__L __K__F__W_      477

6961 actgttgacc taaaggaaaa attttcctcc gatttggatc aatatcccct tggtcgaaag
     _T__V__D__ L__K__E__K __F__S__S_ _D__L__D__ Q__Y__P__L __G__R__K_      497

7021 tttttagttc aggctgggtt acgtcgtagg cctaccatag gacctcgtaa gcgtcctgct
     _F__L__V__ Q__A__G__L __R__R__R_ _P__T__I__ G__P__R__K __R__P__A_      517

7081 gcttccacgt ctactgcatc tactgcatct aggcctgcca aacgtgtacg tatacgtagt
     _A__S__T__ S__T__A__S __T__A__S_ _R__P__A__ K__R__V__R __I__R__S_      537

7141 aagaaaTAAt atgttagcac atatatgtat gtttgtatgt atggttttgt atgttgtatg
     _K__K__$_                                                             539
           <- L1 end
```

❄ ❄ ❄ ❄ Bases 7201 to 7858 not shown. ❄ ❄ ❄ ❄

L1 nuc

```
                6666666666666666
                6666667888889999
                2667780135661159
                1156775672120416
HPV45           CCAAGCGACTGGGCAG   5530-7149
IS816           ----------A-----   6562-7016
IS278           ---G---G--A---G-   6562-7016
IS871           --C---C---A--G--   6562-7016
IS894           ---G--CG--A---G-   6562-7016
IS040           ----ATCGA-A-----   6562-7016
IS788           --CG--CG--A--G--   6562-7016
IS819           ---G-TCG--A----A   6562-7016
IS252           ----ATCGA-A-A---   6562-7016
IS282           ----ATCGA-AA----   6562-7016
IS949           T--G-TCGA-A-----   6562-7016
IS761           -GCG--AGGCA--G--   6562-7016
```

L1 aa

```
                3333^333444444444
                6778889234446678
                4893362961451249
HPV45           DQNSSDQLVTLVDTLL
IS816           ----------*-----
IS278           ---GG--*--*---*-
IS871           --T---H---*--S--
IS894           ---GG-H*--*---*-
IS040           ---NN*H**-*-----
IS788           --TGG-H*--*--S--
IS819           ---GG*H*--*----*
IS252           ---NN*H**-*-N---
IS282           ---NN*H**-*M----
IS949           *--GG*H**-*-----
```

```
LOCUS       HPV52        7942 bp ds-DNA             VRL       04-OCT-1993
DEFINITION  Human papillomavirus type 52 (HPV-52), complete genome.
ACCESSION   X74481
SOURCE      Human papillomavirus type 52 DNA.
  ORGANISM  Human papillomavirus type 52
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
  STANDARD  full staff_review
            Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im
            Neuenheimer Feld 506, W 6900 Heidelberg, FRG
COMMENT     Naturally occurring variants of HPV52 have been studied in the L1
            region.
            Reference: Stewart,A.-C.M., Eriksson,A.M., Manos,M.M., Munoz,N.,
                Bosch,F.X., Peto,J. and Wheeler,C.M.,
                J. Virol. 70, 3127-36 (1996).
                Variant Accession Nos.  U45918-U45923
FEATURES             Location/Qualifiers
     CDS             5565..7154
                     /note="putative"
                     /note="ORF L1 from bp 5532 to 7154"
                     /product="major capsid protein"
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /translation="MVQILFYILVIFYYVAGVNVFHIFLQMSVWRPSEATVYLPPVPV
                     SKVVSTDEYVSRTSIYYYAGSSRLLTVGHPYFSIKNTSSGNGKKVLVPKVSGLQYRVF
                     RIKLPDPNKFGFPDTSFYNPETQRLVWACTGLEIGRGQPLGVGISGHPLLNKFDDTET
                     SNKYAGKPGIDNRECLSMDYKQTQLCILGCKPPIGEHWGKGTPCNNNSGNPGDCPPLQ
                     LINSVIQDGDMVDTGFGCMDFNTLQASKSDVPIDICSSVCKYPDYLQMASEPYGDSLF
                     FFLRREQMFVRHFFNRAGTLGDPVPGDLYIQGSNSGNTATVQSSAFFPTPSGSMVTSE
                     SQLFNKPYWLQRAQGHNNGICWGNQLFVTVVDTTRSTNMTLCAEVKKESTYKNENFKE
                     YLRHGEEFDLQFIFQLCKITLTADVMTYIHKMDATILEDWQFGLTPPPSASLEDTYRF
                     VTSTAITCQKNTPPKGKEDPLKDYMFWEVDLKEKFSADLDQFPLGRKFLLQAGLQARP
                     KLKRPASSAPRTSTKKKKVKR"
BASE COUNT     2540 a   1449 c   1621 g   2332 t
```

❄ ❄ ❄ ❄ Bases 1 to 5520 not shown. ❄ ❄ ❄ ❄

```
     5521 tgttcctaTA Gcccctacag ctccatctac atctattatt gttgATGgta cagattttat
          __V__P__I_ _A__P__T__ A__P__S__T __S__I__I_ _V__D__G__ T__D__F__I      440
                                                     _M__V_ _Q__I__L__      5
         L1 orf start ->                          L1 cds ->

     5581 tttacatcct agttatttttt tactacgtcg caggcgtaaa cgttttccat atttttttac
          __L__H__P_ _S__Y__F__ L__L__R__R __R__R__K_ _R__F__P__ Y__F__F__T      460
          F__Y__I__L __V__I__F_ _Y__Y__V__ A__G__V__N __V__F__H_ _I__F__L__      25

     5641 agatgtccgt gtggcggccT AGtgaggcca ctgtgtacct gcctcctgta cctgtctcta
          __D__V__R_ _V__A__A__ $_                                              466
          Q__M__S__V __W__R__P_ _S__E__A__ T__V__Y__L __P__P__V_ _P__V__S__      45
                                <- L2 end

     5701 aggttgtaag cactgatgag tatgtgtctc gcacaagcat ctattattat gcaggcagtt
          K__V__V__S __T__D__E_ _Y__V__S__ R__T__S__I __Y__Y__Y_ _A__G__S__      65

     5761 ctcgattact aacagtagga catccctatt tttctattaa aaacACCAGT AGTGGTaatg
          S__R__L__L __T__V__G_ _H__P__Y__ F__S__I__K __N__T__S_ _S__G__N__      85
                                                        -> E2 bind <-

     5821 gtaaaaaagt tttagttccc aaggtgtctg gcctgcaata cagggtattt agaattaaat
          G__K__K__V __L__V__P_ _K__V__S__ G__L__Q__Y __R__V__F_ _R__I__K__      105

     5881 tgccggaccc taataaattt ggttttccag atacatcttt ttataaccca gaaacccaaa
          L__P__D__P __N__K__F_ _G__F__P__ D__T__S__F __Y__N__P_ _E__T__Q__      125
```

```
5941 ggttggtgtg ggcctgtaca ggcttggaaa ttggtagggg acagccttta ggtgtgggta
     R__L__V__W __A__C__T_ _G__L__E__ I__G__R__G __Q__P__L_ _G__V__G__      145

6001 ttagtgggca tcctttatta aacaagtttg atgatactga aaccagtaac aaatatgctg
     I__S__G__H __P__L__L_ _N__K__F__ D__D__T__E __T__S__N_ _K__Y__A__      165

6061 gtaaacctgg tatagataat agggaatgtt tatctatgga ttataagcag actcagttat
     G__K__P__G __I__D__N_ _R__E__C__ L__S__M__D __Y__K__Q_ _T__Q__L__      185

