

Human papillomavirus type 52 genomic DNA

GenBank: X74481.1

[FASTA Graphics](#)

[Go to:](#)

LOCUS X74481 7942 bp DNA linear VRL 18-APR-2005
DEFINITION Human papillomavirus type 52 genomic DNA.
ACCESSION X74481
VERSION X74481.1
KEYWORDS E1 gene; E2 gene; E4 gene; E6 gene; E7 gene; early protein; L1 gene; L2 gene; late protein.
SOURCE human papillomavirus 52
ORGANISM [human papillomavirus 52](#)
Viruses; Monodnaviria; Shotokuvirae; Cossaviricota; Papovaviricetes; Zurhausenvirales; Papillomaviridae; Firstpapillomavirinae; Alphapapillomavirus.
REFERENCE 1 (bases 1 to 7942)
AUTHORS Delius,H. and Hofmann,B.
TITLE Primer-directed sequencing of human papillomavirus types
JOURNAL Curr. Top. Microbiol. Immunol. 186, 13-31 (1994)
PUBMED [8205838](#)
REFERENCE 2 (bases 1 to 7942)
AUTHORS Delius,H.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1993) H. Delius, Deutsches Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W 6900 Heidelberg, FRG
FEATURES Location/Qualifiers
source 1..7942
/organism="human papillomavirus 52"
/mol_type="genomic DNA"
/db_xref="taxon:[10618](#)"
/clone="insert in EcoRI site of pUC19"
[gene](#) 102..548
/gene="E6"
[CDS](#) 102..548
/gene="E6"
/codon_start=1
/product="early protein"
/protein_id="[CAA52585.1](#)"
/db_xref="GOA:[P36814](#)"
/db_xref="InterPro:[IPR001334](#)"
/db_xref="UniProtKB/Swiss-Prot:[P36814](#)"
/translation="MFEDPATRPRTLHELCEVLEESVHEIRLQCVQCKKELQRREVYK
FLFTDLRIVYRDNNPYGVCIMCLRFLSKISEYRHYQYSLYGKTLEERVKKPLSEITIR

	CIICQTPLCPPEEKERHVNANKRFHNIMGRWTGRCSECWRPRPVTQV"
gene	553..852
	/gene="E7"
CDS	553..852
	/gene="E7"
	/codon_start=1
	/product="early protein"
	/protein_id="CAA52586.1"
	/db_xref="GOA:P36831"
	/db_xref="InterPro:IPR000148"
	/db_xref="UniProtKB/Swiss-Prot:P36831"
	/translation="MRGDKATIKDYILDLPETTDLHCYEQLGDSSEEDTDGVDRPD
	GQAEQATSNNYIVTYCHSCDSTLRLCIHSTATDLRTLQQMLLGTQVVCPCARL"
gene	864..2807
	/gene="E1"
CDS	864..2807
	/gene="E1"
	/codon_start=1
	/product="early protein"
	/protein_id="CAA52587.1"
	/db_xref="GOA:P36730"
	/db_xref="InterPro:IPR001177"
	/db_xref="InterPro:IPR014000"
	/db_xref="InterPro:IPR014015"
	/db_xref="InterPro:IPR016393"
	/db_xref="UniProtKB/Swiss-Prot:P36730"
	/translation="MEDPEGTEGEREGCTGWFEVEAIEKQTGDNISEDEDENAYDSG
	TDLIDFIDDSNINNEQAEHEAARALFNAQEGEDDLHAVSAVKRKFTSSPESAGQDGVE
	KHGSPRAKHICVNTECVLPKRKPCHVEDSGYGNSEVEAQQMADQVDGQNGDWQSNSSQ
	SSGVGASNSDVSCTSIEDNEENSRTLKSIQNIMCENSIKTTVLFKFKETYGVSFMEL
	VRPFKSNRSSCTDWCIIIGMVTPTVAEGLKVLIQPYSIYAHQLCLTCDRGVLILLIR
	FKCGKNRLTVSKLMSQLLNIPETHMVEPPKLRSATCALYWYRTGLSNISEVYGTTP
	WIEQQTVLQHSFDNSIFDFGEMVQWAYDHDITDDSDIAYKYAQLADVNSNAAFLKSN
	SQAKIVKDCATMCRHYKRAERKHMNIGQWIQYRCRIDDGGDWRPIVRFLRYQDIEFT
	AFLDAFKKFLKGIPKKNCLVLYGPANTGKSYFGMSLIRFLSGCVISYVNSKSHFWLQP
	LTDKAVGMIDDVTPICWTYIDDYMRNALDGNDISVDVKHRALVQIKCPPLILTTNTNA
	GTDPRWPYLHSRLVVFHFKNPFPPFDENGNIYEINNENWKSFFSRTWCKLDLIQEEDK
	ENDGVDTGTGFKCSAGKNTRSIRS"
gene	2743..3849
	/gene="E2"
CDS	2743..3849
	/gene="E2"
	/codon_start=1
	/product="early protein"

