

Human papillomavirus type 16 K3079 DNA, complete genome

GenBank: LC647442.1

[FASTA Graphics](#)

[Go to:](#)

LOCUS LC647442 7905 bp DNA circular VRL 25-AUG-2021

DEFINITION Human papillomavirus type 16 K3079 DNA, complete genome.

ACCESSION LC647442

VERSION LC647442.1

KEYWORDS .

SOURCE Human papillomavirus type 16

ORGANISM [Human papillomavirus type 16](#)

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

REFERENCE 1

AUTHORS Tanaka,K., Iwata,T. and Kukimoto,I.

TITLE Evolutionary history of human papillomavirus 16/18/58 variants
exclusively detected in Japan

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 7905)

AUTHORS Kukimoto,I. and Iwata,T.

TITLE Direct Submission

JOURNAL Submitted (20-AUG-2021) Contact:Iwao Kukimoto National Institute
of

Infectious Diseases, Pathogen Genomics Center; 4-7-1 Gakuen,
Musashi-murayama, Tokyo 208-0011, Japan

COMMENT ##Genome-Assembly-Data-START##

Assembly Method :: VirusTAP

Sequencing Technology :: Illumina MiSeq

##Genome-Assembly-Data-END##

FEATURES Location/Qualifiers

source 1..7905
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="K3079"
/host="Homo sapiens"
/db_xref="taxon:[333760](#)"
/country="Japan"
/note="lineage A3"

[gene](#) 104..559
/gene="E6"

[CDS](#) 104..559
/gene="E6"
/codon_start=1

	/product="E6 protein" /protein_id="BDA81712.1" /translation="MFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIVYRDGNPYAVCDKCLKFYISKISEYRYYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPPEEKQRHLDDKKQRFHNIRGRWTGRCMSCCRSSRTRRETQL"
gene	562..858
	/gene="E7" CDS 562..858 /gene="E7" /codon_start=1 /product="E7 protein" /protein_id="BDA81713.1" /translation="MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFCKCDSTLRCLCVQSTHVDIRTLEDLLMGTGLGIVCPICSQKP"
gene	865..2814
	/gene="E1" CDS 865..2814 /gene="E1" /codon_start=1 /product="E1 protein" /protein_id="BDA81714.1" /translation="MADPAGTNGEEGTGCNGWFYVEAVVEKKTGDAISDDENENDSDTGEDLVDFIVNDNDYLTQAETETAHALFTAQEAKQHRDAVQVLKRKYLGSPLSDISGCVDNNISPRKKAICIEKQSRAAKRRLFESEDSGYGNTVEVETQQMLQVEGRHETETPCSQYSGGSGGGCSQYSSGSGGEGVSRHTICQTPLTNINLVKTSNAKAAMLAKFKELYGVSFSELVRPFKSNKSTCCDWCIAAFGLTPSIADSIKTLLQQYCLYLHIQSLACSWGMMVLLLVRYKCGKNRETIEKLKSKLLCVSPMCMMEIEPPKLRSTAAALYWKYTGISNISEVYGDTPEWIQRQTVLQHSFNDCTFELSQMVMQWAYDNDIVDDSEIAYKYAQLADTNSNASAF LKSNSQAKIVKDCATMCRHYKRAEKKQMSMSQWIKYRCDRVDDGGDWKQIVMFLRYQGVEFMSFLTALKRFLQGIPKKNCILLYGAANTGKSLFGMSLMKFLQGSVICFVNSKSHFWLQPLADAKIGMLDDATVPCWNYIDDNLRLNALDGNLVSMQVHVRPLVQLKCPPLLITSNINAGTDSRWPYLHNRLVVFTFPNEFPFDENGPNVYELNDKNWKSFFSRTWSRSLSHEDEDKENDGDSLPTFKCVSGQNTNTL"
gene	2756..3853
	/gene="E2" CDS 2756..3853 /gene="E2" /codon_start=1 /product="E2 protein" /protein_id="BDA81715.1" /translation="METLCQRLNVCQDKILTHYENDSTDLRDHIDYWKHMRLECAIYYKAREMGFKHINHQVVPTLAVSKNKALQAIELQLTLETIYNSQYSNEKWTQLQDVSLEVYLTAPTGCIAKKHGYTVEVQFDGDICNTMHYTNWTHIYICEEASVTVEGQVDYGYGLYYVHEGIRTYFVQFKDDAEKYSKNKVWEVHAGGQVILCPTSVFSSNEVSSPEIIRQHLANH

