Hepatitis B virus isolate 1745, complete genome

GenBank: MT114169.1

FASTA Graphics PopSet

Go to:

```
LOCUS       MT114169              3182 bp    DNA     circular VRL 20-JUN-2020
DEFINITION  Hepatitis B virus isolate 1745, complete genome.
ACCESSION   MT114169
VERSION     MT114169.1
KEYWORDS    .
SOURCE      Hepatitis B virus
  ORGANISM  Hepatitis B virus
            Viruses; Riboviria; Pararnavirae; Artverviricota; Revtraviricetes;
            Blubervirales; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1  (bases 1 to 3182)
  AUTHORS   McNaughton,A. and Matthews,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-2020) Nuffield Department of Medicine,
University
            of Oxford, Peter Medawar Building, South Parks Road, Oxford, UK
            OX13SY, United Kingdom
COMMENT     ##Assembly-Data-START##
            Assembly Method      :: BWA mem v. v0.7.10
            Sequencing Technology :: Illumina
            ##Assembly-Data-END##
FEATURES             Location/Qualifiers
     source          1..3182
                     /organism="Hepatitis B virus"
                     /mol_type="genomic DNA"
                     /isolate="1745"
                     /isolation_source="Human serum"
                     /host="Homo sapiens"
                     /db_xref="taxon:10407"
                     /country="United Kingdom"
                     /collection_date="01-Sep-2018"
                     /note="genotype: D"
     CDS             join(2307..3182,1..1623)
                     /codon_start=1
                     /product="polymerase"
                     /protein_id="QKQ14804.1"
                     /translation="MPLSYQHFRRLVLLDNEAGPLEEELPRLADEGLNRRVAEDLNLG
                     NLNVSIPWTHKVGNFTGLYSSTVPVFNPHWKTPSFPNIHLHQDIIKKCEQFVGPLTVN
                     EKRRLQLIMPARFYPKVTKYLPLDKGIKPYYPEHLVNHYFHTRHYLHTLWKAGVLYKR
                     ETTHSASFCGSPYSWEQELQHGAESFHQQSSGILSRPSTGSSFQSKHRKSRLGLQSQQ
                     GHLARRQQGRSGSIRAGTHPTARRPFGVEPSGSGHTTNFANKSASCFYQSPVRKAAYS
```

SDSTFEKHSSSGHAVELHNLPSNSARPQSERPVFPCWWLQFRDSKPCSDYCISHIVNL
LQDWGPCDEHGEHHIRTPRTPARVTGGVFLVDKNPHNTSESRLVVDFSQFSRGNYRVS
WPKFAVPNLQSLTNLLSSNLSWLSLDVSAAFYHLPLHPAAMPHLLVGSSGLSRYVARL
SSTSRIFNHQHGTMQDLHDSCSRNLYVSLLLLYKTFGRKLHLYSHPIILGFRKIPMGV
GLSPFLLAQFTSAICSVVRRAFPHCLAFSYMDDVVLGAKNVQHLESLYTAVTNFLMSL
GIHLNPNKTKRWGYSLHFMGYVIGCQGSLPQDHIIQKIKDCFRKLPVNRPIDWKVCQR
IVGLLGFAAPFTQCGYPALKPLYACIQSKQAFTFSPTYKAFLYKQYLNLYPVARQRPG
LCQVFADATPTGWGLVMGHQRMRGTFVAPLPIHTAELLAACFARSRSGANIIGTDNSV
VLSRKYTSFPWLLGCAANWILRGTSFVYVPSALNPADDPSRGRLGLYRPLLRLPFQPT
TGRTSLYADSPSVPSHLPDRVHFASPLHVEWRPP"

CDS join(2848..3182,1..835)
/codon_start=1
/product="large S protein"
/protein_id="QKQ14805.1"
/translation="MGQNLSTSNPLGFFPDHQLDPAFRANTANPDWDFNPNKDTWPDA
NKVGAGAFGLGLTPPHGGLLGWSPQAQGILQTLPTNPPPASTNRQSGRQPTPLTPPLR
NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSHISSIFSRIGDL
ATNMENITSGLLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQS
PTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDCRGMLPVCPLLPGS
STTSTGPCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFS
WLSLLVPFVQWFVGLSPTAWLSVIWMMWYWGPRMYNTLSPFIPLLPIFLCLWVYI"

CDS 1374..1838
/codon_start=1
/product="X protein"
/protein_id="QKQ14802.1"
/translation="MAARLCCQLDPARDVVCLRPVGAESRGRPVSGPLGTLSSPSPSA
VSTNHGAHLSLRGLPVCAFSSAGPCALRFTSARRMETTVNAHQFLPKVLHKRTLGLSV
MSTTDLEAYFKDCLFKDWEESGEELRLKVFVLGGCRHKLVCAPAPFNFFTSA"

CDS 1901..2452
/codon_start=1
/product="core protein"
/protein_id="QKQ14803.1"
/translation="MDIDPYKEFGASVELLSFLPSDFFPSVRDLLDTAAALYRDALES
PEHCSPHHTALRQAILCWGELMTLATWVGGNLEDPTSRDLVVTYVNTNMGLKFRQLLW
FHISCLTFGRETVLEYLVSFGVWIRTPPAYRPSNAPILSTLPETCVVRQRGRSTRRRT
PSPRRRRSQSPRRRRSQSRESQC"