	/protein_id="CAA52588.1" /db_xref="GOA:P36796" /db_xref="InterPro:IPR000427" /db_xref="InterPro:IPR001866" /db_xref="InterPro:IPR009021" /db_xref="InterPro:IPR012677" /db_xref="UniProtKB/Swiss-Prot:P36796" /translation="MESIPARLNAVQEKILDLYEADSNDLNAQIEHWKLTRMECVLFYKAKELGITHIGHQVPPMAVSKAKACQAIELQLALEALNKTOYSTDGWTLQQTSLEMWRAEPQKYFKKHGYTITVQYDNDKNNMDYTNWKEIYLLGECECTIVEGQVDYYGLYYWCDGEKIYFVKFSNDKQYCVTGVWEVHVGGQVIVCPASVSSNEVSTTETAVHLCETSKTSAVSVGAKDTHLQPPQKRRRPDVTDSRNTKYPNNLLRGQQSVDSTTRGLVTATECTNKGRVAHTTCTAPIIHLKGDPNSLKCLRVRVKTHKSLYVQISSTWHWTSNECTNNKLGIVTITYSDETQRQQFLKTVKIPNTVQVIQGVMSL"
gene	4262..5662
	/gene="L2"
CDS	4262..5662
	/gene="L2"
	/codon_start=1
	/product="late protein"
	/protein_id="CAA52589.1"
	/db_xref="GOA:P36763"
	/db_xref="InterPro:IPR000784"
	/db_xref="UniProtKB/Swiss-Prot:P36763"
	/translation="MRYRRSTRHKRASATQLYQTCASGTCPPDVIPKVEGTTIADQLKYGSLGVFFGGLGIGTGAGSGGRAGYVPLSTRPPTSSITTSTIRPPVTVEPIGPLEPSIVSMIEETTFFIESGAPAPSIPSATGFDVTTSANNTPAIINVTSIGESSVQSVSTHLNPTFTEPSIIQPPAPAEASGHVLFSSPTISTHTYEEIPMDTFVTSTDSSSVTSSTPIPGSRPTTRLGLYSRATQQVKVVDPAFMSSPQKLVTYNNPVFEGVDTDETIIIFDRSQLLPAPDPDFLDIIALHRPALTSRRGTVRFSRLGNKATLRTRSGKQIGARVHYHDISPIQPAEVQEDIELQPLLPQSVSPYTINDGLYDVYADSLQQPTFHL PSTLSTHNNFTVPINSGIDFVYQPTMSIESGPDIPLPSPHTHTPFVPIAPTAPSTSIIVDGTDFILHPSYFLLRRRRKRFPYFFTDVRVAA"
gene	5565..7154
	/gene="L1"
CDS	5565..7154
	/gene="L1"
	/codon_start=1
	/product="late protein"
	/protein_id="CAA52590.1"
	/db_xref="GOA:Q05138"
	/db_xref="InterPro:IPR002210"
	/db_xref="InterPro:IPR011222"
	/db_xref="UniProtKB/Swiss-Prot:Q05138"