SAATHTKAVALGTEETQTTIQRPRSEPDGTGNPCHTTKLLHRDSVDSAPILTAFNSSHK
GRINCNSNTTPIVHLKGDANTLKCLRYRFKKHCKLYTAVSSSTWHWTGHNVKHKSAIVT
LTYDSEWQRDQFLSQVKIPKTIITVSTGFMSI"

[gene](#)

<3333..3620

/gene="E4"

[CDS](#)

<3333..3620

/gene="E4"

/codon_start=1

/product="E4 protein"

/protein_id="[BDA81716.1](#)"

/translation="YYVLHLCLAATKYPLLKLLGSTWPTTPPRPIPKPSPWAPKKHRR
LSSDQDQSQTPETPATPLSCCTETQWTVLQSSLHLTAHTKDGLTVIVTLHP"

[gene](#)

3850..4101

/gene="E5"

[CDS](#)

3850..4101

/gene="E5"

/codon_start=1

/product="E5 protein"

/protein_id="[BDA81717.1](#)"

/translation="MTNLDTASTLLACFLLCFCVLLCVCLLIRPLLLSVSTYTSII
LVLLLWITAASAFRCFIVYIVFVYIPLFLIHTHARFLIT"

[gene](#)

4236..5657

/gene="L2"

[CDS](#)

4236..5657

/gene="L2"

/codon_start=1

/product="L2 protein"

/protein_id="[BDA81718.1](#)"

/translation="MRHKRSARKTKRASATQLYKTCKQAGTCPPDII PKVEGKTIADQ
ILQYGSMSGVFFGGLGIGTSGTGGRGTGYIPLGTRPPTATDTLAPVRPLTVDPVGP
PSIVSLVEETS FIDAGAPTSVPSIPPDVSGFSITTSTDTPAILDINNVTVTTHNN
PTFTDPSVLQPPTPAETGGHFTLSSSTISTHNYEEIPMDTFIVSTNPNTVTSSTPI
SRPVARLGLYSRTTQQVKVVDPAFVTTPTKLITYDNPAYEGIDVDNTLYFPSNDNSIN
IAPDPDFLDIVALHRPALTSRRTGIRYSRIGNKQTLRTRSGKSIGAKVHYYYDL
SNIDPAEEIELQTITPSTYTTTSHAASPTSINNGLYDIYADDFITDTSTTPVPSVP
STSLSGYIPANTTIPFGGAYNIPLVSGPDIPINITDQAPSLIPIVPGSPQYTI
IADAGDFYLHP
SYMYLRKRRKRLPYFFSDVSLAA"

[gene](#)

5560..7155

/gene="L1"

[CDS](#)

5560..7155

/gene="L1"

/codon_start=1

/product="L1 protein"

/protein_id="[BDA81719.1](#)"

/translation="MQVTFIYILVITCYENDVNVYHIFQMSLWLPSEATVYLPPVPV
SKVVSTDEYVARTNIYYHAGTSRLLAVGHYPYFIKPNNNKILVPKVSGLQYRVFRIH
LPDPNKFGFPDTSFYNPDTQRLVWACVGVEVGRGQPLGVGISGHPLLNKLLDDTENASA
YAANAGVDNRECISMDYKQTQLCLIGCKPPIGEHWGKGSPTCNVAVNPGDCPPLELIN
TVIQDGDMDVTGFGAMDFTTLQANKSEVPLDICTSICKYPDYIKMVSEPYGDSLFFYL
RREQMFVRHLFNRAGAVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMVTSDAQIFN
KPYWLQRAQGHNNGICWGNQLFVTVDTRSTNMSLCAAISTSETTYKNTNFKEYLRH
GEEYDLQFIFQLCKITLTADVMTYIHSMNSTILEDWNFGLQPPPGGTLEDTYRFVTSQ
AIACQKHTPPAPKEDPLKKYTFWEVNLKEKFSADLDQFPLGRKFLLQAGLKAKPKFTL
GKRKATPTTSSTSTTAKRKKRKL"