6121 gcattttagg atgcaaacct cctataggtg aacattgggg taagggaacc ccttgtaata
     C__I__L__G __C__K__P_ _P__I__G__ E__H__W__G __K__G__T_ _P__C__N__      205

6181 ataattcagg aaatcctggg gattgtcctc ccctacagct cattaacagt gtaatacagg
     N__N__S__G __N__P__G_ _D__C__P__ P__L__Q__L __I__N__S_ _V__I__Q__      225

6241 atggggacat ggtagataca ggatttggtt gcatggattt taataccttg caagctagta
     D__G__D__M __V__D__T_ _G__F__G__ C__M__D__F __N__T__L_ _Q__A__S__      245

6301 aaagtgatgt gcccattgat atatgtagca gtgtatgtaa gtatccagat tatttgcaaa
     K__S__D__V __P__I__D_ _I__C__S__ S__V__C__K __Y__P__D_ _Y__L__Q__      265

6361 tggctagcga gccatatggt gacagtttgt tctttttttct tagacgtgag caaatgtttg
     M__A__S__E __P__Y__G_ _D__S__L__ F__F__F__L __R__R__E_ _Q__M__F__      285

6421 ttagacactt ttttaatagg gccggtacct taggtgaccc tgtgccaggt gatttatata
     V__R__H__F __F__N__R_ _A__G__T__ L__G__D__P __V__P__G_ _D__L__Y__      305

6481 tacaagggtc taactctggc aatactgcca ctgtacaaag cagtgctttt tttcctactc
     I__Q__G__S __N__S__G_ _N__T__A__ T__V__Q__S __S__A__F_ _F__P__T__      325

6541 ctagtggttc tatggtaacc tcagaatccc aattatttaa taaACCGTAC TGGTTacaac
     P__S__G__S __M__V__T_ _S__E__S__ Q__L__F__N __K__P__Y_ _W__L__Q__      345
                                                   -> E2 bind

6601 gtgcgcaggg ccacaataat ggcatatgtt ggggcaatca gttgtttgtc acagttgtgg
     R__A__Q__G __H__N__N_ _G__I__C__ W__G__N__Q __L__F__V_ _T__V__V__      365

6661 ataccactcg tagcactaac atgactttat gtgctgaggt taaaaaaggaa atgcacaTATA
     D__T__T__R __S__T__N_ _M__T__L__ C__A__E__V __K__K__E_ _S__T__Y__      385
                                                         signal ->

6721 AAAatgaaaa ttttaaggaa taccttcgtc atggcgagga atttgattta caatttattt
     K__N__E__N __F__K__E_ _Y__L__R__ H__G__E__E __F__D__L_ _Q__F__I__      405

6781 ttcaattgtg caaaattaca ttaacagctg atgttatgac atacattcat aagatggatg
     F__Q__L__C __K__I__T_ _L__T__A__ D__V__M__T __Y__I__H_ _K__M__D__      425

6841 ccactattt agaggactgg caatttggcc ttaccccacc accgtctgca tctttggagg
     A__T__I__L __E__D__W_ _Q__F__G__ L__T__P__P __P__S__A_ _S__L__E__      445

6901 acacatacag atttgtcact tctactgcta taactgtgca aaaaaacaca ccacctaaag
     D__T__Y__R __F__V__T_ _S__T__A__ I__T__C__Q __K__N__T_ _P__P__K__      465

6961 gaaaggaaga tcctttaaag gactatatgt ttgggaggt ggatttaaaa gaaaagtttt
     G__K__E__D __P__L__K_ _D__Y__M__ F__W__E__V __D__L__K_ _E__K__F__      485

7021 ctgcagattt agatcagttt cctttaggta ggaagttttt gttacaggca gggctacagg
     S__A__D__L __D__Q__F_ _P__L__G__ R__K__F__L __L__Q__A_ _G__L__Q__      505

7081 ctaggcccaa actaaaacgc cctgcatcat cggccccacg tacctccaca aagaagaaaa
     A__R__P__K __L__K__R_ _P__A__S__ S__A__P__R __T__S__T_ _K__K__K__      525

7141 aggttaaaag gTAAccattg tctgttgggt aattgtctgt gtcatgtatg tgttgtgtat
     K__V__K__R __$_                                                      529
                <- L1 end
```

❄ ❄ ❄ ❄ Bases 7201 to 7942 not shown. ❄ ❄ ❄ ❄

L1 nuc

```
                6666666666666666666
                6777777888999999999
                9001169234123445889
                8131244438705149032
HPV52           GTAAGTACGTCTTAAAGCT   5565-7154
IS1001          -----CGT--A--------   6603-7051
IS121           A--------CA----G---   6603-7051
IS370           ------G-A-A---G----   6603-7051
IS705           A---------A--------   6603-7051
IS1015          -G----GT--A--------   6603-7051
IS464           --CGA-----AGAG--AGC   6603-7051
```

L1 aa

```
                333334444444444444444
                7788801222555566777
                8903300038127905236
HPV52           EVKSSFKYKIVTTQKKKDF
IS1001          -----***--*--------
IS121           *--------**----*---
IS370           ------*-*-*---*----
IS705           *---------*--------
IS1015          -*----**--*--------
IS464           --TDD-----****--*E*
```

```
LOCUS       HPV58        7824 bp ds-DNA             VRL       24-JAN-1992
DEFINITION  Human papillomavirus type 58 (HPV-58), complete genome.
ACCESSION   D90400
SOURCE      Human papillomavirus type 58 DNA recovered from an invasive
            cervical carcinoma, patient specimen GN479.
  ORGANISM  Human papillomavirus type 58
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
COMMENT     Naturally occurring variants of HPV58 have been studied in the L1
            region.
            Reference: Stewart,A.-C.M., Eriksson,A.M., Manos,M.M., Munoz,N.,
                Bosch,F.X., Peto,J. and Wheeler,C.M.,
                J. Virol. 70, 3127-36 (1996).
                Variant Accession Nos. U45924-U45929
FEATURES             Location/Qualifiers
     CDS             5565..7139
                     /note=" L1 ORF from bp 5559 to 7139"
                     /product="major capsid protein"
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /translation="MVLILCCTLAILFCVADVNVFHIFLQMSVWRPSEATVYLPPVPV
                     SKVVSTDEYVSRTSIYYYAGSSRLLAVGNPYFSIKSPNNNKKVLVPKVSGLQYRVFRV
                     RLPDPNKFGFPDTSFYNPDTQRLVWACVGLEIGRGQPLGVGVSGHPYLNKFDDTETSN
                     RYPAQPGSDNRECLSMDYKQTQLCLIGCKPPTGEHWGKGVACNNNAAATDCPPLELFN
                     SIIEDGDMVDTGFGCMDFGTLQANKSDVPIDICNSTCKYPDYLKMASEPYGDSLFFFL
                     RREQMFVRHFFNRAGKLGEAVPDDLYIKGSGNTAVIQSSAFFPTPSGSIVTSESQLFN
                     KPYWLQRAQGHNNGICWGNQLFVTVVDTTRSTNMTLCTEVTKEGTYKNDNFKEYVRHV
                     EEYDLQFVFQLCKITLTAELMTYIHTMDSNILEDWQFGLTPPPSASLQDTYRFVTSQA
                     ITCQKTAPPKEKEDPLNKYTFWEVNLKEKFSADLDQFPLGRKFLLQSGLKAKPRLKRS
                     APTTRAPSTKRKKVKK"
BASE COUNT     2487 a   1388 c   1576 g   2373 t
```

❄ ❄ ❄ ❄ Bases 1 to 5520 not shown. ❄ ❄ ❄ ❄

