

Human papillomavirus type 18 complete sequence

GenBank: AY262282.1

[FASTA Graphics](#)

[Go to:](#)

LOCUS AY262282 7857 bp DNA circular VRL 06-DEC-2005

DEFINITION Human papillomavirus type 18 complete sequence.

ACCESSION AY262282

VERSION AY262282.1

KEYWORDS .

SOURCE human papillomavirus 18

ORGANISM [human papillomavirus 18](#)

Viruses; Monodnaviria; Shotokuvirae; Cossaviricota;
Papovaviricetes; Zurhausenvirales; Papillomaviridae;
Firstpapillomavirinae; Alphapapillomavirus.

REFERENCE 1 (bases 1 to 7857)

AUTHORS Narechania,A., Chen,Z., DeSalle,R. and Burk,R.D.

TITLE Phylogenetic incongruence among oncogenic genital alpha human
papillomaviruses

JOURNAL J. Virol. 79 (24), 15503-15510 (2005)

PUBMED [16306621](#)

REFERENCE 2 (bases 1 to 7857)

AUTHORS Chen,Z. and Burk,R.D.

TITLE The newly modified full genome sequence of HPV18 prototype (Cole,
1987), with E6, E7, E1, E2, E4, E5, L2 and L1 ORFs

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 7857)

AUTHORS Burk,R.D.

TITLE Direct Submission

JOURNAL Submitted (25-MAR-2003) Microbiology and Immunology, Albert
Einstein College of Medicine of Yeshiva University, 1300 Morris
Park Ave., Ullmann 515, Bronx, NY 10461, USA

FEATURES Location/Qualifiers

source 1..7857

/organism="human papillomavirus 18"

/mol_type="genomic DNA"

/db_xref="taxon:[333761](#)"

[gene](#) 105..581

/gene="E6"

[CDS](#) 105..581

/gene="E6"

/codon_start=1

/product="E6 protein"

/protein_id="[AAP20594.1](#)"

/translation="MARFEDPTRRPYKLPDLCTELNTSLQDIEITCVYCKTVLELTEV"

```

FEFAFKDLFVVYRDSIPHAACHKCIDFYSRIREL RHYSDSVYGD TLEKLTNTGLYNLL
IRCLRCQKPLNPAEKLRLHLEKRRFHNIAGHYRGQCHSCCNRARQERLQRRRETQV"

gene      590..907
          /gene="E7"

CDS       590..907
          /gene="E7"
          /codon_start=1
          /product="E7 protein"
          /protein_id="AAP20595.1"
          /translation="MHGPKATLQDIVLHLEPQNEIPVDLLCHEQLSDSEEENDEIDGV
NHQHLPARRAEPQRHTMLCMCKCEARIELVVESSADDLRAFQQLFLNTLSFVCPWCA
SQQ"

gene      914..2887
          /gene="E1"

CDS       914..2887
          /gene="E1"
          /codon_start=1
          /product="E1 protein"
          /protein_id="AAP20596.1"
          /translation="MADPEGTDGEGTGCNGWFYVQAIVDKKTGDVISDDEDENATDTG
SDMVD FIDTQGTFC EQAELETAQALFHAQEVHND AQVLHVLKRKFAGGSTENSPLGER
LEVDT ELS PRLQEISLNSGQKKAKRRLFTISDSGYGCSEVEATQIQVTN GEHGGNVC
SGGST EAIDNGGTEGNSSVDGTS DNSNIENVNPQCTIAQLKDLLKVNNKQGAMLAVF
KDTYGLSFTDLVRNFKSDKTTCTDWVTAIFGVNPTIAEGFKTLIQPFILYAHIQCLDC
KWGVLILALLRYKCGKSRLTVAKGLSTLLHVPETCMLIQPPKLRSSVAALYWYRTGIS
NISEVMGDTPEWIRQRTIIQH GIDDSNFDLSEMVQWAFDNELTDES DMAFEYALLADS
NSNAA AFLKSNCQAKYLKDCATMCKHYRRAQKRQMNMSQWIRFRCSKIDEGGDWRPIV
QFLRYQQIEFITFLGALKSFLKGT PKKNCLVFCGPANTGKSYFGMSFIHF IQGAVISF
VNSTSHFWLEPLTDTKVAMLD DATTCWTFYFD TYMRNALDGNPISIDRKHKPLIQ LKC
PPILLTTNIHPAKDNRWPYLES RITVFEFPNAFPFDKNGNPVYEINDKNWKCF FERTW
SRDLH EEEEEADTEGNPF GTFKCVAGQNRPL"

gene      2817..3914
          /gene="E2"

CDS       2817..3914
          /gene="E2"
          /codon_start=1
          /product="E2 protein"
          /protein_id="AAP20597.1"
          /translation="MQTPKETLSERLSALQDKIIDHYENDSKDIDSQIQYWQLIRWEN
AIFFAAREHGIQTLNHQVVPAYNISKSKAHKAIELQMALQGLAQSA YKTEDWTLQDTC
EELWNTEPTHCFKKGQTVQVYFDG NKDNCMTYVAWDSVYYMTDAGTW DKTATCVSHR
GLYYVKEGYNTFYIEFKSECEKYGNTGTWEVHFGNNVIDCND SMCSTSDDTVSATQLV
KQLQH TSPYSSTVSVGTAKTYGQ TSAATRPGHCGLA EKQHCGPVNPLLGAATPTGNN
KRRKLC SGNTTPIIHLKGDRNSL KCLRRLRKHSDHYRDISSTWHWTGAGNEKTGILT

