

## Human papillomavirus type 31 (HPV-31) complete genome

GenBank: J04353.1

[FASTA Graphics](#)

[Go to:](#)

LOCUS PPH31A 7912 bp DNA circular VRL 18-MAR-1994  
DEFINITION Human papillomavirus type 31 (HPV-31) complete genome.  
ACCESSION J04353  
VERSION J04353.1  
KEYWORDS complete genome.  
SOURCE human papillomavirus 31  
ORGANISM [human papillomavirus 31](#)  
Viruses; Monodnaviria; Shotokuvirae; Cossaviricota;  
Papovaviricetes; Zurhausenvirales; Papillomaviridae;  
Firstpapillomavirinae; Alphapapillomavirus.  
REFERENCE 1 (bases 1 to 7912)  
AUTHORS Goldsborough,M.D., DiSilvestre,D., Temple,G.F. and Lorincz,A.T.  
TITLE Nucleotide sequence of human papillomavirus type 31: a cervical  
neoplasia-associated virus  
JOURNAL Virology 171 (1), 306-311 (1989)  
PUBMED [2545036](#)  
COMMENT Draft entry and computer-readable copy of sequence [1] kindly  
submitted by M.D.Goldsborough, 05-JUL-1989.  
FEATURES Location/Qualifiers  
source 1..7912  
/organism="human papillomavirus 31"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:[10585](#)"  
[regulatory](#) 19..24  
/regulatory\_class="TATA\_box"  
[regulatory](#) 69..74  
/regulatory\_class="TATA\_box"  
[gene](#) 108..557  
/gene="E6"  
[CDS](#) 108..557  
/gene="E6"  
/note="ORF E6 from bp 39 to 557"  
/codon\_start=1  
/product="transforming protein"  
/protein\_id="[AAA46950.1](#)"  
/translation="MFKNPAERPRKLHELSSALEIPYDELRLNCVYCKGQLTETEVLDF  
FAFTDLTIVYRDDTPHGVCTKCLRFYSKVSEFRWYRYSVYGTTLKLTNKGICDLLIR  
CITCQRPLCPEEKQRHLDKRRFHNIGGRWTGRCIACWRRPRTETQV"  
[misc feature](#) 228..236  
/gene="E6"

```

                                /standard_name="Splice donor"
misc feature    403..414
                                /gene="E6"
                                /standard_name="Splice acceptor"
gene           560..856
                                /gene="E7"
CDS           560..856
                                /gene="E7"
                                /note="ORF E7 from bp 545 to 856"
                                /codon_start=1
                                /product="transforming protein"
                                /protein_id="AAA46951.1"
                                /translation="MRGETPTLQDYVLDLQPEATDLHCYEQLPSSDEEDVIDSPAGQ
                                AEPDTSNYNIVTFCCQCKSTLRCLCVQSTQVDIRILQELLMGSGFIVPCNCSTRL"
gene           862..2751
                                /gene="E1"
CDS           862..2751
                                /gene="E1"
                                /note="ORF E1 from bp 850 to bp 2751"
                                /codon_start=1
                                /product="replication protein"
                                /protein_id="AAA46952.1"
                                /translation="MADPAGTDGEGTGCNGWFYVEAVIDRQTGDNISEDENEDSSDTG
                                EDMVDFIDNCNVYNNQAEAETAQALFHAQEAEEHAEAVQVLKRKYVGSPLSDISSCVD
                                YNISPRKKAICIENNSKTAKRRLFELPDSGYGNTEVETQQMVQVEEQQTTLSCNGSDG
                                THSERENETPTRNILQVLKTSNGKAAMLGKFKELYGVSELMELIRPFQSNKSTCTDWCV
                                AAFGVTGTVAEGFKTLQPYCLYCHLQSLACSWGVMMLMVRFKCAKNRITIEKLLK
                                LLCISTNCMLIQPPKLRSTAAALYWYRTGMSNISDVYGETPEWIERQTVLQHSFNDDT
                                FDLQMVQWAYDNDVMDSEIAYKYQLADSDSNACAFKLSNSQAKIVKDCGTMCRHY
                                KRAEKRQMSMGQWIKSRCDKVSDEGDWRDIVKFLRYQQIEFVSFLSALKLFLKGVPKK
                                NCILIHGAPNTGKSYFGMSLISFLQGCII SYANSKSHFWLQPLADAKIGMLDDATTPC
                                WHYIDNYLRNALDGNPVSIDVKHKALMQLKCPLLITSNINAGKDDRWPYLHSRLVVF
                                TFPNPFPPFDKNGNPVYELSDKNWKSFFSRTWCRLNLHEEEDKENDGDSFSTFKCVSGQ
                                NIRTL"
gene           2693..3811
                                /gene="E2"
CDS           2693..3811
                                /gene="E2"
                                /note="ORF E2 from bp 2663 to 3811"
                                /codon_start=1
                                /product="regulatory protein"
                                /protein_id="AAA46953.1"
                                /translation="METLSQRLNVCQDKILEHYENDSKRLCDHIDYWKHIRLECVMY
                                KAREMGIHSINHQVVPALSVSKAKALQAIELQMMLETLNNTTEYKNEDWTMQQTSLELY

