Human gammaherpesvirus 4, complete genome

NCBI Reference Sequence: NC_007605.1

FASTA Graphics

LOCUS NC_007605 171823 bp DNA circular VRL 13-AUG-2018
DEFINITION Human gammaherpesvirus 4, complete genome.
ACCESSION NC_007605
VERSION NC_007605.1
DBLINK BioProject: PRJNA485481
KEYWORDS RefSeq; complete genome.
SOURCE Human gammaherpesvirus 4 (Epstein-Barr virus)
ORGANISM Human gammaherpesvirus 4
Viruses; Herpesvirales; Herpesviridae; Gammaherpesvirinae;
Lymphocryptovirus.
REFERENCE 1 (bases 1 to 171823)
AUTHORS de Jesus,O., Smith,P.R., Spender,L.C., Elgueta Karstegl,C.,
Niller,H.H., Huang,D. and Farrell,P.J.
TITLE Updated Epstein-Barr virus (EBV) DNA sequence and analysis of a
promoter for the BART (CST, BARF0) RNAs of EBV
JOURNAL J. Gen. Virol. 84 (PT 6), 1443-1450 (2003)
PUBMED 12771413
REFERENCE 2 (bases 1 to 171823)
AUTHORS Parker,B.D., Bankier,A., Satchwell,S., Barrell,B. and
Farrell,P.J.
TITLE Sequence and transcription of Raji Epstein-Barr virus DNA spanning
the B95-8 deletion region
JOURNAL Virology 179 (1), 339-346 (1990)
PUBMED 2171209
REFERENCE 3 (sites)
AUTHORS Hatfull,G., Bankier,A.T., Barrell,B.G. and Farrell,P.J.
TITLE Sequence analysis of Raji Epstein-Barr virus DNA
JOURNAL Virology 164 (2), 334-340 (1988)
PUBMED 2835854
REFERENCE 4 (bases 1 to 171823)
AUTHORS Baer,R., Bankier,A.T., Biggin,M.D., Deininger,P.L., Farrell,P.J.,
Gibson,T.J., Hatfull,G., Hudson,G.S., Satchwell,S.C., Seguin,C.,
Tuffnell,P.S. and Barrell,B.G.
TITLE DNA sequence and expression of the B95-8 Epstein-Barr virus genome
JOURNAL Nature 310 (5974), 207-211 (1984)
PUBMED 6087149
REFERENCE 5 (sites)
AUTHORS Jones,M.D., Foster,L., Sheedy,T. and Griffin,B.E.
TITLE The EB virus genome in Daudi Burkitt's lymphoma cells has a
deletion similar to that observed in a non-transforming strain
(P3HR-1) of the virus

JOURNAL EMBO J. 3 (4), 813-821 (1984)
PUBMED 6327290
REFERENCE 6 (sites)
AUTHORS Jeang,K.T. and Hayward,S.D.
TITLE Organization of the Epstein-Barr virus DNA molecule. III. Location of the P3HR-1 deletion junction and characterization of the NotI repeat units that form part of the template for an abundant 12-O-tetradecanoylphorbol-13-acetate-induced mRNA transcript
JOURNAL J. Virol. 48 (1), 135-148 (1983)
PUBMED 6310141
REFERENCE 7 (bases 1 to 171823)
CONSRTM NCBI Genome Project
TITLE Direct Submission
JOURNAL Submitted (10-NOV-2005) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA
REFERENCE 8 (bases 1 to 171823)
AUTHORS Farrell,P.J. and Davison,A.J.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2005) Department of Virology, Ludwig Institute for Cancer Research, Faculty of Medicine, Imperial College, St Mary's Campus, Norfolk Place, London W2 1PG, UK
COMMENT VALIDATED REFSEQ: This record has undergone validation or preliminary review. This record has been curated by Drs. Andrew Davison (Institute of Virology, Glasgow, UK) and Paul Farrell (Imperial College School of Medicine, London, UK). The reference sequence was derived from AJ507799.
CONSTRUCTION: This sequence was assembled from data for the B95-8 (V01555) and Raji (M35547) strains with corrections. The number of IR1 (W) repeats in V01555 has been reduced from 11.6 to a more typical 7.6, and the missing B95-8 sequence has been restored to give a sequence more representative of wild type EBV.
NUMBERING: Like the modified B95-8 sequence in V01555, this sequence starts 1 bp to the left of the EcoRI site separating EcoRI
Dhet from EcoRI I (i.e. the first A of AGAATTC).
TRANSCRIPTION: Long-range splicing of the EBNA and BART genes generates multiple, complex mRNAs. A single mRNA structure is shown
for each gene, but alternatives may initiate at other promoters and
may involve various exon combinations. TATA and polyA signals are shown as non-experimental where predicted or as experimental where supporting data are available. Predictions of TATA or polyA signals