```
     5521 tattcctata tctccactaa ctccttttaa taccaTAAtt gtggATGgtg ctgattttat
          __I__P__I_ _S__P__L__ T__P__F__N __T__I__I_ _V__D__G__ A__D__F__M      446
                                                     _M__V_ _L__I__L__      5
                                       L1 orf start ->     -> L1 cds

     5581 gttgcaccct agctatttta ttttgcgtcg cagacgtaaa cgttttccat atttttttgc
          __L__H__P_ _S__Y__F__ I__L__R__R __R__R__K_ _R__F__P__ Y__F__F__A      466
          C__C__T__L __A__I__L_ _F__C__V__ A__D__V__N __V__F__H_ _I__F__L__      25

     5641 agatgtccgt gtggcggccT AGtgaggcca ctgtgtacct gcctcctgtg cctgtgtcta
          __D__V__R_ _V__A__A__ $_                                            472
          Q__M__S__V __W__R__P_ _S__E__A__ T__V__Y__L __P__P__V_ _P__V__S__      45
                                <- L2 end

     5701 aggttgtaag cactgatgaa tatgtgtcac gcacaagcat ttattattat gctggcagtt
          K__V__V__S __T__D__E_ _Y__V__S__ R__T__S__I __Y__Y__Y_ _A__G__S__      65

     5761 ccagactttt ggctgttggc aatccatatt tttccatcaa aagtcccaat aacaataaaa
          S__R__L__L __A__V__G_ _N__P__Y__ F__S__I__K __S__P__N_ _N__N__K__      85

     5821 aagtattagt tcccaaggta tcaggcttac agtatagggt ctttagggtg cgtttacctg
          K__V__L__V __P__K__V_ _S__G__L__ Q__Y__R__V __F__R__V_ _R__L__P__      105

     5881 atcccaataa atttggtttt cctgatacat ctttttataa ccctgataca caacgtttgg
          D__P__N__K __F__G__F_ _P__D__T__ S__F__Y__N __P__D__T_ _Q__R__L__      125

     5941 tctgggcatg tgtaggcctt gaaataggta ggggacagcc attgggtgtt ggcgtaagtg
          V__W__A__C __V__G__L_ _E__I__G__ R__G__Q__P __L__G__V_ _G__V__S__      145
```

```
6001 gtcatcctta tttaaataaa tttgatgaca ctgaaaccag taacagatat cccgcacagc
     G__H__P__Y __L__N__K_ _F__D__D__ T__E__T__S __N__R__Y_ _P__A__Q__      165

6061 cagggtctga taacagggaa tgcttatcta tggattataa acaaacacaa ttatgtttaa
     P__G__S__D __N__R__E_ _C__L__S__ M__D__Y__K __Q__T__Q_ _L__C__L__      185

6121 ttggctgtaa acctcccact ggtgagcatt ggggtaaagg tgttgcctgt aacaataatg
     I__G__C__K __P__P__T_ _G__E__H__ W__G__K__G __V__A__C_ _N__N__N__      205

6181 cagctgctac tgattgtcct ccattggaac tttttaattc tattattgag gatggtgaca
     A__A__A__T __D__C__P_ _P__L__E__ L__F__N__S __I__I__E_ _D__G__D__      225

6241 tggtagatac agggtttgga tgcatggact ttggtacatt gcaggctaat aaaagtgatg
     M__V__D__T __G__F__G_ _C__M__D__ F__G__T__L __Q__A__N_ _K__S__D__      245

6301 tgcctattga tatttgtaac agtacatgca aatatccaga ttatttaaaa atggccagtg
     V__P__I__D __I__C__N_ _S__T__C__ K__Y__P__D __Y__L__K_ _M__A__S__      265

6361 aaccttatgg ggatagtttg ttctttttc ttagacgtga gcagatgttt gttagacact
     E__P__Y__G __D__S__L_ _F__F__F__ L__R__R__E __Q__M__F_ _V__R__H__      285

6421 tttttaatag ggctggaaaa cttggcgagg ctgtcccgga tgacctttat attaaagggt
     F__F__N__R __A__G__K_ _L__G__E__ A__V__P__D __D__L__Y_ _I__K__G__      305

6481 ccggtaatac tgcagttatc caaagtagtg catttttcc aactcctagt ggctctatag
     S__G__N__T __A__V__I_ _Q__S__S__ A__F__F__P __T__P__S_ _G__S__I__      325

6541 ttacctcaga atcacaatta tttaataagc cttattggct acagcgtgca caaggtcata
     V__T__S__E __S__Q__L_ _F__N__K__ P__Y__W__L __Q__R__A_ _Q__G__H__      345

6601 acaatggcat ttgctggggc aatcagttat ttgttaccgt ggttgatacc actcgtagca
     N__N__G__I __C__W__G_ _N__Q__L__ F__V__T__V __V__D__T_ _T__R__S__      365

6661 ctaatatgac attatgcact gaagtaacta aggaaggtac aTATAAAAat gataatttta
     T__N__M__T __L__C__T_ _E__V__T__ K__E__G__T __Y__K__N_ _D__N__F__      385
                                        signal ->

6721 aggaatatgt acgtcatgtt gaagaatatg acttacagtt tgtttttcag ctttgcaaaa
     K__E__Y__V __R__H__V_ _E__E__Y__ D__L__Q__F __V__F__Q_ _L__C__K__      405

6781 ttacactaac tgcagagata atgacatata tacatactat ggattcaat atttggagg
     I__T__L__T __A__E__I_ _M__T__Y__ I__H__T__M __D__S__N_ _I__L__E__      425

6841 actggcaatt tggtttaaca cctcctccgt ctgccagttt acaggacaca tatagatttg
     D__W__Q__F __G__L__T_ _P__P__P__ S__A__S__L __Q__D__T_ _Y__R__F__      445

6901 ttacctccca ggctattact tgccaaaaaa cagcaccccc taaagaaaag gaagatccat
     V__T__S__Q __A__I__T_ _C__Q__K__ T__A__P__P __K__E__K_ _E__D__P__      465

6961 taaataaata tacttttgg gaggttaact taaaggaaaa gttttctgca gatctagatc
     L__N__K__Y __T__F__W_ _E__V__N__ L__K__E__K __F__S__A_ _D__L__D__      485

7021 agtttccttt gggacgaaag tttttattac aatcaggcct taaagcaaag cccagactaa
     Q__F__P__L __G__R__K_ _F__L__L__ Q__S__G__L __K__A__K_ _P__R__L__      505

7081 aacgttcggc ccctactacc cgtgcaccat ccaccaaacg caaaaaggtt aaaaaaTAAt
     K__R__S__A __P__T__T_ _R__A__P__ S__T__K__R __K__K__V_ _K__K__$_      524
                                                              <- L1 end
```

❄ ❄ ❄ ❄ Bases 7141 to 7824 not shown. ❄ ❄ ❄ ❄

HPV-58 Variants

L1 nuc

```
                6666666667
                6667788880
                4991922281
                1271827816
HPV58           GGGGAGCAAA   5565-7139
IS068           A---------   6588-7036
IS573           ------A---   6588-7036
IS1021          ----G-A-G-   6588-7036
IS131           -----AAG--   6588-7036
IS404           ----GAAG--   6588-7036
IS417           -AAAGAAG-G   6588-7036
```

L1 aa

```
                3333444444
                5778122238
                9683201294
HPV58           VKGDIDSNLL
IS068           *---------
IS573           ------*---
IS1021          ----V-*-*-
IS131           -----N*D--
IS404           ----VN*D--
IS417           -*DNVN*D-*
```