/translation="MVQILFYILVIFYVAGVNVFHIFLQMSVWRPSEATVYLPPVPV
SKVVSTDEYVSRTSIYYYAGSSRLTLVGHYPYFSIKNTSSGNGKKVLVPKVSGLQYRVF
RIKLDPDNKFGFDPDTSFYNPETQRLVWACTGLEIGRGQPLGVGISGHPLLNKFDDTET
SNKYAGKPGIDNRECLSM DYKQTQLCILGCKPPIGEHWGKGTPCNNNSGNPGDCPPLQ
LINSVIQDGMVD TGFGCMDFNTLQASKSDVPIDICSSVCKYPDYLQMASEPYGDSLF
FFLRREQMFVRHFFNRAGTLGDPVPGDLYIQGSNSGNTATVQSSAFFPTPSGSMVTSE
SQLFNKPYWLQRAQGHNNGICWGNQLFVTVD TTRSTNMTLCAEVKKESTYKNENFKE
YLRHGEEDLQFIFQLCKITLTADVM TYIHKMDATILEDWQFGLTPPPSASLED TYRF
VTSTAITCQKNTPPKGKEDPLKDYMFWEVDLKEKFSADLDQFPLGRKFLLQAGLQARP
KLKR PASSAPRTSTKKKKVKR"

ORIGIN

1 taaattataa tcttatacta gtaaaaaata ggggtgtaacc gaaaacggtc agaccgaaac
61 cgggtgtatat atatagaaca cagtgtagct aacgcacggc catgtttgag gatccagcaa
121 caccgaccccg gaccctgcac gaattgtgtg aggtgctgga agaatcggtg catgaaataa
181 ggctgcagtg tgtgcagtc aaaaaagagc tacaacgaag agaggtatac aagtttctat
241 ttacagattt acgaatagta tatagagaca ataatccata tggcgtgtgt attatgtgcc
301 tacgcttttt atctaagata agtgaatata ggcattatca atattcactg tatgggaaaa
361 cattagaaga gagggtaaaa aaaccattaa gtgaaataac tattagatgt ataatttgtc
421 aaacgccatt atgtcctgaa gaaaaagaaa gacatgttaa tgcaaacaag cgatttcata
481 atattatggg tcgttgga gggcgctgtt cagagtgttg gagacccga cctgtgaccc
541 aagtgtaacg tcatgcgtgg agacaaagca actataaaaag attatatatt agatctgcaa
601 cctgaaacaa ctgacctaca ctgctatgag caattagtg acagctcaga tgaggaggat
661 acagatggtg tggaccggcc agatggacaa gcagaacaag ccacaagcaa ttactacatt
721 gtgacataatt gtcacagttg tgatagcaca ctacggctat gcattcatag cactgcgacg
781 gaccttcgta ctctacagca aatgctgttg ggcacattac aagttgtgtg ccccggtgt
841 gcacggctat aaacaacct gcaatggagg accctgaagg tacagaggc gaaagggagg
901 gatgtacagg ctggtttgaa gtagaggcaa taatagaaaa acaaacagga gataacattt
961 cagaggacga ggatgaaaat gcatatgata gtggaacaga tctaatagat tttatagatg
1021 attcaaatat aaataatgaa caggcagaac atgaggcagc cggggcattg tttaatgcac
1081 aggaagggga ggatgattta catgctgtgt ctgcagtaaa acgaaagttt acaagcagtc
1141 cggaaagtgc tgggcaagat ggtgtagaaa aacatggtag tccgctgca aaacacattt
1201 gtgtaaatac agagtgtgtt ttacaaaaac gcaaaccatg tcacgtagaa gacagcggct
1261 atggcaatag tgaagtggaa gcgcagcaga tggcagacca ggtagacggg caaatggcg
1321 actggcaaag taacagtagt caatcaagtg ggggtggggc tagtaattca gatgtaagtt
1381 gtactagtat agaggacaat gaggaaaata gtaatagaac gctaaaaagc atacaaaata
1441 ttatgtgcga aaatagcata aaaacaactg tattatttaa atttaaagaa acatatggtg
1501 ttagctttat ggaattagta agaccattta aaagtaatag aagtagttgt acagattggg
1561 gtattatagg aatgggagta acaccatcag ttgcagaagg attaaaagta ttaatacagc
1621 cctatagcat atatgccc attgcaatgtt taacatgtga cagaggcgtg cttatactgc
1681 tgctaattag gtttaaatgt ggaaaaaaca gattaacagt gtccaaacta atgtcacagc
1741 tgttaaatat accagaaaca catatggtaa tagaaccacc aaaattacga agtgctacct
1801 gtgcattata ttggtataga acaggtttgt ctaatattag tgaggatat ggtaccaccc
1861 cagaatggat agaacaacaa acagtattac agcatagctt tgacaatagc atattcgatt
1921 ttggagaaat ggtgcaatgg gcatatgatc atgatataac agatgatagt gacatagcat