```

LTAPTGCLKKHGYSVEVQFDGDVHNTMHYTNWKFYLYCIDGQCTVVEGQVNCCKGIYYV  
HEGHITYFVNFTEEAKKYGTGKKWEVHAGGQVIVFPESVFSDEISFAGIVTKLPTAN  
NTTTSNSKTCALGTSEGVRATTSTKRPRTEPEHRNTHHPNKLLRGDSVDSVNCGVIS  
AAACTNQTRAVSCPATPTPIIHLKGDANILKCLRYRLSKYKQLYEQVSSTWHWTCTDGK  
HKNAIVTLTYISTSQRDDFLNTVKIPNTVSVSTGYMTI"

gene

3270..3578

/gene="E4"

CDS

<3270..3578

/gene="E4"

/note="ORF E4 from bp 3270 to 3578"

/codon\_start=1

/protein\_id="[AAA46949.1](#)"

/translation="LFFLNLYLAVTKYPLLGLLQSYQQPTTPPHRIKPKAPWAPVKVC  
GRRRLSDQEQSQSTETPTTPTSCCEATPWTVSTVGLSVQLHAQTKQGLSVVLQLHL  
"

gene

3816..4070

/gene="E5"

CDS

3816..4070

/gene="E5"

/note="ORF E5 from bp 3804 to 4070"

/codon\_start=1

/protein\_id="[AAA46954.1](#)"

/translation="MIELNISTVSIIVLCFLLCFCVLLFVCLVIRPLVLSVSVYATLLL  
LIVILWVIATSPLRCFCIYVVFYIPLFVIHTHASFLSQQ"

repeat region

4099..4134

regulatory

4138..4143

/regulatory\_class="polyA\_signal\_sequence"

/note="putative"

repeat region

4143..4158

gene

4171..5571

/gene="L2"

CDS

4171..5571

/gene="L2"

/note="ORF L2 from bp 4060 to bp 5571"

/codon\_start=1

/product="minor capsid protein"

/protein\_id="[AAA46955.1](#)"

/translation="MRSKRSTKRTKRASATQLYQTCKAAGTCPSDVIKIEHTTIADQ  
ILRYGSMGVFFGGLGIGSGSGTGGRGTGYVPLSTRPSTVSEASIPRPPVSIDPVGPLD  
PSIVSLVEESGIVDVGAPAPIPHPPTTSGFDIATTADTTPAILDVTSTHENPTFTD  
PSVLQPPTPAETSGHLLSSSSISTHNYEEIPMDTFIVSTNNENITSSSTPIPGVRRPA  
RLGLYSKATQQVKVIDPTFLSAPKQLITYENPAYETVNAEESLYFSNTSHNIAPDPDF  
LDIIALHRPALTSRRNTVRYRSLGNKQTLRTRSGATIGARVHYYYDISSINPAGESIE  
MQPLGASATTTSTLNDGLYDIYADTDFTVDTPATHNVSPSTAVQSTSAVSAYVPTNTT

VPLSTGFDIPIFSGPDVPIEHAPTQVFFPLAPTTQVVSIFVDGGDFYLHPSYYMLKR  
RRKRVSYYFFTDVSVAA"

gene 5552..7066

/gene="L1"