ORIGIN

```
      1 ctccacaacc ttccatcaaa ctctgcaaga ccccagagtg agaggcctgt atttccctgc
     61 tggtggctcc agttcaggga cagtaaaccc tgttccgact actgtatctc acatatcgtc
    121 aatcttctcc aggattgggg accttgcgac gaacatggag aacatcacat caggactcct
    181 aggacccctg ctcgtgttac aggcggggtt tttcttgttg acaagaatcc tcacaatacc
    241 tcagagtcta gactcgtggt ggacttctct caattttcta gggggaacta ccgtgtgtct
    301 tggccaaaat tcgcagtccc caacctccaa tcactcacca acctcctgtc ctccaacttg
    361 tcctggttat cgctggatgt gtctgcggcg ttttatcatc ttcctcttca tcctgctgct
```

```
 421 atgcctcatc ttcttgttgg ttcttctgga ctgtcgaggt atgttgcccg tttgtcctct
 481 acttccagga tcttcaacca ccagcacggg accatgcagg acctgcacga ctcctgctca
 541 aggaacctct atgtatccct cctgttgctg tacaaaacct tcggacggaa attgcacctg
 601 tattcccatc ccatcatcct gggctttcgg aaaattccta tgggagtggg cctcagcccg
 661 tttctcctgg ctcagtttac tagtgccatt tgttcagtgg ttcgtagggc tttcccccac
 721 tgcttggctt tcagttatat ggatgatgtg gtattggggg ccaagaatgt acaacacctt
 781 gagtcccttt ataccgctgt taccaatttt cttatgtctt tgggtataca tttaaaccct
 841 aacaaaacaa aaagatgggg gtactcttta catttcatgg gctatgtcat tggatgtcag
 901 ggttcattgc cacaagatca catcatacag aaaatcaaag attgttttag gaaacttcct
 961 gtcaacaggc ctattgattg gaaagtctgt caacgtattg tgggtctttt gggttttgct
1021 gcccccttta cgcaatgtgg ttatcctgct ttaaagcctt tgtacgcctg tattcaatct
1081 aagcaggctt tcactttctc gccaacttac aaggcctttc tgtataaaca atacctgaac
1141 ctttaccccg ttgcccggca acggccaggt ctgtgccaag tgtttgctga cgcaaccccc
1201 actggctggg gcttggtcat gggccatcag cgcatgcgtg gaacctttgt ggctcctctg
1261 ccgatccata ctgcggaact cctagccgct tgctttgctc gcagcaggtc tggagcaaac
1321 attatcggga cagataactc tgtcgttctc tcccgcaaat atacatcgtt tccatggctg
1381 ctaggctgtg ctgccaactg gatcctgcgc gggacgtcgt ttgtctacgt cccgtcggcg
1441 ctgaatcccg cggacgaccc gtctcggggc cgcttgggac tctatcgtcc ccttctccgt
1501 ctgccgtttc aaccaaccac ggggcgcacc tctctttacg cggactcccc gtctgtgcct
1561 tctcatctgc cggaccgtgt gcacttcgct tcacctctgc acgtcgaatg gagaccaccg
1621 tgaacgccca ccaattcttg cccaaggtct tacataagag gactcttgga ctctctgtaa
1681 tgtcaacgac cgaccttgag gcatacttca aagactgttt gtttaaagac tgggaggagt
1741 cgggggagga gcttagatta aaggtgtttg tattaggagg ctgtaggcat aaattggtct
1801 gcgcaccagc accattcaac tttttcacct ctgcctaatc atctcttgtt catgtcctac
1861 tgttcaagcc tccaagctgt gccttgggtg gctttggggc atggacatcg atccttataa
1921 agaatttgga gcttctgtgg agttactctc ttttttgcct tctgacttct ttccttccgt
1981 acgagatctt ctagataccg ccgcagctct gtatcgggat gccttagagt ctcctgaaca
2041 ttgttcgcct caccatactg cactcaggca agcaattctt tgctgggggg aactaatgac
2101 tctagccacc tgggtgggtg gtaatttgga agatccaaca tccagggacc tagtagtcac
2161 ttatgttaac actaatatgg gcctaaagtt caggcaacta ttgtggtttc acatttcttg
2221 tctcactttt ggaagagaaa cggtcttaga gtatttggtg tcttttggag tgtggattcg
2281 cactcctcca gcttatagac catcaaatgc ccctatctta tcaacacttc cggagacttg
2341 tgttgttaga caacgaggca ggtccactag aagaagaact ccctcgcctc gcagacgaag
2401 gtctcaatcg ccgcgtcgca gaagatctca atctcgggaa tctcaatgtt agtattcctt
2461 ggactcataa ggtgggaaac tttacggggc tttactcttc tactgttcct gtctttaacc
2521 ctcattggaa aacaccctct tttcctaata tccatttaca tcaagacatt atcaaaaaat
2581 gtgaacaatt tgtaggccca ctcacagtca atgagaaaag aagactacaa ttgattatgc
2641 ctgctaggtt ctatccaaaa gttaccaaat atttgccatt ggataagggc attaaacctt
2701 attatccaga acatctagtt aatcattact tccacaccag acattattta cacactctat
2761 ggaaggcggg tgtattatat aagagagaaa caacacatag cgcttcattt tgtgggtcac
2821 catattcttg ggaacaagag ctacagcatg gggcagaatc tttccaccag caatcctctg
2881 ggattctttc ccgaccatca actggatcca gctttcagag caaacaccgc aaatccagat
2941 tgggacttca atcccaacaa ggacacctgg ccagacgcca acaaggtagg agcgggagca
3001 ttcgggctgg gactcacccc accgcacgga ggccttttgg ggtggagccc tcaggctcag
```

3061 ggcatactac aaactttgcc aacaaatccg cctcctgctt ctaccaatcg ccagtcagga

3121 aggcagccta ctcctctgac tccacctttg agaaacactc atcctcaggc catgcagtgg

3181 aa