```
LOCUS       HPV73MM9      458 bp ds-DNA             VRL       16-OCT-1994
DEFINITION  Human papillomavirus isolate MM9, partial L1 cds, My09/My11
            region.
ACCESSION   U12491
SOURCE      Human papillomavirus DNA recovered from a gential swab sample, isolate
            MM9 (PAP238a).
  ORGANISM  Human papillomavirus
            Viridae; ds-DNA nonenveloped viruses; Papovaviridae;
            Papillomavirus.
COMMENT     Naturally occurring variants of HPV73(MM9) have been studied in the
            L1 region.
            Reference: Stewart,A.-C.M., Eriksson,A.M., Manos,M.M., Munoz,N.,
                Bosch,F.X., Peto,J. and Wheeler,C.M.,
                J. Virol. 70, 3127-36 (1996).
                Variant Accession Nos. U45935-U45937
FEATURES             Location/Qualifiers
     CDS             <1..>458
                     /note="putative"
                     /product="major capsid protein"
                     /gene="L1"
                     /note="putative"
                     /codon_start=1
                     /translation="AQGHNNGICWHNQLFLTVVDTTRSTNFSVCVGTQASSSTTTYAN
                     SNFKEYLRHAEEFDLQFVLQLCKISLTTEVMTYIHSMNSTILEEWNFGLTPPPSGTLE
                     ETYRYVTSHAISCQRPQPPKETEDPYAKLSFWDVDLKEKFSAELDQYPLG"
BASE COUNT      145 a     79 c     84 g    150 t

        1 gcacagggtc ataataatgg tatttgttgg cataatcaat tatttttaac tgttgtagat
          _A__Q__G__ H__N__N__G __I__C__W_ _H__N__Q__ L__F__L__T __V__V__D_      20
  L1 cds ->
          -> MY11 PCR primer <-

       61 actactagaa gcactaattt ttctgtatgt gtaggtacac agctagtag ctctactaca
          _T__T__R__ S__T__N__F __S__V__C_ _V__G__T__ Q__A__S__S __S__T__T_      40

      121 acgtatgcca actctaattt taaggaatat ttaagacatg cagaagagtt tgatttacag
          _T__Y__A__ N__S__N__F __K__E__Y_ _L__R__H__ A__E__E__F __D__L__Q_      60

      181 tttgttcttc agttatgtaa aattagttta actactgagg taatgacata tatacattct
          _F__V__L__ Q__L__C__K __I__S__L_ _T__T__E__ V__M__T__Y __I__H__S_      80

      241 atgaattcta ctatattgga agagtggaat tttggtctta ccccaccacc gtcaggtact
          _M__N__S__ T__I__L__E __E__W__N_ _F__G__L__ T__P__P__P __S__G__T_      100

      301 ttagaggaaa catatagata tgtaacatca cattgctatta gttgccaacg tcctcaacct
          _L__E__E__ T__Y__R__Y __V__T__S_ _H__A__I__ S__C__Q__R __P__Q__P_      120

      361 cctaaagaaa cagaggaccc atatgccaag ctatcctttt gggatgtaga tcttaagaa
          _P__K__E__ T__E__D__P __Y__A__K_ _L__S__F__ W__D__V__D __L__K__E_      140

      421 aagttttctg cagaattaga tcagtatccc cttggacg
          _K__F__S__ A__E__L__D __Q__Y__P_ _L__G_        152
                                          L1 cds ->
                              -> MY09 PCR primer  <-
```

L1 nuc

```
                11113334
                04780371
                24073357
HPV73MM9        GGTCATGG   1-456
IS324           A---GGT-   1-456
IS223           --AT-GTA   1-456
IS601           -AAT-GTA   1-456
```

L1 aa