ORIGIN

1 actacaataa ttcatgtata aaactaaggg cgtaaccgaa atcggttgaa ccgaaaccgg
61 ttagtataaa agcagacatt ttatgcacca aaaaagaact gcaatgtttc aggaccaca
121 ggagcgaccc agaaagtac cacagttatg cacagagctg caaacaacta tacatgatat
181 aatattagaa tgttgtgtact gcaagcaaca gttactgcga cgtgaggtat atgactttgc
241 ttttcgggat ttatgcatag tatatagaga tgggaatcca tatgctgtat gtgataaatg
301 tttaaagttt tattctaaaa ttagtgagta tagatattat tggtatagtt tgtatggaac
361 aacattagaa cagcaataca acaaaccggt gtgtgatttg ttaattaggt gtattaactg
421 tcaaaagcca ctgtgtcctg aagaaaagca aagacatctg gacaaaaagc aaagattcca
481 taatataagg ggtcgggtga ccggtcgatg tatgtcttgt tgcagatcat caagaacacg
541 tagagaaacc cagctgtaat catgcatgga gatacaccta cattgcatga atatatgtta
601 gatttgcaac cagagacaac tgatctctac tggtatgagc aattaaatga cagctcagag
661 gaggaggatg aaatagatgg tccagctgga caagcagaac cggacagagc ccattacaat
721 attgtaacct ttgtttgcaa gtgtgactct acgcttcggg tgtgctgaca aagcacacac
781 gtagacattc gtactttgga agacctgtta atgggcacac taggaattgt gtgccccatc
841 tgttctcaga aaccataatc taccatggct gatcctgcag gtaccaatgg ggaagagggt
901 acgggatgta atggatggtt ttatgtagag gctgtagtgg aaaaaaaac aggggatgct
961 atatcagatg acgagaacga aaatgacagt gatacagggt aagatttggt agattttata
1021 gtaaatgata atgattattt aacacaggca gaaacagaga cagcacatgc gttgtttact
1081 gcacaggaag caaaacaaca tagagatgca gtacagggtc taaaacgaaa gtatttggtg
1141 agtccactta gtgatattag tggatgtgta gacaataata ttagtcctag attaaaagct
1201 atatgtatag aaaaacaaag tagagctgca aaaaggagat tatttgaaag cgaagacagc
1261 gggatgga atactgaagt ggaaactcag cagatgttac aggtagaagg gcgccatgag
1321 actgaaacac catgtagtca gtatagtggt ggaagtgggg gtggttgagc tcagtacagt
1381 agtggaagtg ggggagaggg tgtagtgtaa agacacacta tatgccaaac accacttaca
1441 aatattttta atgtactaaa aactagtaat gcaaaggcag caatgttagc aaaattttaa
1501 gagttatacg gggtaggtt ttcagaatta gtaagaccat taaaagtaa taaatcaacg
1561 tggtgcgatt ggtgtattgc tgcatttga cttacacca gtatagctga cagtataaaa
1621 aactattac aacaatattg tttatattta cacattcaaa gtttagcatg ttcattggga
1681 atggttgtgt tactattagt aagatataaa tgtggaaaa atagagaaac aattgaaaaa
1741 ttgctgtcta aactattatg tgtgtctcca atgtgtatga tgatagagcc tccaaaattg
1801 cgtagtagag cagcagcatt atattggtat aaaacaggta tatcaaatat tagtgaagtg
1861 tatggagaca cgccagaatg gatacaaaga caaacagtat tacaacatag ttttaattg
1921 tgtacatttg aattatcaca gatggtacaa tgggcctacg ataattgacat agtagacgat