1981 ataaatatgc acagtttagca gatgtaaata gcaatgctgc agcattccta aaaagcaatt
2041 cgcaagcaaa aatagtaaaag gactgtgcaa ccatgtgtag acattataaa cgggcagaaa
2101 gaaaacatat gaatatggga caatggatac agtatagatg tgatagaata gatgatggtg
2161 gagattggag gcctatagta agatTTTTTaa gatatcaaga catagaattt acagcctttt
2221 tagacgcatt taaaaaattt ttaaaaggta tacctaaaaa aaattgttta gtattatatg
2281 gacctgcaaa cacaggaaaa tcatattttg gaatgagttt aattaggttc ttaagtggat
2341 gtgtaatatc ctatgtaaac tcaaaaagcc atttttggct acaaccatta acagatgcaa
2401 aagtgggtat gatagatgat gtaacaccta tatgttggac atatatagat gattatatga
2461 gaaatgcact ggatggaaat gatatatcag tagatgtaaa gcatagagcc ttagtacaaa
2521 taaaatgccc accattaatt ttaacaacaa atacaaatgc aggaacagat cctagggtggc
2581 catattttaca tagtagattg gttgtgtttc atttcaaaaa ccattttcca tttgatgaaa
2641 atggcaatcc tatatatgaa attacaacg aaaattggaa atcctttttc tcaaggacgt
2701 ggtgcaaatt agattttaata caggaagagg acaaggaaaa cgatggagtc gataccggca
2761 cgttttaaat cagtgcagga aaaaatacta gatctatacg aagctgatag taatgaccta
2821 aacgcacaaa ttgaacattg gaaattgact cgaatggaat gtgttttgtt ttacaaagca
2881 aaggaactgg gaataactca tataggccac caggtgggtc caccaatggc agtgtctaag
2941 gcaaaggcct gccaaactat tgaactacaa ttggcattgg aggcattaaa caaaacacaa
3001 tatagcacag atggatggac attacaacaa acaagtctag aaatgtggcg tgcagaacca
3061 caaaaatact ttaaaaaaca tgggtataca ataacagtgc aatacgataa tgataaaaac
3121 aatactatgg attatacaaa ctggaaggaa atttatttac ttgggtgagtg tgaatgtaca
3181 attgtagaag gacaagtaga ttactatggg ttatattatt ggtgtgatgg agaaaaata
3241 tattttgtaa aatttagtaa cgatgcaaag caatattgtg taacaggagt atgggaagta
3301 catgtgggtg gtcaggtaat tgtttgtcct gcatctgtat ctagtaacga agtatccact
3361 actgaaactg ctgtccacct atgcaccgaa acctccaaga cctccgcagt gtccgtgggt
3421 gccaaagaca cacacctaca accaccacag aaacgacgac gaccagacgt cacagactcc
3481 agaaacacca agtaccceaa caaccttttg cggggacaac aatccgtgga cagtactaca
3541 cggggactcg tcaactgcaac tgagtgcaca aacaaaggac gggttgcaca tacaacttgt
3601 actgcaccta taatacacct aaaagggtgat cctaatagtt taaaatgttt aagatatag
3661 gtaaaaacac ataaaagtgt gtatgttcaa atttcatcta cctggcattg gaccagtaat
3721 gaatgtacaa ataataaact aggtattgta acaataacgt acagtgatga aacacaacgt
3781 caacaatttt taaaaactgt taaaatacca aatactgtgc aagttataca aggtgtcatg
3841 tcattgtgat atttgtacat atgtatatat gtatatgtgt atggtaaaca cccaacacaa
3901 gccaatattg ctgctattgt gtatatataa caatgttagg attatttcta ttttgtttta
3961 ttttgcttat ggtgttttgt gcagtgtcta ggccgctctt gctatctata tcgggtgatg
4021 cgcagggtgt ggtgctgggt cttttgctat gggatatctat tgggtcacca tttaaagtgt
4081 tttttttgta cctactgttt ttatattttc caatgttttg tattcactgt catgcacagt
4141 atttggcaca actgcaataa ctgtacatgt agattggcta catgcatata tgcaaaatat
4201 actttttcac tttttagatt tgtctaataa atacttttat attttttaat agcttgtcgc
4261 aatgagatac agacggtcta cacggcacia acgtgcttct gcaacacagc tatatcaaac
4321 atgcaaagcc tctggcacct gccccccga tggtattcct aaagtggaag gcacaactat
4381 tgcagatcaa ctttttaaat atggcagcct aggggtgttt tttggaggtt tgggtatagg
4441 tacagggtga ggctctgggt gtagggcagg ctatgtgcca ttgtccactc gtccctccac
4501 tagtagtatt accacgtcca ccattcgtcc ccctgtaact gtagaacca tttgtccctt
4561 agaaccatct atagtttcta tgatagaaga aacaacattt attgagtctg gcgcacctgc