```
                00001111
                34560123
                48731159
HPV73MM9        QKFLLHEK
IS324           *---*QD-
IS223           --YF-QD*
IS601           -*YF-QD*
```

Sources for Variant Sequences

VARIANT	SOURCE	ACCESSION NUMBER(S)
BPV1R		
BPU23379	[34]	U23379
PPBVARA	[32]	L07797
PPBVARB	[32]	L07798
PPBVARC	[32]	L07799
PPBVARD	[32]	L07800
CRPV		
CRPVSyl	[41]	U09467, U09493-U09497
HPV1a		
HPV1var	[26]	U06714
HPV5		
E1	[22]	U49465
E5	[22]	U49477
E7	[22]	U49466
E9	[22]	U49467
E11	[22]	U49476
E13	[22]	U4979
E14	[22]	U4981
E16	[22]	U4978
E17	[22]	U49464
FC1	[22]	U49472
FC2	[22]	U49473
HPV5b	[43]	D90252
HPV5d	[31]	M18452, M18453, M18454, M22961
NAf1	[22]	U49469
NAf2	[22]	U4980
NAf3	[22]	U49474
NAf5	[22]	U49470
NA1	[22]	U49475
Sam4	[22]	U49468
WAf2	[22]	U49471
5a2	[3]	M72882, X74644, X74645, X64618
5a3	[4, 3]	M73052, M73053, M73458, X74646, X74647, X64620
5a4	[4, 3, 8]	M73054, M73055, M72884, M80460, X74648, X74649, X64619
5a5	[4, 3]	M73058, M73059, M7345X7, 74650, X74651, X64621
5a6	[4, 3, 8]	M73051, M73056, M73057, M73456, X74652, X74653, X64622
5a7	[3]	X64641
5a8	[3]	X64642
5a9	[3, 8]	X64643

Sources for Variant Sequences cont.

VARIANT	SOURCE	ACCESSION NUMBER(S)
HPV6		
6R	[21]	M22108
1082	[35]	L36837
1083	[35]	L36839
1084	[35]	L36840
1086	[35]	L36841
1094	[35]	L36842
Alaska-C36	[17]	
AM6-1	[13]	
B6-1	[13]	
B6-5	[13]	
B6-15/1	[13]	
BAM	[21]	MM22107
G6-6	[13]	
G6-42	[13]	
G6-78	[13]	
Georgia-B1	[17]	
Georgia-B4	[17]	
Georgia-B5	[17]	
Georgia-B6	[17]	
Georgia-G1	[17]	
Georgia-G2	[17]	
Georgia-G3	[17]	
Georgia-G4	[17]	
Georgia-G5	[17]	
Georgia-G6	[17]	
Georgia-G7	[17]	
HPV6A	[16]	L41216
HPV-6A	[13]	
HPV-6C	[13]	
IN6-6	[13]	
India-D4	[17]	
India-D5	[17]	
India-D7	[17]	
India-D9	[17]	
J6-8	[13]	
ML2	[21]	M22106
NY6-1	[13]	
NY6-16	[13]	
NY6-19	[13]	
PPHRRA	[21]	M20561
Philippines-A4	[17]	
Philippines-A6	[17]	
SN6-1	[13]	
SN6-3	[13]	
SN6-6A	[13]	
SN6-11	[13]	
T70	[7]	L22694
W50	[7]	L22693

Sources for Variant Sequences cont.

VARIANT	SOURCE	ACCESSION NUMBER(S)
HPV11		
6c	[28]	M26656
B11-7	[13]	
B11-15D2	[13]	
HPV11P	[25]	
HPV11VC	[25]	
J11-2	[13]	
G11-30	[13]	
G11-79	[13]	
NY11-9	[13]	
NY11-14	[13]	
PPH11A	[5]	J04351
SN11-1	[13]	
S11-27	[13]	
HPV16R		
114B	[23]	
114K	[23]	
1194	[19, 18]	
1268	[19]	
649	[19]	
77	[19]	
84	[19]	
A00446	[27]	
ALABAMA-B23	[17]	
ALABAMA-B24	[17]	
ALASKA-C32	[17]	
ALASKA-C51	[17]	
AM-6	[14]	L22662
AN-10	[14]	L22674
AN10a	[14]	L22673
AN-12	[14]	L22665
AP-4	[14]	L22669
B20	[6]	
B22	[6]	
B23	[6]	
B25	[6]	
B29	[6]	
B30	[6]	
B33	[6]	
B61	[6]	
BAVINC0	[1]	
BAVINC3	[1]	
BB1	[2, 15]	M83847
BB3	[2, 15]	M83846
BB8A	[2, 15]	M83844
BB11	[2, 15]	M95524
BB14	[2, 15]	M83848
BB19	[2, 15]	M83849
BB24	[2, 15]	M83834, M83845

Sources for Variant Sequences cont.

VARIANT	SOURCE	ACCESSION NUMBER(S)
HPV16R cont.		
BT6	[38]	
BT7	[38]	
BT8	[38]	
BT9	[38]	
BT10	[38]	
BT11	[38]	
BT12	[38]	
BT15	[38]	
BT20	[38]	
BT22	[38]	
BT23	[38]	
CASKI	[2]	M83776, M83840, M83871
DC141	[38]	
DC207	[38]	
DC212	[38]	
DC255	[38]	
DC269	[38]	
DT4	[38]	
DT24	[38]	
DT42	[38]	
F94_2	[9]	
F94_27	[9]	
F94_42	[9]	
F94_44	[9]	
F94_73	[9]	
F94_76	[9]	
F94_80	[9]	
F94_85	[9]	
F94_89	[9]	
F94_113	[9]	
F94_135	[9]	
F94_143	[9]	
F94_161	[9]	
F94_179	[9]	
GALUC64	[11]	U33069
GALUC73	[11]	U33068
GALUC91	[11]	U33119
GALUC106	[11]	U33066
GALUC136	[11]	U33067
GALUC153	[11]	U33065
GALU1601	[10]	U14511
GALU1603	[10]	U14512
GALU1607	[10]	U14513
GALU1627	[10]	U14514
GALU1629	[10]	U14515
GALU1649	[10]	U14516
GB1=WV_2965	[6]	
GB3=WV_3096	[6]	
GB4=WV_3097	[6]	
GB5=WV_3270	[6]	

Sources for Variant Sequences cont.

VARIANT	SOURCE	ACCESSION NUMBER(S)
HPV16R cont.		
GB6=WV_3471	[6]	
GB10	[2]	M83851
GB10=WV_4042	[6]	
GB11=WV_4043	[6]	
GB12=WV_4044	[6]	
GB13	[2]	M83850
GB13=WV_75	[6]	
GB14=WV_1024	[6]	
GB21A	[2]	M83852
GEORGIA-B9	[17]	
H84	[18]	
HDS69.10/16/89	[42]	
HDS81A.02/24/92	[42]	
HDS81A.06/19/91	[42]	
HDS81A.10/04/89	[42]	
HDS81A.11/07/90	[42]	
HDS81B.02/24/92	[42]	
HDS81B.06/19/91	[42]	
HDS81B.10/04/89	[42]	
HDS81B.11/07/90	[42]	
HDS84A.03/26/90	[42]	
HDS84A.07/23/90	[42]	
HDS84B.03/26/90	[42]	
HDS84C.03/26/90	[42]	
HDS84C.07/23/90	[42]	
HDS84D.07/23/90	[42]	
HDS85.06/12/90	[42]	
HDS106A.02/13/90	[42]	
HDS106A.05/08/90	[42]	
HDS106B.02/13/90	[42]	
HDS106C.05/08/90	[42]	
HDS109.07/25/90	[42]	
HDS109.11/30/90	[42]	
HDS117A.04/09/92	[42]	
HDS117A.08/12/92	[42]	
HDS117B.04/09/92	[42]	
HDS138.05/01/91	[42]	
HPK_IB	[6]	
HPK_II	[6]	
IND-4	[14]	L22667
IND-5	[14]	L22670
IND-7	[14]	L22666
J-5	[14]	L22668
MICHIGAN-B8	[17]	
MISSOURI-B11	[17]	
NM.4094	[44]	U34078,U34107,U34136,U34165
NM.9999	[44]	U34079,U34108,U34137,U34166
NM.T197	[44]	U34080,U34109,U34138,U34167
NM.T446	[44]	U34081,U34110,U34139,U34168
NM.T455	[44]	U34082,U34111,U34140,U34169

Sources for Variant Sequences cont.

VARIANT	SOURCE	ACCESSION NUMBER(S)
HPV16R cont.		
NM.T529	[44]	U34083,U34112,U34141,U34170
O162	[18]	
O191	[18]	
O204	[18]	
O356	[18]	
O424	[18]	
O425	[18]	
O462	[18]	
O489	[18]	
O512	[18]	
O559	[18]	
O561	[18]	
OR.0198	[44]	U34084,U34113,U34142,U34171
OR.1783	[44]	
OR.1905	[44]	U34085,U34114,U34143,U34172
OR.2087	[44]	U34086,U34115,U34144,U34173
OR.3136	[44]	U34087,U34116,U34145,U34174
OR.3473	[44]	U34088,U34117,U34146,U34175
OR.3759	[44]	U34089,U34118,U34147,U34176
OR.4072	[45]	
OR.4451	[45]	
OR.4541	[44]	U34090,U34119,U34148,U34177
OR.4716	[44]	U34091,U34120,U34149,U34178
OR.4724	[44]	U34092,U34121,U34150,U34179
OR.4997	[44]	U34093,U34122,U34151,U34180
OR.5110	[44]	U34094,U34123,U34152,U34181
OR.5428	[44]	U34095,U34124,U34153,U34182
OR.5691	[44]	
OR.6106	[44]	U34096,U34125,U34154,U34183
OR.6170	[44]	U34097,U34126,U34155,U34184
OR.6311	[44]	U34098,U34127,U34156,U34185
OR.7145	[44]	U34099,U34128,U34157,U34186
OR.7574	[44]	U34100,U34129,U34158,U34187
OR.7587	[44]	U34101,U34130,U34159,U34188
OR.7632	[44]	U34102,U34131,U34160,U34189
OR.7754	[45]	
OR.7875	[45]	
OR.7908	[45]	
OR.8160	[44]	U34103,U34132,U34161,U34190
OR.8329	[44]	U34104,U34133,U34162,U34191
OR.8392	[44]	
OR.8863	[45]	
OR.8987	[44]	U34105,U34134,U34163,U34192
OR.9237	[44]	U34106,U34135,U34164,U34193
P-1	[14]	L22663
PANAMA-136	[17]	