CDS 5552..7066

/gene="L1"

/note="ORF L1 from bp 5516 to 7066"

/codon\_start=1

/product="major capsid protein"

/protein\_id="AAA46956.1"

/translation="MSLWRPSEATVYLPVPVSKVVSTDEYVTRTNIYYHAGSARLLT  
VGHPYYSIPKSDNPKKIVVPKVSGLQYRVFRVRLPDPNKFQFPDTSFYNPETQRLVWA  
CVGLEVGRGQPLGVGISGHPLLNKFDDTENSRYAGGPGTDNRECISMDYKQTQLCLL  
GCKPPIGEHWGKSPCSNNAITPGDCPPELEKNSVIQDGMVDTGFGAMDFTALQDTK  
SNVPLDICNSICKYPDY LKMVAEPYGDTLFFYLRRQMFVRHFFNRSQVTSVPTDL  
YIKGSGSTATLANSTYFPTPSGSMVTSDAQIFNKPYWMQRAQGHNGICWGNQLFVTV  
VDTRSTNMSVCAAIANSDDTFKSSNFKEYLRHGEEFDLQFIFQLCKITLSADIMTYI  
HSMNPAILEDWNFGLTTPPSGSLEDYRFVTSQAITCQKTAPQPKEDPFKDYVFEV  
NLKEKFSADLDQFPLGRKFLQAGYRARPKFKAGKRSAPSASTTTTAKRKKTKK"

regulatory 7227..7231

/regulatory\_class="polyA\_signal\_sequence"

/note="putative"

repeat region 7291..7302

repeat region 7314..7333

/note="putative"

protein bind 7406..7420

/standard\_name="glucocorticoid responsive element"

/bound\_moiety="hormone receptor"

protein bind 7477..7488

/bound\_moiety="E2"

/function="gene transcription"

regulatory 7542..7549

/regulatory\_class="enhancer"

/standard\_name="keratinocyte-dependant enhancer"

protein bind 7868..7879

/bound\_moiety="E2"

/function="gene transcription"

#### ORIGIN

1 taataataat aatcttagta taaaaaagta gggagtgacc gaaagtggg aaccgaaaac  
61 ggttggtata taaagcacat agtattttgt gcaaacctac agacgccatg ttcaaaaatc  
121 ctgcagaaaag acctcggaaa ttgcatgaac taagctcggc attggaaata ccctacgatg  
181 aactaagatt gaattgtgtc tactgcaaag gtcagttaac agaaacagag gtattagatt  
241 ttgcatttac agatttaaca atagtatata gggacgacac accacacgga gtgtgtacaa  
301 aatgtttaag attttattca aaagtaagtg aatttagatg gtatagatat agtgtgtatg