1981 agtgaaattg catataaata tgcacaattg gcagacacta atagtaatgc aagtgccttt
2041 ctaaaaagta attcacaggc aaaaattgta aaggattgtg caacaatgtg tagacattat
2101 aaacgagcag aaaaaaaca aatgagtatg agtcaatgga taaaatatag atgtgatag
2161 gtagatgatg gaggtgattg gaagcaaatt gttatgtttt taaggatatca aggtgtagag
2221 tttatgtcat ttttaactgc attaaaaaga tttttgcaag gcatacctaa aaaaaattgc
2281 atattactat atgggtgcagc taacacaggc aaatcattat ttggtatgag tttaatgaaa
2341 tttctgcaag ggtctgtaat atgttttgta aattctaaaa gccatttttg gttacaacca
2401 ttagcagatg ccaaaatagg tatgttagat gatgctacag tgccctgttg gaactacata
2461 gatgacaatt taagaaatgc attggatgga aatttagttt ctatggatgt aaagcataga
2521 ccattggtag aactaaaatg ccctccatta ttaattacat ctaacattaa tgctggtaca
2581 gattctaggt ggcttattt acataataga ttggtggtgt ttacatttcc taatgagttt
2641 ccatttgacg aaaacggaaa tccagtgtat gagcttaatg ataagaactg gaaatccttt
2701 ttctcaagga cgtggtccag attaagtttg cacgaggacg aggacaagga aaacgatgga
2761 gactctttgc caacgtttaa atgtgtgtca ggacaaaata ctaacacatt atgaaaatga
2821 tagtacagac ctacgtgacc atatagacta ttggaaacac atgcgccatg aatgtgctat
2881 ttattacaag gccagagaaa tgggatttaa acatattaac caccagggtg tgccaacact
2941 ggctgtatca aagaataaag cattacaagc aattgaactg caactaacgt tagaaacaat
3001 atataactca caatatagta atgaaaagtg gacattacaa gacgttagcc ttgaagtgt
3061 ttttaactgca ccaacaggat gtataaaaaa acatggatat acagtggaag tgcagtttga
3121 tggagacata tgcaatacaa tgcattatac aaactggaca catatatata tttgtgaaga
3181 agcatcagta actgtggtag agggcgaagt tgactattat ggtttatatt atgttcatga
3241 aggaatacga acatattttg tgcagtttaa agatgatgca gaaaaatata gtaaaaataa
3301 agtatgggaa gttcatgcgg gtggtcaggc aatattatgt cctacatctg tgtttagcag
3361 caacgaagta tcctctcctg aaattattag gcagcacttg gccaccact ccgccgcgac
3421 ccataccaaa gccgtcgct tgggcaccga agaaacacag acgactatcc agcgaccaag
3481 atcagagcca gacaccgga acccctgcc caccactaag ttgttgca gagactcagt
3541 ggacagtgtc ccaatcctca ctgcatttaa cagctcacac aaaggacgga ttaactgtaa
3601 tagtaacact acaccatag tacattttaa aggtgatgct aatactttaa aatgtttaag
3661 atatagattt aaaaagcatt gtaaattgta tactgcagtg tcgtctacat ggcattggac
3721 aggacataat gtaaaacata aaagtgcaat tgttacactt acatatgata gtgaatggca
3781 acgtgaccaa tttttgtctc aagttaaaat accaaaaact attacagtgt ctactggatt
3841 tatgtctata tgacaaatct tgatactgca tccacaacat tactggcgtg ctttttgctt
3901 tgcttttgtg tgcttttgtg tgtctgccta ttaatacgtc cgctgctttt gtctgtgtct
3961 acatacacat cattaataat attggtatta ctattgtgga taacagcagc ctctgcgttt
4021 aggtgtttta ttgtatatat tgtatttgtt tatataccat tatttttaat ccatacacat
4081 gcacgctttt taattacata atgtatatgt acataatgta attgttacat ataattgttg
4141 tataccataa cttactattt tttctttttt atttttatat ataatttttt ttttgtttgt
4201 ttgtttgttt ttttaataaac tgttatcact taacaatgcg acacaaacgt tctgcaaaac
4261 gcacaaaacg tgcacggt acccaacttt ataaaacatg caaacaggca ggtacatgtc
4321 cacctgacat tatacctaag gttgaaggca aaactattgc tgatcaaata ttacaatatg
4381 gaagtatggg tgtatttttt ggtgggttag gaattggaac agggcgggt acaggcggac
4441 gcaactggta tattccattg ggaacaaggc ctcccacagc tacagatata cttgctcctg
4501 taagaccccc tttaacagta gatcctgtgg gcccttctga tccttctata gtttctttag
4561 tggaagaaac tagttttatt gatgctggtg caccaacatc tgtaccttcc attccccag