are not made for some genes.
NOMENCLATURE: The original gene nomenclature has been retained.
Genes presumably inherited from the common ancestor of alpha-,
beta- and gammaherpesviruses (core genes), from the common
ancestor
of beta- and gammaherpesviruses (betagamma genes), or from the
common ancestor of gammaherpesviruses (gamma genes) are indicated.
A standard protein nomenclature has been applied so that orthologs
have the same name in all herpesviruses.
CONFIDENCE: Evidence for encoded protein functions is more
convincing for some ORFs than others, and rests on a combination
of
factors, including evolutionary conservation, expression and
functional data. The most questionable ORFs are annotated as such.
COMPLETENESS: full length.

FEATURES Location/Qualifiers

source 1..171823
/organism="Human gammaherpesvirus 4"
/mol_type="genomic DNA"
/strain="B95-8"
/strain="Raji"
/host="Homo sapiens"
/db_xref="taxon:10376"
/type="1"
/common="Epstein-Barr virus"
/acronym="EBV"
/acronym="HHV-4"

source join(1..139723,151555..171823)
/organism="Human gammaherpesvirus 4"
/mol_type="genomic DNA"
/strain="B95-8"
/host="Homo sapiens"
/db_xref="taxon:10376"
/type="1"
/common="Epstein-Barr virus"
/acronym="EBV"

source 139724..151554
/organism="Human gammaherpesvirus 4"
/mol_type="genomic DNA"
/strain="Raji"
/host="Homo sapiens"
/db_xref="taxon:10376"
/type="1"
/common="Epstein-Barr virus"

/acronym="EBV"

略

regulatory 23568..23573

/regulatory_class="TATA_box"

/gene="EBNA-LP"

/locus_tag="HHV4_EBNA-LP"

/experiment="experimental evidence, no additional details recorded"

/note="W promoter (Wp) used in earlier stages of latent infection"

repeat_region 24289..27360

/rpt_family="IR1; W"

/rpt_type=direct

gene <24829..25980

/gene="BWRF1"

/locus_tag="HHV4_BWRF1.5"

/db_xref="GeneID:3783737"

CDS <24829..25980

/gene="BWRF1"

/locus_tag="HHV4_BWRF1.5"

/note="questionable whether a protein function is encoded"

/codon_start=1

/product="protein BWRF1"

/protein_id="YP_401640.1"

/db_xref="UniProtKB/TrEMBL:Q8AZK3"

/db_xref="GeneID:3783737"

/translation="VWEAEGRPRPGEVEGDRPGLCWQSPGDPLRPSGPGRSPSAPQTD
PRVSRQGPASSGAAGSPPQAPQTRVSASRADRPRAWRLLGASRRGWFCPSLCPSEEPG
TSGTPEPLGPASRRPPGLRSPLSPVKPKECLRGATLGAQAPESRGQGHLRVPPRVPGQ
PEGPRQPGRPQRPVPRPFPGLQSPGCPPEGTLGVPSPPLQARASPSRRGASLGPQVQP
HRDPSGPDPPTGPSLCPPAPLQPSLHPRPQLLASPGPPGQPEGPRQPGRVAFPLPWPL
LPASHPSPLSLPPHRVHQAGRRDPGGPVSVPPAAAQSLPPGKGASFSPPSLRPSLLCT
VCKVQPPTPVHGSRAQPRPPLPTVDRPSVHPGHPRPPVSTPVPSRGDFM"

略

ORIGIN

```
       1 agaattcgtc ttgctctatt cacccttact tttcttcttg cccgttctct ttcttagtat
      61 gaatccagta tgcctgcctg taattgttgc gccctacctc ttttggctgg cggctattgc
     121 cgcctcgtgt ttcacggcct cagttagtac cgttgtgacc gccaccggct tggccctctc
     181 acttctactc ttggcagcag tggccagctc atatgccgct gcacaaagga aactgctgac
     241 accggtgaca gtgcttactg cggttgtcac ttgtgagtac acacgcacca tttacaatgc
     301 atgatgttcg tgagattgat ctgtctctaa cagttcactt cctctgcttt tctcctcagt
     361 ctttgcaatt tgcctaacat ggaggattga ggacccacct tttaattctc ttctgtttgc
     421 attgctggcc gcagctggcg gactacaagg catttacggt tagtgtgcct ctgttatgaa
     481 atgcaggttt gacttcatat gtatgccttg gcatgacgtc aactttactt ttatttcagt
```