361 gaacaacatt agaaaaattg acaaacaaaag gtatatgtga tttgttaatt aggtgtataa  
421 cgtgtcaaaag accgttgtgt ccagaagaaa aacaaagaca tttggataaa aagaaacgat  
481 tccacaacat aggaggaag tggacaggac gttgcatagc atgttggaga agacctcgta  
541 ctgaaaccca agtgtaaaca tgcgtggaga aacacctacg ttgcaagact atgtgttaga  
601 tttgcaacct gaggcaactg acctccactg ttatgagcaa ttacccgaca gctcagatga  
661 ggaggatgtc atagacagtc cagctggaca agcagaaccg gacacatcca attacaatat  
721 cgttaccttt tgttgtcagt gtaagtctac acttcgtttg tgtgtacaga gcacacaagt  
781 agatattcgc atattgcaag agctgttaat gggctcattt ggaatcgtgt gcccactg  
841 ttctactaga ctgtaactac aatggctgat ccagcaggta cagatgggga ggggacggga  
901 tgcaatggtt ggttttatgt agaagcagta attgacagac agacagggga caacatttca  
961 gaggacgaaa atgaagacag tagtgatact ggggaggata tggttgactt tattgacaat  
1021 tgtaatgtat acaacaatca ggcagaagca gagacagcac aggcatgtt tcatgcacag  
1081 gaagcggagg aacatgcaga ggctgtgcag gttctaaaac gaaagtatgt aggtatgctt  
1141 ttaagtgata ttagtagttg tgtggattat aatattagtc cacggtaaa agctatatgc  
1201 atagaaaata acagtaaaac agcaaaacga agactccttg aacttccaga cagcgggtat  
1261 ggcaatactg aagtggaaac gcagcagatg gtacaggtag aggagcaaca aacaacatta  
1321 agttgtaatg gtagtgcgg gacacatagt gaacgagaga atgaaactcc aacacgtaat  
1381 atattgcaag tgttaaaaac tagcaatggt aaagctgcta tgttaggtaa atttaaagaa  
1441 ttatatggtg taagttttat ggaactaatt aggccatttc aaagcaataa aagcacatgt  
1501 actgattggt gtgtagctgc gtttggagtt acaggtacag ttgcagaaggt atttaaaacc  
1561 ctattgcaac catattgttt gtattgcat ttacaaagt tagcatgttc ctggggcatg  
1621 gttatgtaa tgcttgtgag atttaaatgt gcaaaaaata gaataacaat tgaaaaatta  
1681 ttagaaaaat tattgtgtat atctacaaat tgtatgtaa ttcagccacc caaattacgt  
1741 agcacagctg cagcattata ttggtacaga acaggaatgt caaacattag cgatgtatat  
1801 ggtgaaacac cagaatggat agaaagacaa acagtattac agcatagttt taatgacaca  
1861 acatttgatt tgtcccaaat ggtacaatgg gcatatgaca atgatgttat ggatgatagt  
1921 gaaattgcct ataaatagc acaattagct gacagtgata gtaatgcatg tgcattttta  
1981 aaaagtaatt cgcaggcaaa aatagttaaa gattgtggaa caatgtgtag acattataaa  
2041 cgagcagaaa aacgacaaat gtccatggga cagtggatta aaagtagatg tgacaaagtt  
2101 agtgacgaag gtgactggag ggacatagta aagtttttaa gatatcaaca aatagaattt  
2161 gtgtcatttt tatctgcatt aaagctgttt ttaaaggag tgccaaagaa aaactgtatt  
2221 ttaatacatg gtgcacctaa tacaggtaaa tcatatcttg gaatgagcct tattagcttt  
2281 ttacaaggat gtataatc atatgcaaat tcaaaaagtc atttttggtt acaaccactg  
2341 gctgatgcta aaataggcat gttagatgat gctacaacgc catgttggca ttatatagac  
2401 aattacctac gaaatgcact agatggcaac cctgtatcta tagatgtaaa gcataaagct  
2461 ttaatgcagt taaaatgtcc tcttttattg attacatcta atataaatgc aggtaaggat  
2521 gacagatggc catacctaca tagcagactg gtggttttta catttccaaa tccatttcca  
2581 tttgacaaaa acggaaatcc agtatatgaa ttaagtgata aaaactggaa atcctttttc  
2641 tcaaggacgt ggtgcagatt aaatgtgcac gaggaagagg acaaagaaaa cgatggagac  
2701 tctttctcaa cgttttaaatg tgtgtcagga caaaatatta gaacattatg aaaatgatag  
2761 taaacgactt tgtgatcata tagactattg gaaacatatt cgacttgaat gtgtattaat  
2821 gtataaagca agagaaatgg gaatacacag tattaaccac caggtggtgc cagcgttgtc  
2881 agtatcaaa gccaagcct tacaagctat tgaactacaa atgatgttgg aaacattaaa  
2941 taacactgaa tacaaaaatg aggactggac aatgcagcaa acaagtcttg aactgtattt