```

VTYHSETQRTKFLNTVAIPDSVQILVGYMTM"
gene 3418..3684
/gene="E4"
CDS 3418..3684
/gene="E4"
/codon_start=1
/product="E4 protein"
/protein_id="AAP20598.1"
/translation="MTLCAVPVTTRYPLLSLLNSYSTPPHRIPAPCPWAPQRPTARRR
LLHDLDTVDSRRSSIVDLSTHFSVQLHLQATTKDGNVSVVTLRL"

gene 3936..4157
/gene="E5"
CDS 3936..4157
/gene="E5"
/codon_start=1
/product="E5 protein"
/protein_id="AAP20599.1"
/translation="MLSLIFLFCFCVCMYVCCHVPLLPSVCMCAYAWVLVFVYIVVIT
SPATAFTVYVFCFLLPMLLLHIHAILSQ"

gene 4244..5632
/gene="L2"
CDS 4244..5632
/gene="L2"
/codon_start=1
/product="L2 protein"
/protein_id="AAP20600.1"
/translation="MVSHRAARRKRASVTDLYKTCKQSGTCPPDVVPKVEGTTLADKI
LQWSSLGIFLGGLGIGTGSGTGGRGTGYIPLGGRSNTVVDVGPTRPVVPVIEPVGPTDPS
IVTLIEDSSSVVTSGAPRPTFTGTSGFDITSAGTTTPAVLDITPSSTSVSISTTNFTNP
AFSDPSIIEVPQTGEVAGNVFVGTPTSNGTHYEEIPLQTFASSGTGEEPISSTPLPTV
RRVAGPRLYSRAYQQVSVANPEFLTRPSSLITYDNPAFEPVDTTLTDFDRSDVPDSDF
MDIIRLHRPALTSRRGTVRF SRLGQRATMFTRSGTQIGARVHFYHDISPIAPSPEYIE
LQPLVSATEDNDLFDIYADDMDPAVPVPSRSTTSFAFFKYSPTISSASSYSNVTVPLT
SSWDVPVYTGPDITLPSTTSVWPIVSPTAPASTQYIGIHGTHYYLWPLYFIPKKRKR
VPYFFADGFVAA"

gene 5430..7136

/gene="L1"