4621 tccatctatt ccatcagcaa cagggtttga tgttacaaca tctgcaaata atactcctgc
 4681 aataattaat gtaacatcta taggtgaatc atctgtacaa tcagtttcta cacatttaaa
 4741 tctacattc actgaacat ctataatata gcccccgga cctgcagaag catctggtca
 4801 tgtattgttt tctagtccaa ctattagtag acacacctat gaagaaatcc ctatggatac
 4861 atttgttacc tctactgaca gcagcagtg aacaagtagt acacctattc caggggtctcg
 4921 ccctacgaca cgcccttggt tatatagccg tgccacacaa caggttaag tagtcgaccc
 4981 tgcttttatg tcatcaccac agaaattagt aacatataac aatcctgttt ttgagggcgt
 5041 tgatacagat gaaactataa tttttgatcg ttcacaactt ttacctgcac cggatcctga
 5101 ttttttagac attatagctt tgcataggcc tgcattaacc tctcgaagag gtactgttag
 5161 gtttagcagg cttggtaata aggccaccct acgtacacgt agtggaagac aaattggggc
 5221 acgggtacat tattatcatg atattagtag tatccagcct gctgaagttc aggaagacat
 5281 agaattgcaa cctttattac cacagtctgt gtccccttac actattaatg atgggttgta
 5341 tgatgtgtat gcagattcct tgcagcaacc cagtttccac ttacctcca cactttctac
 5401 ccataataat actttcactg tacctattaa tagtggtatt gactttgtat atcaaccac
 5461 tatgtccatt gagtcaggtc ctgacattcc attacctcgt ttaccacac atactccttt
 5521 tgttcctata gccctacag ctccatctac atctattatt gttgatggta cagattttat
 5581 tttacatcct agttattttt tactacgtcg caggcgtaaa cgttttccat atttttttac
 5641 agatgtccgt gtggcgccct agtgaggcca ctgtgtacct gcctcctgta cctgtctcta
 5701 aggttgtaag cactgatgag tatgtgtctc gcacaagcat ctattattat gcaggcagtt
 5761 ctcgattact aacagtagga catccctatt tttctatta aaacaccagt agtggtaatg
 5821 gtaaaaaagt tttagtctcc aaggtgtctg gcctgcaata cagggtatit agaattaaat
 5881 tgccggaccc taataaattt ggttttccag atacatcttt ttataacca gaaacccaaa
 5941 gggttggtgt ggccctgtaca ggcttgga aa ttggtagggg acagccttta ggtgtgggta
 6001 ttagtgggca tcttttatta aacaagtttg atgatactga aaccagtaac aaatatgctg
 6061 gtaaacctgg tatagataat agggaatgtt tatctatgga ttataagcag actcagttat
 6121 gcatttttag atgcaaacct cctatagggt aacattgggg taagggaacc ccttgtaata
 6181 ataattcagg aaatcctggg gattgtcctc ccctacagct cattaacagt gtaatacagg
 6241 atggggacat ggtagataca ggatttggtt gcatggattt taataacctg caagctagta