PANAMA-148	[17]	
PANAMA-156	[17]	
PANAMA-200	[17]	
PANAMA-206	[17]	

Sources for Variant Sequences cont.

VARIANT	SOURCE	ACCESSION NUMBER(S)
HPV16R cont.		
PANAMA-349	[17]	
PANAMA-A12	[17]	
PANAMA-A13	[17]	
PANAMA-B20	[17]	
PANAMA-B4	[17]	
PANAMA-C10	[17]	
PANAMA-C20	[17]	
PANAMA-D6	[17]	
PANAMA-F15	[17]	
PANAMA-F30	[17]	
PANAMA-F38	[17]	
PANAMA-F43	[17]	
S23	[33]	
S27	[33]	
S29	[33]	
S83	[33]	
S93	[33]	
S99	[33]	
S108	[33]	
SA-9	[14]	L22661
SB1	[2, 15]	M83838
SB2	[2, 15]	M83870
SB5	[2, 15]	M83778, M83869
SB7	[2]	M83839,M83875
SB10	[2, 15]	M83877
SB13	[2, 15]	M83836, M83872, M83873
SB14	[2, 15]	M83837
SB16	[2, 15]	M83842, M83880
SB17	[2, 15]	M83876
SB21	[2, 15]	M83841
SB21A	[2]	M83843, M83878, M83879
SH-2a	[14]	L22671
SIHA	[2]	M83777, M83874
SL-6a	[14]	L22664
SN-4	[14]	L22672
SS1A	[2]	M83835
T3	[33]	
T17	[33]	
T45	[33]	
T49	[33]	
TB1	[2]	M83885
TB1=A6	[6]	
TB3=A16	[6]	
TB4	[2]	M83855, M83881
TB4=A21	[6]	
TB5=A24	[6]	
TB6=A27	[6]	
TB7=A32	[6]	
TB8	[2]	M83853
TB8=A36	[6]	

Sources for Variant Sequences cont.

VARIANT	SOURCE	ACCESSION NUMBER(S)
HPV16R cont.		
TB9=A42	[6]	
TB11=A46	[6]	
TB12=A50	[6]	
TB13	[2]	M83854, M83884
TB14=A54	[6]	
TB15	[2]	M83859
TB15=A56	[6]	
TB16	[2]	M83858, M83883
TB16=A61	[6]	
TB17=A62	[6]	
TS3	[2]	M83865
TS3A	[2]	M83866
TS3B	[2]	M83867
TS3C	[2]	M83868
TS5	[2]	M83860
TS5A	[2]	M83861
TS5B	[2]	M83862, M83863
TS5C	[2]	M83864
TS6A	[2]	M83856
TS6B	[2]	M83857
HPV18		
AM18-1	[29]	L22615
A4	[40]	
A5	[40]	
A8	[40]	
A13	[40]	
A20	[40]	
A23	[40]	
A29	[40]	
A52	[40]	
A18-1	[30]	U59162
A18-2	[30]	U59163
C18-5	[30]	U59164
C18-7	[30]	U59165
B6	[40]	
B16	[40]	
B17	[40]	
B21	[40]	
B37	[40]	
B44	[40]	
B54	[40]	
C4-1	[29]	L22635
C41	[40]	
G18-1	[29]	L22617
G18-2	[29]	L22619
HELA	[40]	
HPV18v1E67	[20]	M20324
HPV18v2E67	[20]	M20325

Sources for Variant Sequences cont.

VARIANT	SOURCE	ACCESSION NUMBER(S)
HPV18 cont.		
HPV18v3E67	[37]	X04773
HPV18v4E6	[24]	X04354
HPV18v5E67	[36]	M26798
IN18-1	[30]	U59161
IS002	[39]	U45889
IS168	[39]	U45892
IS172	[39]	U45894
IS326	[39]	U45890
IS664	[39]	U45891
IS768	[39]	U45893
J18-1	[29]	L22616
NY18-1	[30]	U59156
NY18-2	[30]	U59156
NY18-6	[30]	U59158
NY18-11	[30]	U59159
NY18-13	[30]	U59160
Sc18-2	[29]	L22622
Sc18-3	[29]	L22618
Sc18-4a	[29]	L22621
Sc18-4b	[29]	L22623
SW756	[40]	
T18-3	[29]	L22626
T18-4	[29]	L22628
T18-5	[29]	L22625
T18-7	[29]	L22624
T18-8	[29]	L22627
T18-9	[29]	L22620
T18-10	[29]	L22634
T18-12	[29]	L22633
T18-16	[29]	L22631
T18-17	[29]	L22632
T18-18a	[29]	L22629
T18-18b	[29]	L22630
2	[12]	
3	[12]	
4	[12]	
8	[12]	
12	[12]	
14	[12]	
16	[12]	
23	[12]	
24	[12]	
25	[12]	
27	[12]	
33	[12]	
40	[12]	
42	[12]	
47	[12]	

Sources for Variant Sequences cont.

VARIANT	SOURCE	ACCESSION NUMBER(S)
HPV18 cont.		
101	[40]	
161	[40]	
173	[40]	
2031	[40]	-
3068	[40]	
4039	[40]	
4040	[40]	
4050	[40]	
4052	[40]	
4055	[40]	
5076	[40]	
HPV33		
IS267	[39]	U45895
IS549	[39]	U45896
IS827	[39]	U45897
HPV39		
IS015	[39]	U45903
IS065	[39]	U45902
IS073	[39]	U45900
IS114	[39]	U45901
IS214	[39]	U45904
IS270	[39]	U45899
IS351	[39]	U45905
HPV45		
IS040	[39]	U45907
IS252	[39]	U45909
IS278	[39]	U45912
IS282	[39]	U45908
IS761	[39]	U45916
IS788	[39]	U45915
IS816	[39]	U45906
IS819	[39]	U45910
IS871	[39]	U45914
IS894	[39]	U45913
IS949	[39]	U45911
HPV52		
IS121	[39]	U45919
IS370	[39]	U45920
IS464	[39]	U45923
IS705	[39]	U45918
IS1001	[39]	U45921
IS1015	[39]	U45922

Sources for Variant Sequences cont.

VARIANT	SOURCE	ACCESSION NUMBER(S)
HPV58		
IS068	[39]	U45924
IS131	[39]	U45927
IS404	[39]	U45928
IS417	[39]	U45929
IS573	[39]	U45925
IS1021	[39]	U45926
HPV73MM9		
IS223	[39]	U45936
IS324	[39]	U45935
IS601	[39]	U45937

References

[1] P.J. Bavin, P.G. Walker, and V.C. Emery. Sequence microheterogeneity in the long control region of clinical isolates of human papillomavirus type 16. *Journal of Medical Virology*, 39(4):267–72, 1993. Sequences constructed by editing reference sequence (HPV16R) according to figures in article, with corrections from authors (V.C. Emery p.c.). Medline: 93260414.

[2] S. Y. Chan, L. Ho, C. K. Ong, V. Chow, B. Drescher, M. Durst, Meulen ter J., L. Villa, J. Luande, and H. N. Mgaya et al. Molecular variants of human papillomavirus type 16 from four continents suggest ancient pandemic spread of the virus and its coevolution with humankind. *J Virol*, 66:2057–66, 1992. OTE: (Medline: 92194439), (Genbank: M83776 M83777 M83778 M83779 M83780 M83781 M83782 M83783 M83784 M83785 M83786 M83787 M83788 M83789 M83790 M83791 M83792 M83793 M83794 M83795 M83796 M83797 M83798 M83799 M83800 M83801 M83802 M83803 M83804 M83840).

[3] M. C. Deau, M. Favre, S. Jablonska, L. A. Rueda, and G. Orth. Genetic heterogeneity of oncogenic human papillomavirus type 5 HPV5 and phylogeny of HPV5 variants associated with epidermodysplasia verruciformis. *J Clin Microbiol*, 31:2918–26, 1993. Sequences obtained from GenBank. Medline: 94086789, Genbank: X74618 X74619 X74620 X74621 X74622 X74641 X74642 X74643 X74644 X74645 X74646 X74647 X74648 X74649 X74650 X74651 X74652 X74653.

[4] M. C. Deau, M. Favre, and G. Orth. Genetic heterogeneity among human papillomaviruses HPV associated with epidermodysplasia verruciformis: evidence for multiple allelic forms of HPV5 and HPV8 E6 genes. *Virology*, 184:492–503, 1991. Sequences obtained from GenBank. Medline: 91361540, Genbank: M72880 M72881 M72882 M72883 M72884 M73051 M73052 M73053 M73054 M73055 M73056 M73057 M73058 M73059 M73456 M73457 M73458.

[5] S. C. Dollard, L. T. Chow, J. W. Kreider, T. R. Broker, N. L. Lill, and M. K. Howett. Characterization of an HPV type 11 isolate propagated in human foreskin implants in nude mice. *Virology*, 171:294–7, 1989. Sequences obtained from GenBank. Medline: 89299475, Genbank: J04351 J04351.

[6] D. Eschle, M. Durst, J. ter Meulen, J. Luande, HC. Eberhardt, M. Pawlita, and L. Gissmann. Geographical dependence of sequence variation in the E7 gene of human papillomavirus type 16. *Journal of General Virology*, 73(7):1829–32, 1992. Sequences constructed by editing reference sequence (HPV16R) according to figures in article. Medline: 92333270.