CDS 5430..7136
/gene="L1"
/codon_start=1
/product="L1 protein"
/protein_id="AAP20601.1"
/translation="MCLYTRVLILHYHLLPLYGPLYHPRPLPLHSILVYVMVHIIICGH
YIILFLRNVNVPFI FLQMALWRPSDNTVYLPPPSVARVVNTDDYVTRTSIFYHAGSSR

LLTVGNPYFRVPAGGGNKQDIPKVSAYQYRVFRVQLPDPNKFLPDTSIYNPETQRLV
WACAGVEIGRGQPLGVGLSGHPFYNKLDDTESSHAATSNVSEDVRDNVSVDYKQTQLC
ILGCAPAIGEHWAKGTACKSRPLSQGDCPPLELKNTVLEDGDMVDTGYGAMDFSTLQD
TKCEVPLDICQSICKYPDYLQMSADPYGDSMFFCLRREQLFARHFWNRAGTMGDTVPO
SLYIKGTGMRASPGSCVYSPSPSGSIVTSDSQLFNKPYWLHKAQGHNNGVCWHNQLFV
TVVDTRSTNLTICASTQSPVPGQYDATKFKQYSRHVEEYDLQFIFQLCTITLTADVM
SYIHSMNSSILEDWNFGVPPPTTSLVDITYRFVQSVAITCQKDAAPAENKDPYDKLKF
WNVDLKEKFSLDLDQYPLGRKFLVQAGLRKPTIGPRKRSAPSATTSSKPAKRVRVRA
RK"

ORIGIN

1 attaatactt ttaacaattg tagtatataa aaaaggaggt aaccgaaaac ggtcgggacc
61 gaaaacggtg tatataaaag atgtgagaaa cacaccacaa tactatggcg cgctttgagg
121 atccaacacg gcgaccctac aagctacctg atctgtgcac ggaactgaac acttcactgc
181 aagacataga aataacctgt gtatatgtca agacagtatt ggaacttaca gaggtatttg
241 aatttgcatt taaagattta tttgtggtgt atagagacag tataccgcat gctgcatgcc
301 ataatgtat agatttttat tctagaatta gagaattaag acattattca gactctgtgt
361 atggagacac attggaaaaa ctaactaaca ctgggttata caatttatta ataagggtcc
421 tgcggtgcc aaaaacctg aatccagcag aaaaacttag acaccttaac gaaaaacgac
481 gatttcacaa catagctggg cactatagag gccagtgcc ttcgtgctgc aaccgagcac
541 gacaggaacg actccaacga cgcagagaaa cacaagtata atattaagta tgcattggacc
601 taaggcaaca ttgcaagaca ttgtattgca tttagagccc caaatgaaa ttccggttga
661 ccttctatgt cagcagcaat taagcgactc agaggaagaa aacgatgaaa tagatggagt
721 taatcatcaa catttaccag cccgacgagc cgaaccacaa cgtcacacaa tgttgtgtat
781 gtgttgtaag tgtgaagcca gaattgagct agtagtagaa agctcagcag acgaccttcg
841 agcattccag cagctgtttc tgaacaccct gtcctttgtg tgtccgtggt gtgcatcca
901 gcagtaagca acaatggctg atccagaagg tacagacggg gagggcacgg gttgtaacgg
961 ctgggtttat gtacaagcta ttgtagacaa aaaaacagga gatgtaatat cagatgacga
1021 ggacgaaaat gcaacagaca cagggtcgga tatggtagat tttattgata cacaaggaac
1081 attttgtgaa caggcagagc tagagacagc acaggcattg ttccatgcgc aggaggtcca
1141 caatgatgca caagtgttg atgttttaaa acgaaagttt gcaggaggca gcacagaaaa
1201 cagtccatta ggggagcggc tggaggtgga tacagagtta agtccacggt tacaagaaat
1261 atctttaaat agtgggcaga aaaaggcaaa aaggcggctg tttacaatat cagatagtgg
1321 ctatggctgt tctgaagtgg aagcaacaca gattcaggta actacaaatg gcgaacatgg
1381 cggcaatgta tgtagtggcg gcagtacgga ggctatagac aacgggggca cagagggcaa
1441 caacagcagt gtagacggt caagtgacaa tagcaatata gaaaatgtaa atccacaatg
1501 taccatagca caattaaaag acttggttaa agtaaacaat aaacaaggag ctatgttagc
1561 agtatttaaa gacacatatg ggctatcatt tacagattta gtagaaatt taaaagtga
1621 taaaaccacg tgtacagatt gggttacagc tatatttgga gtaaacccaa caatagcaga
1681 aggatttaaa acactaatac agccatttat attatatgcc catattcaat gtctagactg
1741 taaatgggga gtattaatat tagccctgtt gcgttacaaa tgtggtaaga gtagactaac
1801 agttgctaaa ggtttaagta cgttggtaca cgtacctgaa acttgtagt taattcaacc
1861 accaaaattg cgaagtagtg ttgcagcact atattggtat agaacaggaa tatcaaatat
1921 tagtgaagta atgggagaca cacctgagtg gatacaaaga ctactatta tacaacatgg
1981 aatagatgat agcaattttg atttgtcaga aatggtacaa tgggcatttg ataagtagct