3001 aactgcacct acaggggtgtt taaaaaaca tggatatact gtagaggtgc aatttgatgg  
3061 tgatgtacac aacaccatgc attatactaa ctggaaatth atatacctat gtatagatgg  
3121 ccaatgtact gttgtggaag ggcaagttaa ttgtaagggc atttattatg tacatgaagg  
3181 acatataaca tattttgtaa attttacaga agaggcaaaa aaatatggga ctggtaaaaa  
3241 atgggaagtg catgcggggtg gtcaggtaat tgthtttctt gaatctgtat ttagcagtga  
3301 cgaaatatcc tttgctggga ttgttacaaa gctaccaaca gccacaaca ccaccacatc  
3361 gaattccaaa acctgcgcct tgggcaccag tgaaggtgtg cggcgggcga cgacgtctac  
3421 taagcgacca agaacagagc cagagcacag aaacaccac caccacaaca agttgttgcg  
3481 aggcgactcc gtggacagtg tcaactgtgg ggthtatcagt gcagctgcat gcacaaacca  
3541 aacaagggct gtcagttgtc ctgcaactac acctataata cacttaaaag gtgatgcaaa  
3601 tatattaata tgtttaagat ataggctgtc aaaatataaa caattgtatg aacaagtgtc  
3661 atctacatgg cattggacat gtacagatgg aaaacataaa aatgctattg taacctaac  
3721 atatataagt acatcacaaa gagacgattt tttaaatact gtaaaaatac ctaacacagt  
3781 atcagtgtca acaggatata tgactattta gcctaattgat tgaactaaat atthctacag  
3841 taagcattgt gctatgctth ttgctthgtc thtgtgtgtc actatthgtg tgtcttgtca  
3901 tacgtccact tgtgtgtgtc gtgtcggat atgcaacact actattatta atthgtgatt  
3961 tatgggttat tgcaacctct ccattacgtt gthtttgtat atatgttgtg thtatatata  
4021 thccattatt tgtaattcat acacatgcat cthththtaag tcaacagtaa cthththact  
4081 tgtgtatact gttgtthtga ttggtattgg tathgttatt ggtattgta thgttataat  
4141 aaactththt actthththt tathattacc atgcggtcca aacgtctac aaaacgcact  
4201 aaacgtgcgt ctgctacaca atthatacaa acatgtaaa cagcaggtac thgtccatca  
4261 gacgttatac ctaaaataga acatactacc atthcagacc aaatathaa gtatggtagt  
4321 atgggtgtth thththgtgg gthgggtatt gggthcggct ctggtactgg gggthcact  
4381 ggatatgtcc ctcttagtac acgtccttct acagtatctg aggcaggat acctattaga  
4441 ccaccagtta gcattgacct tgtaggtccc thggacctc ctatagtaa thctgttgaa  
4501 gaatctgaa thgttgatgt tggthcccct gthcctatac cacacctcc tacaacatct  
4561 gggthtgaca thgttacaac tgcagacaca acacctgca ththtagatgt aacaagtgt  
4621 agcacacatg aaaatcctac ththactgat ccatctgtat tgcagcctcc tacacctga  
4681 gaaacatcag gtcattthact actthcatca tcatctatta gcacacata thtataggaa  
4741 atacctatgg atacaththt tgtthctact aataatgaaa acataacaag tagcacacct  
4801 atthcaggg thgcccgtcc tgcagthth gggthtatata gtaaggctac acaacaagta  
4861 aaagthattg atthcaagth thcttagtgc ccaaacagc thattacata tgaaaacct  
4921 gcctatgaaa ctgthaaatgc tgaagaatct ththacttht ccaatacat gcataatata  
4981 gccctgatc ccgactthct agatathata gcattacata ggctgcctc thctcagct  
5041 aggaacactg thtagatath tagactaggt aataaacaah cththgcgac thgtagtgt  
5101 gthactattg thgcaaggt gcathathat ththgathata gthgathth thctgcaggt  
5161 gaaagtattg aaatgcaacc ththagggcg ththgcaacta thactthctac ththaatgat  
5221 ggctthathg acaththatg agacactgat ththactgtg athacacctg cacacataat  
5281 gththcccct thactgtgt acagthcaca ththctgtgt ththctatgt acctacaaat  
5341 acctgtgtc cactaagth aggtththgac atthccatath ththctgggc thgtgtacct  
5401 athagcatg cacctacaca gthththcca ththcctthg ccctacaa cccacaagt  
5461 ththththth thgtatgggg ththththth ththcacta gthththth ththaaact  
5521 cgacgthaa gtgtatcata ththththth ththctctct thggcgctc gcgaggctac  
5581 ththctacta ccacctgtcc cagththth agthththth acggatgath atgthaacg