[7] A. Farr, H. Wang, M. S. Kasher, and A. Roman. Relative enhancer activity and transforming potential of authentic human papillomavirus type 6 genomes from benign and malignant lesions. *J Gen Virol*, 72 Pt 3:519–26, 1991. Sequences obtained from GenBank. Medline: 91170932 GenBank: L22693 L22694.

[8] M. Favre. 1995. Personal Communication.

[9] Y. Fujinaga, K. Okazawa, A. Nishikawa, Y. Yamakawa, M. Fukushima, I. Kato, and K. Fujinaga. Sequence variation of human papillomavirus type 16 E7 in preinvasive and invasive cervical neoplasias. *Virus Genes*, 9(1):85–92, 1994. Sequences constructed by editing reference sequence (HPV16R) according to figures in article.

[10] D. G. Haegert, D. F. Galutira, and B. H. Younghusband. Sequence variation in the E6 gene of human papillomavirus type 16. Unpublished. Sequences obtained from GenBank. Genbank: U14511 U14512 U14513 U14514 U14515 U14516, 1995.

[11] D. G. Haegert, B. H. Younghusband, D. F. Galutira, and K. N. Egger. Sequence variations, and phylogenetic analysis of human papillomavirus type 16 in the canadian province of newfoundland. Unpublished. Sequences obtained from GenBank. Genbank: U33065 U33066 U33067 U33068 U33069 U33119, 1995.

[12] J. L. Hecht, A. S. Kadish, G. Jiang, and R. D. Burk. Genetic characterization of the human papillomavirus HPV 18 E2 gene in clinical specimens suggests the presence of a subtype with decreased oncogenic potential. *Int J Cancer*, 60:369–376, 1995. Sequences constructed by editing reference sequence according to figures in article. Medline: 95130263.

[13] P. A. Heinzel, S. Y. Chan, L. Ho, M. O'Connor, P. Balaram, M. S. Campo, K. Fujinaga, N. Kiviat, J. Kuypers, and H. Pfister et al. Variation of human papillomavirus type 6 HPV-6 and HPV-11 genomes sampled throughout the world. *J Clin Microbiol*, 33:1746–1754, 1995. Sequences obtained from authors.

[14] L. Ho, S. Y. Chan, R. D. Burk, B. C. Das, K. Fujinaga, J. P. Icenogle, T. Kahn, N. Kiviat, W. Lancaster, and P. Mavromara-Nazos et al. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J Virol*, 67:6413–23, 1993. Sequences obtained from GenBank. Medline: 94016822, Genbank: L22661 L22662 L22663 L22664 L22665 L22666 L22667 L22668 L22669 L22670 L22672 L22673 L22674 L22771.

[15] L. Ho, S.Y. Chan, V. Chow, T. Chong, S.K. Tay, L.L. Villa, and H.U. Bernard. Sequence variants of human papillomavirus type 16 in clinical samples permit verification and extension of epidemiological studies and construction of a phylogenetic tree. *Journal of Clinical Microbiology*, 29(9):1765–72, 1991. Medline: 92129569 GenBank: M83776 M83777 M83778 M83834 M83835 M83836 M83837 M83838 M83839 M83841.

[16] K. J. Hofmann, J. C. Cook, J. G. Joyce, D. R. Brown, L. D. Schultz, H. A. George, M. Rosolowsky, K. H. Fife, and K. U. Jansen. Sequence determination of human papillomavirus type 6a and assembly of virus-like particles in Saccharomyces cerevisiae. *Virology*, 209:506–518, 1995. Sequences obtained from GenBank. Medline: 95297152, Genbank: L41216.

[17] J. P. Icenogle, P. Sathya, D. L. Miller, R. A. Tucker, and W. E. Rawls. Nucleotide and amino acid sequence variation in the L1 and E7 open reading frames of human papillomavirus type 6 and type 16. *Virology*, 184:101–7, 1991. Sequences constructed by editing reference sequence according to figures in article. Medline: 91335738.

[18] J.P. Icenogle, K. A. Clancy, and S. Y. Lin. Sequence variation in the capsid protein genes of human papillomavirus type 16 and type 31. *Virology*, 214:664–669, 1995. Medline: 96130211, Genbank: K02718 U37217 U37410.

[19] J.P. Icenogle, M. Laga, D. Miller, A.T. Manoka, R.A. Tucker, and W.C. Reeves. Genotypes and sequence variants of human papillomavirus DNAs from human immunodeficiency virus type 1-infected women with cervical intraepithelial neoplasia. *Journal of Infectious Diseases*, 166(6):1210–6, 1992. Sequences constructed by editing reference sequence (HPV16R) according to figures in article. Medline: 93056684.

[20] Y. Inagaki, Y. Tsunokawa, N. Takebe, H. Nawa, S. Nakanishi, M. Terada, and T. Sugimura. Nucleotide sequences of cDNAs for human papillomavirus type 18 transcripts in HeLa cells. *J Virol*, 62:1640–6, 1988. Sequences obtained from GenBank. Medline: 88188247, Genbank: M20324 M20325.

[21] M. S. Kasher and A. Roman. Characterization of human papillomavirus type 6b DNA isolated from an invasive squamous carcinoma of the vulva. *Virology*, 165:225–33, 1988. Sequences obtained from GenBank. Medline: 88265862 GenBank: M20561.

[22] M. Kawase, G. Orth, S. Jablonska, C. Blanchet-Bardon, L.-A. Rueda, and M. Favre. Variability and phylogeny of the L1 capsid protein gene of human papillomavirus type 5: contribution of clusters of nonsynonymous mutations and of a 30-nucleotide duplication. *Virology*, 221(1):189–198, 1996.

[23] R. Kirnbauer, J. Taub, H. Greenstone, R. Roden, M. Durst, L. Gissmann, and D.R.; Schiller JT Lowy. Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like

particles. *Journal of Virology*, 67(12):6929–36, 1993. Sequences constructed by editing reference sequence (HPV16R) according to figures in article. Medline: 94047301.

[24] G. Matlashewski, L. Banks, J. Wu-Liao, P. Spence, D. Pim, and L. Crawford. The expression of human papillomavirus type 18 E6 protein in bacteria and the production of anti-E6 antibodies. *J Gen Virol*, 67 Pt 9:1909–16, 1986. Sequences obtained from GenBank. Medline: 86306665, Genbank: X04354.

[25] R. C. McGlennen, J. Ghai, R. S. Ostrow, K. LaBresh, J. F. Schneider, and A. J. Faras. Cellular transformation by a unique isolate of human papillomavirus type 11. *Cancer Res*, 52:5872–8, 1992. Sequences constructed by editing reference sequence according to figures in article, with corrections from authors (R.S. Ostrow p.c.). Medline: 93007981.

[26] J. Meissner. Complete nucleotide sequencing of an HPV-1a variant and determination of extant errors in the prototype HPV-1a sequence. *Virus Genes*, 9:189–191, 1995. Sequences obtained from GenBank. Medline: 95250312 GenBank: U06714.

