Human adenovirus B1, complete genome

NCBI Reference Sequence: NC_011203.1

FASTA Graphics

LOCUS NC_011203 35343 bp DNA linear VRL 20-DEC-2020
DEFINITION Human adenovirus B1, complete genome.
ACCESSION NC_011203
VERSION NC_011203.1
DBLINK BioProject: PRJNA485481
KEYWORDS RefSeq.
SOURCE Human mastadenovirus B (HAdV-B)
ORGANISM Human mastadenovirus B
Viruses; Varidnaviria; Bamfordvirae; Preplasmiviricota;
Tectiliviricetes; Rowavirales; Adenoviridae; Mastadenovirus.
REFERENCE 1 (bases 1 to 35343)
AUTHORS Sirena,D., Ruzsics,Z., Schaffner,W., Greber,U.F. and Hemmi,S.
TITLE The nucleotide sequence and a first generation gene transfer vector of species B human adenovirus serotype 3
JOURNAL Virology 343 (2), 283-298 (2005)
PUBMED 16169033
REFERENCE 2 (bases 1 to 35343)
CONSRTM NCBI Genome Project
TITLE Direct Submission
JOURNAL Submitted (04-SEP-2008) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA
REFERENCE 3 (bases 1 to 35343)
AUTHORS Hemmi,S., Greber,U.F., Schaffner,W. and Sirena,D.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2005) Institute of Molecular Biology, University of Zurich, Winterthurerstrasse 190, Zurich 8057, Switzerland
COMMENT REVIEWED REFSEQ: This record has been curated by Dr Andrew Davison (MRC Virology Unit, Glasgow, UK) and Dr Balazs Harrach (Veterinary Medical Research Institute, Budapest, Hungary). The reference sequence was derived from DQ086466.
CURATION:
Proteins: A standard nomenclature has been applied so that orthologs have the same name. Functional information may have been propagated from other adenovirus species.
Protein-coding regions: The initiation codon for each CDS is assigned with as much confidence as possible. CDSs presumably inherited from the last common ancestor of all adenoviruses are indicated as genus-common, and all other CDSs as genus-specific. CDSs that have paralogues in one or more genomes are listed as members of named families. The evolutionary history of E3 and E4 is complex, frequently involving generation or loss of paralogues in

whole or in part, accompanied by rapid sequence divergence. In E3, this is evident in a set of tandemly arranged CDSs encoding transmembrane proteins. All of these are listed as members of the MP family, regardless of whether sequence conservation is detectable.

Transcripts: In mastadenoviruses, transcription arises from the early genes (E1A, E1B, E2A, E2B, E3 and E4), the intermediate genes (IX and IVa2) and the late genes (L, U and E2A-L). Certain transcriptional features are supported by experimental data for some of these genes in a few adenoviruses (particularly human adenovirus C). All transcriptional features not marked as experimentally supported are predictions made on the basis of conservation and relevance to CDS expression. The degree of confidence in each prediction depends on the feature under consideration. In primate adenoviruses, predictions for E3, E4 and the 5'-regions of E2A and E2B are generally the least certain. In non-primate mastadenoviruses, this trend is accentuated to the degree that predictions for introns in E3 and E4 are not provided, and those for the 5'-regions of E2A, E2B and IVa2 and for the intron separating the protein-coding exons of 33K are particularly tentative. In human adenovirus C, rightward-oriented transcripts from the late genes (L1-L5) are polyadenylated close downstream from polyadenylation signals at five locations (the 3'-ends of the pIIIa, pX, protease, control protein E3 12.5K and fiber CDSs), and those from the E3 genes (E3A and E3B) at two locations (the 3'-ends of the membrane protein E3 CR1-beta and control protein E3 14.7K CDSs). Only the polyadenylation signals for L3, E3B and L5 are conserved among all mastadenoviruses, whereas those for L1 and L2 are conserved in most mastadenoviruses and those for L4 and E3A in a few. In addition, potential additional polyadenylation signals exist at the 3'-ends of the penton base, pVII, V, pVI and fiber-2 CDSs in some mastadenoviruses, particularly those with non-primate hosts; the corresponding genes are listed as L1A, L1B, L1C, L2A and L5A, respectively. The gene corresponding to E3B is termed E3 in genomes lacking the E3A polyadenylation signal. In some instances where a TATA signal is not convincingly evident (particularly for E2A/E2B), a tentative assignment is made and marked as nominal.
COMPLETENESS: full length.