5641 aaccaacata tattatcacg caggcagtg c taggctgctt acagtaggcc atccatatta  
5701 ttccatacct aaatctgaca atcctaaaaa aatagttgta ccaaagggtg caggattaca  
5761 atatagggta tttagggttc gtttaccaga tccaaacaaa tttggatttc ctgatacatc  
5821 tttttataat cctgaaactc aacgcttagt ttgggcctgt gttggtttag aggtaggtcg  
5881 cgggcagcca ttaggtgtag gtattagtg tcatccatta ttaaataaat ttgatgacac  
5941 tgaaaactct aatagatatg ccgggtgtcc tggcactgat aatagggaat gtatatcaat  
6001 ggattataaa caaacacaac tgtgtttact tggttgcaa ccacctattg gagagcattg  
6061 gggtaaaggt agtccttgta gtaacaatgc tattaccctt ggtgattgct ctccattaga  
6121 attaaaaaat tcagttatac aagatgggga tatggttgat acaggccttg gagctatgga  
6181 ttttactgct ttacaagaca ctaaaagtaa tgttcctttg gacatttgta attctatttg  
6241 taaatatcca gattatctta aaatggttgc tgagccatat ggcgatacat tattttttta  
6301 tttacgtagg gaacaaatgt ttgtaaggca tttttttaat agatcaggca cggttggtga  
6361 atcgggtcct actgacttat atattaaagg ctccggttca acagctactt tagctaacag  
6421 tacatacttt cctacaccta gcggctccat ggttacttca gatgcacaaa tttttaataa  
6481 accatattgg atgcaacgtg ctcagggaca caataatggt atttggtggg gcaatcagtt  
6541 atttgttact gtggtagata ccacacgtag taccaatatg tctgtttggt ctgcaattgc  
6601 aaacagtgat actacattta aaagtagtaa ttttaagag tatttaagac atgggtgagga  
6661 atttgattta caatttata ttcagttatg caaaataaca ttatctgcag acataatgac  
6721 atatattcac agtatgaatc ctgctatttt ggaagattgg aattttggat tgaccacacc  
6781 tcctcaggt tctttggagg atacctatag gtttgtcacc tcacaggcca ttacatgtca  
6841 aaaaactgcc ccccaaaagc ccaaggaaga tccatttaa gattatgtat tttgggaggt  
6901 taatttaaaa gaaaagtttt ctgcagattt agatcagttt ccaactgggtc gcaaattttt  
6961 attacaggca ggatataggg cacgtcctaa atttaaagca ggtaaacgta gtgcaccctc  
7021 agcatctacc actacaccag caaacgtaa aaaaactaaa aagtaatgga tgtgtatgta  
7081 atacatgtgt ctgtatgtgt atgtgcttgt gctgtattgt atatgtgtgt gtttgtgtgt  
7141 tatatatggt atatgtatgt ttatgtatgc gtgtgtactt gtatatatgt atagtatggt  
7201 atgtgtgtat gtatgctatg tatgttaata aatatgtgta tacctgtgtg tgttgtgtat  
7261 gttgtcctta tataccctt attagtaaca tactattact atttataaa ctattgttcc  
7321 tacttgttcc tacttgttcc tgctcctccc aatagtcatg tacttatttc tgcctataat  
7381 ttaggtgtca cgccatagta aaagttgtac acccgggtccg ttttttgcaa ctaaagetac  
7441 tccattttga ttttatgcag ccattttaaa tccctaaccg ttttcggttg cattgtttaa  
7501 acatgctagt acaactatgc tgatgcagta gttctgcggg ttttggttcc ctgaatacta  
7561 gtttttgcca acattctggc ttgtagtttc ctgcctaaca caccttgcca acatataatc  
7621 cagtccaact ttgcaattat actatgaatc atgtttgttt aaatacaact gtagtccaac  
7681 tatgtgtcat gcacatatat tatattatcc tacacacctt aaactgcttt taggcacata  
7741 tttttagat tatctatatc cttgattgca gtgctggctt ttgcacatgt taaactgcc  
7801 aaggttgtgt catgcattat aaataagttg tatgttactc atataattaa ttgcatatag  
7861 gtattacacc gttttcgggt acagttttac aagcaattgt tctttttata ct

//