4621 atgtatcagg atttagtatt actacttcaa ctgataccac acctgctata ttagatatta
4681 ataatactgt tactactggt actacacata ataatcccac tttcactgac ccatctgtat
4741 tgcagcctcc aacacctgca gaaactggag ggcattttac actttcatca tocactatta
4801 gtacacataa ttatgaagaa attcctatgg atacatttat tgtttagcaca aaccctaaca
4861 cagtaactag tagcacaccc ataccagggt ctgcgccagt ggcacgccta ggattatata
4921 gtgcgacaac acaacaagtt aaagttgtag accctgcttt tgtaaccact cccactaaac
4981 ttattacata tgataatcct gcatatgaag gtatagatgt ggataatata ttatattttc
5041 ctagtaatga taatagtatt aatatagctc cagatcctga ctttttggat atagttgctt
5101 tacataggcc agcattaacc tctaggcgta ctggcattag gtacagtaga attggtaata
5161 aacaaacact acgtactcgt agtggaaaat ctatagggtc taaggtagatc tattattatg
5221 atttaagtaa tattgatcct gcagaagaaa tagaattaca aactataaca ccttctacat
5281 atactaccac ttcacatgca gcctcaccta cttctattaa taatggatta tatgatattt
5341 atgcagatga ctttattaca gatacttcta caaccccggt accatctgta cctctacat
5401 ctttatcagg ttatattcct gcaaatataa caattccttt tgggtggtgca tacaatatc
5461 ctttagtatc aggtcctgat ataccatta atataactga ccaagctcct tcattaattc
5521 ctatagttcc agggctctca caatatacaa ttattgctga tgcaggtagc ttttatttac
5581 atcctagtta ttacatgtta cgaaaacgac gtaaacgttt accatatttt ttttcagatg
5641 tctctttggc tgcctagtga ggccactgtc tacttgccctc ctgtcccagt atctaagggt
5701 gtaagcacgg atgaatatgt tgcacgcaca aacatatatt atcatgcagg aacatccaga
5761 ctacttgtag ttggacatcc ctattttcct attaaaaaac ctaacaataa caaaatatta
5821 gttcctaaag tatcaggatt acaatacagg gtatttagaa tacatttacc tgaccccaat
5881 aagtttggtt ttcttgacac ctcatcttat aatccagata cacagcggct ggtttgggccc
5941 tgtgtagggtg ttgaggtagg tcgtggtcag ccattagggtg tgggcattag tggccatcct
6001 ttattaaata aattggatga cacagaaaat gctagtgtct atgcagcaaa tgcagggtgtg
6061 gataatagag aatgtatatc tatggattac aaacaaacac aattgtgttt aattggttgc
6121 aaaccaccta taggggaaca ctggggcaaa ggatcccat gtaccaatgt tgcagtaaat
6181 ccagggtgatt gtccaccatt agagttaata aacacagtta ttcaggatgg tgatatgggt
6241 gatactggct ttggtgctat ggactttact acattacagg ctaacaaaag tgaagttcca
6301 ctggatatatt gtacatctat ttgcaaatat ccagattata ttaaaatggt gtcagaacca
6361 tatggcgaca gcttattttt ttatttacga agggaaacaaa tgtttgttag acatttattt
6421 aatagggctg gtgctgttgg tgaaaatgta ccagacgatt tatacattaa aggtctctggg
6481 tctactgcaa atttagccag ttcaaattat tttcctacac ctagtgggtc tatgggtacc
6541 tctgatgccc aaatattcaa taaaccttat tggttacaac gagcacaggg ccacaataat
6601 ggcatttggt ggggtaacca actatttggt actgttggtg atactacacg cagtacaaat
6661 atgtcattat gtgctgcat atctacttca gaaactacat ataaaaatac taactttaag
6721 gagtacctac gacatgggga ggaatatgat ttacagttta tttttcaact gtgcaaaata
6781 accttaactg cagacgttat gacatacata cattctatga attccactat tttggaggac
6841 tggaaatttg gtctacaacc tccccagga ggcacactag aagatactta taggtttgta
6901 acatcccagg caattgcttg tcaaaaacat acacctccag cacctaaga agatcccctt
6961 aaaaaatata ctttttggga agtaaattta aaggaaaagt tttctgcaga cctagatcag
7021 tttccttttag gacgcaaatt tttactacaa gcaggattaa aggccaaacc aaaatttaca
7081 ttaggaaaac gaaaagctac acccaccacc tcatctacct ctacaactgc taaacgcaaa
7141 aaacgtaagc tgtaagtatt gtatgtatgt tgaattagtg ttgtttgttg tttatatgtt
7201 tgtatgtgct tgtatgtgct tgtaaatatt aagttgtatg tgtgtttgta tgtatggtat

7261 aataaacacg tgtgtatgtg tttttaaatg cttgtgtaac tatttgtgtca tgcaacataa
7321 ataaacttat tgtttcaaca cctactaatt gtgttggtgt tattcattgt atataaacta
7381 tatttgctac atcctgtttt tgttttatat atactatatt ttgtagcgcc agcggccatt
7441 ttgtagcttc aaccgaattc ggttgcatgc tttttggcac aaaatgtgtt tttttaaata
7501 gttctatgtc agcaactata gtttaaaact gtacgtttcc tgcttgccat gcgtgccaaa
7561 tccctgtttt cctgacctgc actgcttgcc aaccattcca ttgtttttta cactgcacta
7621 tgtgcaacta ctgaatcact atgtacattg tgtcatataa aataaatcac tatgcgccaa
7681 cgccttacat accgctgtta ggcacatatt tttggcttgt tttaactaac ctaattgcat
7741 atttggcata aggttttaaac ttctaaggcc aactaaatgt caccctagtt catacatgaa
7801 ctgtgtaaag gttagtcata cattgttcat ttgtaaaact gcacatgggt gtgtgcaaac
7861 cgtttttaggt tacacattta caagcaactt atataataat actaa

//