 6301 aaagtgatgt gccattgat atatgtagca gtgtatgtaa gtatccagat tatttgcaaa
 6361 tggctagcga gccatagtgt gacagtttgt tcttttttct tagacgtgag caaatgtttg
 6421 ttagacactt ttttaatagg gccggtacct taggtgaccc tgtgccaggt gatttatata
 6481 tacaagggtc taactctggc aatactgcca ctgtacaaag cagtgccttt tttcctactc
 6541 ctagtggttc tatggtaacc tcagaatccc aattatttaa taaaccgtac tggttacaac
 6601 gtgcgcaggg ccacaataat ggcatatgtt ggggcaatca gttgtttgtc acagttgtgg
 6661 ataccactcg tagcactaac atgactttat gtgctgaggt taaaaaggaa agcacatata
 6721 aaaatgaaaa ttttaaggaa taccttcgtc atggcgagga atttgattta caatttattt
 6781 ttcaattgtg caaaattaca ttaacagctg atgttatgac atacattcat aagatggatg
 6841 ccaactattt agaggactgg caatttggtc ttacccacc accgtctgca tctttggagg
 6901 acacatacag atttgtcact tctactgcta taacttgta aaaaaacaca ccacctaaag
 6961 gaaaggaaga tcttttaaag gactatatgt tttgggaggt ggatttaaaa gaaaagtttt
 7021 ctgcagattt agatcagttt ctttttaggta ggaagttttt gttacaggca gggctacagg
 7081 ctaggcccaa actaaaacgc cctgcatcat cggcccccag tacctccaca aagaagaaaa
 7141 aggttaaaag gtaaccattg tctgttggtt aattgtctgt gtcattgtat tgtgtgtat
 7201 gtcaaacaca ggttaaaagg taaccattgt ttgttatgta attgttttgt gtgtgtactg

7261 tgttgtttgc atgttatgta tgtgtgtgca tgtttgttgt atttgtcagt tctgtatgt
7321 atgttttgtg tatgtattaa taaagtactg tatttactaa actatttata gtagtcctat
7381 gttatgttat ggttgcaccc acatgagtaa caatacagtt gtcctaatac tattgcatct
7441 cctgccctac cctgtgtccc ctgccctacc ctgtgtccta ctttgttaca ctactaatta
7501 gccttatact ctccattttg taccattttg tactatccac cattttaaat cctaaccgaa
7561 ttoggttggc cttggcacia ctttggttgt cttggcaca gtaacaacta tttttatata
7621 agtttcagca aactgcttaa tcctttggtt tctgcagtc cactggtcta cacttgttgt
7681 ccgcctaata ctgacttctt gctgactcac aggtcctgca gtgcagctaa acaatacatt
7741 gcctaacatt gcatgtttta aactgctttt aggcacatat tttattttaa ctttcaatgc
7801 actaattaca gtgttggcct acacaagtac atcctacgcc aaatatgtct tgtaaacat
7861 gattaaatac tgttactcac caggtgtgca ctacacgacc gggtacggtt accgtacca
7921 caaccacttt tttttataat ta

//