541 tctggtgatg cttgtgctcc tgatactagc gtacagaagg agatggcgcc gtttgactgt
601 ttgtggcggc atcatgtttt tggcatgtgt acttgtcctc atcgtcgacg ctgttttgca
661 gctgagtccc ctccttggag ctgtaactgt ggtttccatg acgctgctgc tactggcttt
721 cgtcctctgg ctctcttcgc caggggggcct aggtactctt ggtgcagccc ttttaacatt
781 ggcagcaggt aagccacacg tgtgacattg cttgcctttt tgccacatgt tttctggaca
841 caggactaac catgccatct ctgattatag ctctggcact gctagcgtca ctgattttgg
901 gcacacttaa cttgactaca atgttccttc tcatgctcct atggacactt ggtaagtttt
961 cccttccttt aactcattac ttgttctttt gtaatcgcag ctctaacttg gcatctcttt
1021 tacagtggtt ctcctgattt gctcttcgtg ctcttcatgt ccactgagca agatccttct
1081 ggcacgactg ttcctatatg ctctcgcact cttgttgcta gcctccgcgc taatcgctgg
1141 tggcagtatt ttgcaaacaa acttcaagag tttaagcagc actgaattta tacccagtga
1201 gtatctattt gttactcctg tttagttgaa gaaaacaagc tattggattg taacacacat
1261 tttacgcttt gttccttaga tttgttctgc atgttattac tgattgtcgc tggcatactc
1321 ttcattcttg ctatcctgac cgaatggggc agtggaaata gaacatacgg tccagttttt
1381 atgtgcctcg gtggcctgct caccatggta gccggcgctg tgtggctgac ggtgatgtct
1441 aacacgcttt tgtctgcctg gattcttaca gcaggattcc tgattttcct cattggtaag
1501 tgtgacacca acaggtgttg ccttgttatg tcaccgttct gacacatgac ttacatgggt
1561 ttggcttttg taggctttgc cctctttggg gtcattagat gctgccgcta ctgctgctac
1621 tactgcctta cactggaaag tgaggagcgc ccaccgaccc catatcgcaa cactgtataa
1681 aggtaagtat tattaaattt tagagacact atcacgtgta acttgacgtg caaggatgga
1741 agagaggggc agggaaacgc aaatgccggt tgcccggtat gggggcccgt ttattatggt
1801 aaggctcttc gggcaagatg gagaggcaaa catacaggag gaaaggctat atgagctact
1861 ctctgaccca cgctccgcgc tcggcctaga cccggggccc ctgattgctg agaacctgct
1921 gctagtggcg ctgcgtggca ccaacaacga tcccaggcct cagcgtcagg agagggccag
1981 agaactggcc ctcgttggca ttctactagg aaacggcgag cagggtgaac acttgggcac
2041 ggagagtgcc ctggaggcct caggcaacaa ctatgtgtat gcctacggac cagactggat
2101 ggcaaggcct tccacatggt ccgcggaaat ccagcaattc ctgcgactcc tgggcgccac

略

20641 cgtcttgccc tgcgcccggg gccaccttca tcaccgtcgc tgactccgcc atccaagcct
20701 aggggagacc gaagtgaagg ccctggacca acccggcccg ggccccccgg tatcgggcca
20761 gaggtaagtg gactttaatt ttttctgcta agcccaacac tccaccacac ccaggcacac
20821 actacacaca cccacccgtc tcagggtccc ctcggacagc tcctaagaag gcaccggtcg
20881 cccagtccta ccagaggggg ccaagaaccc agacgagtcc gtagaagggt cctcgtccag
20941 caagaagagg aggtggtaag cggttcacct tcaggggtaa gtaacctgac ctctccaggg
21001 ctcacataaa gggaggctta gtatacatgc ttcttgcttt tcacaggaac ctgggggcta
21061 gtctgggtgg gattaggctg cctcaagttg catcagccag ggcttcatgc cctcctcagt
21121 tccctagtcc ccgggcttca ggcccccctcc gtccccgtcc tccagagacc cgggcttcag
21181 gccctgcctc tcctgttacc cttttagaac cacagcctgg acacatgtgc cagacgcctt
21241 ggcctctaag gccctcgggt ccccctggac cccggcctca gcaaccctgc tgctcccctc
21301 ctgccacccc agcctccccc cctccccgtc ccccttcgct cctgatcctc ccccggtccc
21361 cagtagggcc gcctgccccc ctgcacccag tacctgcccc tcttggccac gcaccccggg
21421 ccaggccacc ttagacccgg ccaagcccca tccctgaaga cccagcggcc attctctctg
21481 gtaacgagca gagaagaagt agaggcccgc ggccattggg cccagattga gagaccagtc
21541 caggggcccg aggttggagc cagcgggcac ccgaggtccc agcacccggt ccctccgggg