FEATURES Location/Qualifiers
source 1..35343
/organism="Human mastadenovirus B"
/mol_type="genomic DNA"
/strain="GB"
/serotype="Human adenovirus 3"

/host="Homo sapiens"
/db_xref="taxon:108098"
/type="1"
/acronym="HAdV-3"
/acronym="HAdV-B1"
repeat_region 1..136
/standard_name="ITR"
/rpt_type=inverted
gene 480..1499
/gene="E1A"
/locus_tag="HAdVB1_gp01"
/db_xref="GeneID:6870531"
regulatory 480..485
/regulatory_class="TATA_box"
/gene="E1A"
/locus_tag="HAdVB1_gp01"
CDS join(576..1155,1250..1455)
/gene="E1A"
/locus_tag="HAdVB1_gp01"
/function="gene regulation"
/function="cell cycle regulation"
/function="transformation"
/note="alternative splicing utilizing nonconserved sites
may produce additional proteins; genus-specific"
/codon_start=1
/product="control protein E1A"
/protein_id="YP_002213765.1"
/db_xref="GeneID:6870531"
/translation="MRHLRFLPQEVISSETGIEILEFVVNTLMGDDPEPPVQPFDPPT
LHDLYDLEIDGPEDPNEEAVNGFFTDSMLLAADEGLDINPPPETLVTPGVVVESGIGG
KKLPDLGAAEMDLRCYEEGFPPSDDEDGETEQSIHTAVNEGVKAASDVFKLDCPELPG
HGCKSCEFHRNNTGMKELLCSLCYMRMHCHFIYSPVSDDESPSPDSTTSPPEIQAPAP
ANVCKPIPVKPKPGKRPAVDKLEDLLEGGDGPLDLSTRKLPRQ"
intron 1156..1249
/gene="E1A"
/locus_tag="HAdVB1_gp01"
regulatory 1494..1499
/regulatory_class="polyA_signal_sequence"
/gene="E1A"
/locus_tag="HAdVB1_gp01"
gene 1549..3918
/gene="E1B"
/locus_tag="HAdVB1_gp02"
/db_xref="GeneID:6870515"

```
regulatory     1549..1554
               /regulatory_class="TATA_box"
               /gene="E1B"
               /locus_tag="HAdVB1_gp02"
CDS            1603..2139
               /gene="E1B"
               /locus_tag="HAdVB1_gp02"
               /function="apoptosis"
               /function="transformation"
               /note="genus-specific"
               /codon_start=1
               /product="control protein E1B 19K"
               /protein_id="YP_002213766.1"
               /db_xref="GeneID:6870515"
               /translation="MEVWAILEDLRQTKLLLENASDGVSGLWRFWFGGDLARLVFRIK
               QDYREEFEKLLDDSPGLFEALNLGHQAHFKEKVLSVLDFSTPGRTAAAVAFLTFILDK
               WIRQTHFSKGYVLDFIAAALWRTWKARRMRTILDYWPVQPLGVAGILRHPPTMPAVLQ
               EEQQEDNPRAGLDPPVEE"
CDS            1908..3386
               /gene="E1B"
               /locus_tag="HAdVB1_gp02"
               /function="apoptosis"
               /function="DNA damage response"
               /function="RNA metabolism and transport"
               /function="transformation"
               /note="alternative splicing utilizing nonconserved sites
               may produce additional proteins; genus-specific"
               /codon_start=1
               /product="control protein E1B 55K"
               /protein_id="YP_002213767.1"
               /db_xref="GeneID:6870515"
               /translation="MDPPNSLQQGIRFGFHSSSFVENMEGSQDEDNLRLLASAASGSS
               RDTETPTDHASGSAGGAAGGQSESRPGPSGGGVADLFPELRRVLTRSTTSGQNRGIKR
               ERNPSGNNSRTELALSLMSRRRPETVWWHEVQSEGRDEVSILQEKYSLEQLKTCWLEP
               EDDWEVAIRNYAKISLRPDKQYRITKKINIRNACYISGNGAEVIIDTQDKAVFRCCMM
               GMWPGVVGMEAVTLMNIRFKGDGYNGIVFMANTKLILHGCSFFGFNNTCVEAWGQVSV
               RGCSFYACWIATSGRVKSQLSVKKCMFERCNLGILNEGEARVRHCAATETGCFILIKG
               NASVKHNMICGHSDERPYQMLTCAGGHCNILATVHIVSHARKKWPVFEHNVITKCTMH
               IGGRRGMFMPYQCNMNHVKVMLEPDAFSRVSLTGIFDMNIQLWKILRYDDTKPRVRAC
               ECGGKHARFQPVCVDVTEDLRPDHLVLACTGAEFGSSGEETD"
gene           3419..3918
               /gene="IX"
               /locus_tag="HAdVB1_gp03"
               /db_xref="GeneID:6870516"
```

```
regulatory     3419..3424
               /regulatory_class="TATA_box"
               /gene="IX"
               /locus_tag="HAdVB1_gp03"
CDS            3480..3896
               /gene="IX"
               /locus_tag="HAdVB1_gp03"
               /function="capsid morphogenesis"
               /note="genus-specific"
               /codon_start=1
               /product="capsid protein IX"
               /protein_id="YP_002213768.1"
               /db_xref="GeneID:6870516"
               /translation="MSGSASFEGGVFSPYLTGRLPPWAGVRQNVMGSTVDGRPVQPAN
               SSTLTYATLSSSPLDAAAAAAATAAANTILGMGYYGSIVANSSSSNNPSTLAEDKLLV
               LLAQLEALTQRLGELSKQVAQLREQTESAVATAKSK"
regulatory     3892..3897
               /regulatory_class="polyA_signal_sequence"
               /note="E1B, IX"
regulatory     3909..3914
               /regulatory_class="polyA_signal_sequence"
               /note="E1B, IX"
regulatory     3913..3918
               /regulatory_class="polyA_signal_sequence"
               /note="E1B, IX"
gene           complement(3947..26643)
               /gene="E2B"
               /locus_tag="HAdVB1_gp04"
               /db_xref="GeneID:6870517"
gene           complement(3947..5572)
               /gene="IVa2"
               /locus_tag="HAdVB1_gp05"
               /db_xref="GeneID:6870518"
regulatory     complement(3947..3952)
               /regulatory_class="polyA_signal_sequence"
               /note="IVa2, E2B"
CDS            complement(join(3948..5281,5560..5572))
               /gene="IVa2"
               /locus_tag="HAdVB1_gp05"
               /function="gene regulation"
               /function="DNA encapsidation"
               /function="capsid morphogenesis"
               /note="putative ATPase; genus-common"
               /codon_start=1
```