2041 gacagatgaa agcgatatgg catttgaata tgccttatta gcagacagca acagcaatgc
2101 agctgccttt ttaaaaagca attgccaaagc taaatattta aaagattgtg ccacaatgtg
2161 caaacattat aggcgagccc aaaaacgaca aatgaatatg tcacagtggg tacgatttag
2221 atgttcacaaa atagatgaag ggggagattg gagaccaata gtgcaattcc tgcgatacca
2281 acaaatagag tttataacat ttttaggagc cttaaaatca tttttaaaag gaacccccaa
2341 aaaaaattgt ttagtatttt gtggaccagc aaatacagga aaatcatatt ttggaatgag
2401 ttttatacac tttatacaag gagcagtaat atcatttgtg aattccacta gtcatttttg
2461 gttggaaccg ttaacagata ctaagggtggc catgttagat gatgcaacga ccacgtgttg
2521 gacatacttt gataacctata tgagaaatgc gttagatggc aatccaataa gtattgatag
2581 aaagcacaaa ccattaatac aactaaaatg tcctccaata ctactaacca caaatataca
2641 tccagcaaa gataatagat ggccatattt agaaagtaga ataacagtat ttgaatttcc
2701 aaatgcattt ccatttgata aaaatggcaa tccagtatat gaaataaatg acaaaaattg
2761 gaaatgtttt tttgaaagga catggtccag attagatttg cacgaggaag aggaagatgc
2821 agacaccgaa ggaaaccctt tcggaacgtt taagtgcgtt gcaggacaaa atcatagacc
2881 actatgaaaa tgacagtaaa gacatagaca gccaaataca gtattggcaa ctaatacgtt
2941 gggaaaatgc aatattcttt gcagcaaggg aacatggcat acagacatta aaccaccagg
3001 tgggtgccagc ctataacatt tcaaaaagta agcacataa agctattgaa ctgcaaatgg
3061 ccctacaagg ccttgacaaa agtgcataca aaaccgagga ttggacactg caagacacat
3121 gcgaggaact atggaataca gaacctactc actgctttta aaaaggtggc caaacagtac
3181 aagtatattt tgatggcaac aaagacaatt gtatgaccta tgtagcatgg gacagtgtgt
3241 attatatgac tgatgcagga acatgggaca aaacggctac ctgtgtaagt cacaggggat
3301 tgtattatgt aaaggaaggg tacaacacgt tttatataga atttaaaagt gaatgtgaaa
3361 aatatgggaa cacaggtacg tgggaagtac attttgggaa taatgtaatt gattgtaatg
3421 actctatgtg cagtaccagt gacgacacgg tatccgctac tcagcttggt aaacagctac
3481 agcacacccc ctcaccgtat tccagcacgg tgtccgtggg caccgcaaag acctacggcc
3541 agacgtcggc tgctacacga cctggacact gtggactcgc ggagaagcag cattgtggac
3601 ctgtcaaccc acttctcgtt gcagctacac ctacaggcaa caacaaaaga cggaaactct
3661 gtagtggtaa cactacgcct ataatacatt taaaagggtga cagaaacagt ttaaaatgtt
3721 tacggtacag attgcgaaaa catagcgacc actatagaga tatatcatcc acctggcatt
3781 ggacaggtgc aggcaatgaa aaaacaggaa tactgactgt aacataccat agtgaaacac
3841 aaagaacaaa atttttaaat actgttgcaa ttccagatag tgtacaaata ttggtgggat
3901 acatgacaat gtaatacata tgctgtagta ccaatatgtt atcacttatt tttttatttt
3961 gcttttgtgt atgcatgtat gtgtgctgcc atgtcccgt tttgccatct gtctgtatgt
4021 gtgcgtatgc atgggtattg gtatttgtgt atattgtggt aataacgtcc cctgccacag
4081 cattcacagt atatgtattt tgttttttat tgcccatgtt actattgcat atacatgcta
4141 tattgtcttt acagtaattg tataggttgt tttatacagt gtattgtaca ttgtatattt
4201 tgttttatac cttttatgct ttttgtattt ttgtaataaa agtatgggat ccaccgtgc
4261 cgcacgacgc aaacgggctt cggtaactga cttatataaa acatgtaaac aatctggtac
4321 atgtccacct gatgttggtc ctaagggtgga gggcaccacg ttagcagata aaatattgca
4381 atgggtcaagc cttggtatat ttttgggtgg acttggcata ggtactggca gtggtacag
4441 gggtcgtaca gggtaacatt cattgggtgg gcgttccaat acagtgggtg atgttgggtcc
4501 tacacgtccc ccagtgggta ttgaacctgt gggccccaca gacctatcta ttgttacatt
4561 aatagaggac tccagtgtgg ttacatcagg tgcacctagg cctacgttta ctggcacgtc
4621 tgggtttgat ataacatctg cgggtacaac tacacctgcg gttttggata tcacaccttc