[27] G. Meneguzzi. Pharmaceutical composition useful in the prevention or treatment of papillomavirus-induced tumours. Patent: WO 9010459-A 21 20-SEP-1990; Transgene S.A. Sequence obtained from GenBank. Acc. No. A00446.

[28] L. Metcalfe, S. L. Chen, and P. Mounts. Structural analysis of human papillomavirus type 6c isolates from condyloma acuminatum and juvenile-onset and adult-onset laryngeal papillomata. *Virus Genes*, 3:11–27, 1989. Sequences obtained from GenBank. Medline: 90051107 GenBank: M26656.

[29] C. K. Ong, S. Y. Chan, M. S. Campo, K. Fujinaga, P. Mavromara-Nazos, V. Labropoulou, H. Pfister, S. K. Tay, J. ter Meulen, and L. L. Villa et al. Evolution of human papillomavirus type 18: an ancient phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic groups. *J Virol*, 67:6424–31, 1993. Sequences obtained from GenBank. Medline: 94016823, Genbank: L22615 L22616 L22617 L22618 L22619 L22620 L22621 L22622 L22623 L22624 L22625 L22626 L22627 L22628 L22629 L22630 L22631 L22632 L22633 L22634 L22635.

[30] C.-K. Ong, S. Nee, A. Rambaut, H.-U. Bernard, and P. H. Harvey. Sequences obtained from GenBank, accession numbers U59156–U59165, 1996.

[31] R. S. Ostrow, K. R. Zachow, and A. J. Faras. Molecular cloning and nucleotide sequence analysis of several naturally occurring HPV-5 deletion mutant genomes. *Virology*, 158:235–8, 1987. Sequences obtained from GenBank. Medline: 87207667 GenBank: M22961.

[32] N. Otten, C. von Tscharner, S. Lazary, D. F. Antczak, and H. Gerber. DNA of bovine papillomavirus type 1 and 2 in equine sarcoids: PCR detection and direct sequencing. *Arch Virol*, 132:121–31, 1993. Sequences obtained from GenBank. Medline: 93356631 GenBank: L07797 L07798 L07799 L07800.

[33] P. Pushko, T. Sasagawa, J. Cuzick, and L. Crawford. Sequence variation in the capsid protein genes of human papillomavirus type 16. *Journal of General Virology*, 75(4):911–6, 1994. Sequences constructed by editing reference sequence (HPV16R) according to figures in article, with corrections to sequence numbering from authors (L. Crawford p.c.).

[34] S. W. Reid, K. T. Smith, and W. F. Jarrett. Detection, cloning and characterisation of papillomaviral DNA present in sarcoid tumours of Equus asinus. *Vet Rec*, 135:430–432, 1994. Sequences obtained from GenBank. Medline: 95149443 GenBank: U32584.

[35] A. Roman and D. Brown. Sequence variation in the extreme 5' end of the human papillomavirus type 6a long control region. *J Infect Dis*, 171:697–700, 1995. Sequences obtained from GenBank. Medline: 95181882, Genbank: L22693 L36837 L36839 L36840 L36841 L36842.

[36] A. Schneider-Gadicke and E. Schwarz. Different human cervical carcinoma cell lines show similar transcription patterns of human papillomavirus type 18 early genes. *EMBO J*, 5:2285–92, 1986. Sequences obtained from GenBank. Medline: 87053870 GenBank: M26798.

[37] K. Seedorf, T. Oltersdorf, G. Krammer, and W. Rowekamp. Identification of early proteins of the human papilloma viruses type 16 HPV 16 and type 18 HPV 18 in cervical carcinoma cells. *EMBO J*, 6:139–44, 1987. Sequences obtained from GenBank. Medline: 87218459, Genbank: X04773.

[38] H.L. Smits, K.F. Traanberg, M.R. Krul, P.R. Prussia, C.L. Kuiken, M.F. Jebbink, J.A. Kleyne, R.H. van den Berg, B. Capone, A. de Bruyn, et al. Identification of a unique group of human papillomavirus type 16 sequence variants among clinical isolates from Barbados. *Journal of General Virology*, 75(9):2457–62, 1994. Sequences constructed by editing reference sequence (HPV16R) according to figures in article, with corrections to sequence numbering from authors (J. ter Schegget p.c.). Medline: 94358750.

[39] A.-C. M. Stewart, A. M. Eriksson, M. M. Manos, N. Munoz, F.X. Bosch, J. Peto, and C. M. Wheeler. Intratype variation in 12 human papillomavirus types: a worldwide perspective. *J Virol*, 70(5):3127–3136, 1996. Sequences obtained from GenBank. Medline: 96186743 GenBank: U45889 U45890 U45891 U45892 U45893 U45894.

[40] J. ter Meulen, A. C. Schweigler, H. C. Eberhardt, J. Luande, H. N. Mgaya, M. Muller, C. Bleul, V. Ulken, H. Ikenberg, and M. Pawlita et al. Sequence variation in the E7 gene of human papillomavirus type 18 in tumor and non-tumor patients and antibody response to a conserved seroreactive epitope. *Int J Cancer*, 53:257–9, 1993. Sequences constructed by editing reference sequence according to figures in article. Medline: 93146645.

[41] X. Wu, W. Xiao, and J. L. Brandsma. Papilloma formation by cottontail rabbit papillomavirus requires E1 and E2 regulatory genes in addition to E6 and E7 transforming genes. *J Virol*, 68:6097–6102, 1994. Sequences obtained from GenBank. Medline: 94335133 GenBank: U09467 U09493 U09494 U09495 U09496 U09497.

[42] L.F. Xi, G.W. Demers, L.A. Koutsky, N.B. Kiviat, J. Kuypers, D.H. Watts, K.K. Holmes, and D.A. Galloway. Analysis of human papillomavirus type 16 variants indicates establishment of a persistant infection. *J. Infect. Dis.*, 172:747–755, 1995. Sequences constructed by editing reference sequence (HPV16R) according to figures in article. Medline: 95386966.

[43] Y. Yabe, A. Sakai, T. Hitsumoto, H. Kato, and H. Ogura. A subtype of human papillomavirus 5 HPV-5b and its subgenomic segment amplified in a carcinoma: nucleotide sequences and genomic organizations. *Virology*, 183:793–8, 1991. Sequences obtained from GenBank. Medline: 91306467 , Genbank: D90252 D90252 D90252 M55473 M55474 M55475 M55476 M55477 M55478 M64316 S42998 S43038.

[44] T. Yamada, C.M. Wheeler, A.L. Halpern, A.-C.M. Stewart, A. Hildesheim, B.B. Rush, and S.A. Jenison. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2 and L1 coding segments. *Journal of Virology*, 69(12):7743–7753, 1995. Sequences available from GenBank. Medline: 96079021 GenBank: U34078–U34193.

[45] T. Yamada, C.M. Wheeler, A.L. Halpern, A.-C.M. Stewart, A. Hildesheim, B.B. Rush, and S.A. Jenison. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2 and L1 coding segments. *Journal of Virology*, 69(12):7743–7753, 1995. Data not shown in article; sequences obtained from authors.