21601 ggcagagaca ggcagggccc cccggcagct ggccccgagg aggcgcccgg agtggggccg
21661 gtcggctggg ctggccgagc ccgggtctgg gaggtctggg gtggcgagcc tgctgtctca
21721 ggaggggcct ggctccgccg ggtggccctg gggtaagtct gggaggcaga gggtcggcct
21781 aggcccgggg aagtggaggg ggatcgcccg ggtctctgtt ggcagagtcc gggcgatcct
21841 ctgagaccct ccgggcccgg acggtcgccc tcagcccccc agacagaccc cagggtctcc
21901 aggcagggtc cggcatcttc aggggcagca ggctcaccac cacaggcccc ccagacccgg
21961 gtctcggcca gccgagccga ccggccccgc gcctggcgcc tcctcggggc cagccgccgg
22021 ggttggttct gcccctctct ctgtccttca gaggaaccag ggacctcggg caccccagag
22081 cccctcgggc ccgcctccag gcgccctcct ggtctccgct cccctctgag ccccgttaaa
22141 cccaaagaat gtctgagggg agccaccctc ggggcccagg ccccagagtc cagaggtcag
22201 gggcacctca gggtgcctcc ccgggtccca ggccagccgg agggaccccg gcagcccggg
22261 cggccccaga ggccggttcc tcgccccttc cccgggcttc agagcccagg atgtccccca
22321 gaagggaccc taggcgtccc ctctcctccc ctccaggccc gagcctctcc ctcgcggaga
22381 ggggcctctt tgggccctca agtccagccc caccgagacc cgagtggccc ggatcccccc
22441 accggccctt ctctctgtcc ccctgctcct ctccaacctt cgctccaccc tagaccccag
22501 cttctggcct ccccgggtcc accaggccag ccggagggac cccggcagcc cgggcgagtc
22561 gccttccctc tcccctggcc tctccttccc gcctcccacc cgagcccccc cagcttgcct
22621 ccccaccggg tccatcaggc cggccggagg gaccccggcg gcccggtgtc agtcccccct
22681 gcagccgccc agtctctgcc tccaggcaag ggcgccagct tttctccccc cagcctgagg
22741 cccagtctcc tgtgcactgt ctgtaaagtc cagcctccca cgcccgtcca cggctcccgg
22801 gcccagcctc gtccaccccc ccccacggtg gacaggccct ctgtccaccc gggccatccc
22861 cgccccccctg tgtccacccc agtcccgtcc agggggggact ttatgtgacc cttgggcctg
22921 gctccccata gactcccatg taagcctgcc tcgagtaggt gcctccagag ccccttttgc
22981 ccccctggcg gcccagcccg acccccgggc gcccccaaac tttgtccaga tgtccagggg
23041 tccccgaggg tgaggcccag ccccctcccg cccctgtcca ctgccccggt ccccccagaa
23101 gcccccaaaa gtagaggctc aggccatgcg cgccctgtca ccaggcctgc caaagagcca
23161 gatctaaggc cgggagaggc agccccaaag cgggtgcagt aacaggtaat ctctggtagt
23221 gatttggacc cgaaatctga cactttagag ctctggagga ctttaaaact ctaaaaatca
23281 aaactttaga ggcgaatggg cgccattttg tccccacgcg cgcataatgg cggacctagg
23341 cctaaaaccc ccaggaagcg ggtctatggt tggctgcgct gctgctatct ttagagggga
23401 aaagaggaat aagcccccag acaggggagt gggcttgttt gtgacttcac caaaggtcag
23461 ggcccaaggg ggttcgcgtt gctaggccac cttctcagtc cagcgcgttt acgtaagcca
23521 gacagcagcc aattgtcagt tctagggagg gggaccactg cccctggtat aaagtggtcc
23581 tgcagctatt tctggtcgca tcagagcgcc aggagtccac acaaatgtaa gagggggtct
23641 tctacctctc cctagccctc cgcccccctcc aaggactcgg gcccagtttc taacttttcc
23701 ccttccctcc ctcgtcttgc cctgcgcccg gggccacctt catcaccgtc gctgactccg
23761 ccatccaagc ctaggggaga ccgaagtgaa ggccctggac caacccggcc cgggcccccc
23821 ggtatcgggc cagaggtaag tggactttaa ttttttctgc taagcccaac actccaccac
23881 acccaggcac acactacaca cacccacccg tctcagggtc ccctcggaca gctcctaaga
23941 aggcaccggt cgcccagtcc taccagaggg ggccaagaac ccagacgagt ccgtagaagg
24001 gtcctcgtcc agcaagaaga ggaggtggta agcggttcac cttcaggggt aagtaacctg
24061 acctctccag ggctcacata aagggaggct tagtatacat gcttcttgct tttcacagga
24121 acctgggggc tagtctgggt gggattaggc tgcctcaagt tgcatcagcc agggcttcat
24181 gccctcctca gttccctagt ccccgggctt caggcccccc ccgtccccgt cctccagaga