4681 gtctacctct gtgtctatct ccacaaccaa ttttaccat cctgcatttt ctgatccgtc
4741 cattattgaa gttccacaaa ctggggaggt ggcaggtaat gtatttgttg gtacccttac
4801 atctggaaca catgggtatg aggaaatacc tttacaaaaca tttgcttctt ctgggtacggg
4861 ggaggaaccc attagtagta ccccatggcc tactgtgcgg cgtgtagcag gtccccgcct
4921 ttacagtagg gcctaccaac aagtgtcagt ggctaaccct gagtttctta cacgtccatc
4981 ctctttaatt acatatgaca acccggcctt tgagcctgtg gacactacat taacatttga
5041 tctctgtagt gatgttcctg attcagattt tatggatatt atccgtctac ataggcctgc
5101 tttaacatcc aggcgtagga ctgttcgctt tagtagatta ggtcaacggg caactatgtt
5161 taccgcagc ggtacacaaa taggtgctag ggttcacttt tatcatgata taagtcctat
5221 tgcaccttc ccagaatata ttgaactgca gccttttagta tctgccacgg aggacaatga
5281 cttgtttgat atatatgcag atgacatgga cctgcagtg cctgtacat cgcgttctac
5341 tacctccttt gcatttttta aatattcgcc cactatatct tctgcctctt cctatagtaa
5401 tgtaacggtc cctttaacct cctctggga tgtgcctgta tacacgggtc ctgatattac
5461 attaccatct actacctctg tatggcccat tgtatcacc acggccctg cctctacaca
5521 gtatatttgt atacatggta cacattatta tttgtggcca ttatattatt ttattcctaa
5581 gaaacgtaaa cgtgttccct atttttttgc agatggcttt gtggcgccct agtgacaata
5641 ccgtatatct tccacctcct tctgtggcaa gagttgtaaa taccgatgat tatgtgactc
5701 gcacaagcat attttatcat gctggcagct ctagattatt aactgttggt aatccatatt
5761 ttagggttcc tgcaggttgt ggcaataagc aggatattcc taaggtttct gcataccaat
5821 atagagtatt tagggtgcag ttacctgacc caaataaatt tggtttacct gatactagta
5881 tttataatcc tgaaacacaa cgttttagtgt gggcctgtgc tggagtggaa attggccgtg
5941 gtcagccttt aggtgttggc cttagtgggc atccatttta taataaatta gatgacactg
6001 aaagttccca tgccgccacg tctaattgtt ctgaggacgt tagggacaat gtgtctgtag
6061 attataagca gacacagtta tgtatttttg gctgtgcccc tgctattggg gaacactggg
6121 ctaaaggcac tgcttgtaaa tcgcgtcctt tatcacagg cgattgcccc cctttagaac
6181 ttaaaaacac agttttggaa gatggtgata tggtagatac tggatatggt gccatggact
6241 ttagtacatt gcaagatact aaatgtgagg taccattgga tatttgtcag tctatttcta
6301 aatatcctga ttattttaca atgtctgcag atccttatgg ggattccatg tttttttgct
6361 tacggcgtga gcagcttttt gctaggcatt tttggaatag agcaggtaact atgggtgaca
6421 ctgtgcctca atccttatat attaaaggca caggatgctg tgcttcacct ggcagctgtg
6481 tgtattctcc ctctccaagt ggctctattg ttacctctga ctcccagttg ttttaataaac
6541 catattggtt acataaggca cagggtcata acaatggtgt ttgctggcat aatcaattat
6601 ttgttactgt ggtagatacc actgcagta ccaatttaac aatatgtgct tctacacagt
6661 ctctgttacc tgggcaatat gatgctacca aatttaagca gtatagcaga catgttgagg
6721 aatatgattt gcagtttatt tttcagttgt gtactattac ttttaactgca gatgttatgt
6781 cctatatcca tagtatgaat agcagtattt tagaggattg gaacttttgt gttccccccc
6841 cgccaactac tagtttggtg gatacatatc gttttgtaca atctgttgct attacctgtc
6901 aaaaggatgc tgcaccggt gaaaataagg atccctatga taagttaaag ttttggaatg
6961 tggatttaaa ggaaaagttt tcttttagact tagatcaata tccccttggc cgtaaatttt
7021 tggttcaggc tggattgctg cgcaagccca ccataggccc tcgcaaactg tctgctccat
7081 ctgccactac gtcttctaaa cctgccaaagc gtgtgcgtgt acgtgccagg aagtaatatg
7141 tgtgtgtgta tatatatata catctattgt tgtgtttgta tgcctgtgt ttgtgtttgt
7201 tgtatgattg cattgtatgg tatgtatggt gtgtgttgta gtgtgtatgt tactatattt
7261 gttggtatgt ggcattaaat aaaatatggt ttgtggttct gtgtgttatg tggttgcgcc

7321 ctagtgagta acaactgtat ttgtgtttgt ggtatgggtg ttgcttggtg ggctatatat
7381 tgtcctgtat ttcaagttat aaaactgcac accttacagc atccatttta tcctacaatc
7441 ctccattttg ctgtgcaacc gatttcggtt gcctttggct tatgtctgtg gttttctgca
7501 caatacagta cgctggcact attgcaaact ttaatctttt gggcactgct cctacataatt
7561 ttgaacaatt ggcgcgccctc tttggcgcat ataaggcgca cctgggatta gtcattttcc
7621 tgtccagggtg cgctacaaca attgcttgca taactatatc cactccctaa gtaataaaac
7681 tgcttttagg cacatatttt agtttgtttt tacttaagct aattgcatac ttggcttgta
7741 caactacttt catgtccaac attctgtcta cccttaacat gaactataat atgactaagc
7801 tgtgcataca tagtttatgc aaccgaaata ggttgggcag cacatactat acttttc

//