24241 cccgggcttc aggccctgcc tctcctgtta ccctttaga accacagcct ggacacatgt
24301 gccagacgcc ttggcctcta aggccctcgg gtccccctgg accccggcct cagcaaccct
24361 gctgctcccc tcctgccacc ccagcctccc cccctccccg tccccttcg ctcctgatcc
24421 tcccccggtc cccagtaggg ccgcctgccc ccctgcaccc agtacctgcc cctcttggcc
24481 acgcaccccg ggccaggcca ccttagaccc ggccaagccc catccctgaa gacccagcgg
24541 ccattctctc tggtaacgag cagagaagaa gtagaggccc gcggccattg ggcccagatt
24601 gagagaccag tccaggggcc cgaggttgga gccagcgggc acccgaggtc ccagcacccg
24661 gtccctccgg ggggcagaga caggcagggc cccccggcag ctggccccga ggaggcgccc
24721 ggagtggggc cggtcggctg ggctggccga gcccgggtct gggaggtctg gggtggcgag
24781 cctgctgtct caggaggggc ctggctccgc cgggtggccc tggggtaagt ctgggaggca
24841 gagggtcggc ctaggcccgg ggaagtggag ggggatcgcc cgggtctctg ttggcagagt
24901 ccgggcgatc ctctgagacc ctccgggccc ggacggtcgc cctcagcccc ccagacagac
24961 cccagggtct ccaggcaggg tccggcatct tcaggggcag caggctcacc accacaggcc
25021 ccccagaccc gggtctcggc cagccgagcc gaccggcccc gcgcctggcg cctcctcggg
25081 gccagccgcc ggggttggtt ctgcccctct ctctgtcctt cagaggaacc agggacctcg
25141 ggcaccccag agcccctcgg gcccgcctcc aggcgccctc ctggtctccg ctcccctctg
25201 agccccgtta aacccaaaga atgtctgagg ggagccaccc tcggggccca ggccccagag
25261 tccagaggtc aggggcacct cagggtgcct ccccgggtcc caggccagcc ggagggaccc
25321 cggcagcccg ggcggcccca gaggccggtt cctcgccccct tccccgggct tcagagccca
25381 ggatgtcccc cagaagggac cctaggcgtc ccctctcctc ccctccaggc ccgagcctct
25441 ccctcgcgga gaggggcctc tttgggccct caagtccagc cccaccgaga cccgagtggc
25501 ccggatcccc ccaccggccc ttctctctgt ccccctgctc ctctccaacc ttcgctccac
25561 cctagacccc agcttctggc ctccccgggt ccaccaggcc agccggaggg accccggcag
25621 cccgggcgag tcgccttccc tctcccctgg cctctccttc ccgcctccca cccgagcccc
25681 ctcagcttgc ctccccaccg ggtccatcag gccggccgga gggaccccgg cggcccggtg
25741 tcagtccccc ctgcagccgc ccagtctctg cctccaggca agggcgccag cttttctccc

略