/product="encapsidation protein IVa2"
/protein_id="YP_002213830.1"
/db_xref="GeneID:6870518"
/translation="METKGRQSAAVLDQQDESEAHPRKRPTRRAPLHRDGNNPDSDAA
TMEGSDPGSAGRPSSDSLLQEPSQPAKRGGLLDRDAIEHITELWDRLELLQQTLSKMP
MADGLKPLKNFSSLQELLSLGGERLLTDLVRENIHVREMMNEVAPLLREDGSCRSLNY
HLQPVIGVIYGPTGCGKSQLLRNLLSSQLITPAPETVFFIAPQVDMIPPSELKAWEMQ
ICEGNYAPGPEGTFIPQSGTLRPKFIKMAYDDLTQEHNYDVSDPRNVFARAAAHGPIA
IIMDECMENLGGHKGVSKFFHAFPSKLHDKFPKCTGYTVLVVLHNMNPRRDLGGNIAN
LKIQSKMHIISPRMHPSQLNRFVNTYTKGLPVAISLLLKDIVQHHALRPCYDWVIYNT
TPEQEALQWSYLHPRDGLMPMYLNIQSHLYRVLEKIHRVLNDRDRWSRAYRARKIK"

CDS complement(join(5051..8623,13848..13856))
/gene="E2B"
/locus_tag="HAdVB1_gp04"
/EC_number="2.7.7.7"
/function="DNA replication"
/note="pol; genus-common"
/codon_start=1
/product="DNA polymerase"
/protein_id="YP_002213831.1"
/db_xref="GeneID:6870517"
/translation="MALVQAHGAGSFHPEASDPGCQPPRRRACESSQGAVPEPTRARR
RRTTASHAAGSRAAPAARCAPSTPLLNMQEATEPPPSKMKNKGTVVAPKGHGTLQAID
ISTNGPVEIKYYLNLPHALQKIMQVNLLTLPTNLTPQRLRTLDSSGLRALVLELRPCR
AEVWTCLPRGLVSMTTIETEDGHADADNIVKREVQAPGLNFPLKFLVKGSQVQLIHKV
HPVNRCEYCGRLYKHKHECSARRREFYFHHINSHSSNWWQEIQFFPIGSHPRTERLFL
TYDVETYTWMGSFGKQLIPFMLVMKLSGDQRLVNIAYDLAIKLKWDRWRQDPQTFYCV
TPEKMAVGQHFRQYRDQLQTALAVDLWSSFLKANPHVHEWALEHYALTDPTDLTFEEL
KKLPHVRGTPRFLELYIVGHNINGFDEIVLAAQVINNRAEVPQPFKITRNFMPRAGKI
LFNDVTFALPNPAYKKRVDFQLWEQGACDDIDFKYQFLKVMVRDTFALTHTSLRNAAQ
AYSLPVEKGCCPYKAVNQFYMLGSYRAEKDGFPLEEYWKDHEEYLLNRELWEKKSQPR
YDIIQETLDYCALDVLITAELVAKLQESYAHFIRDSVGLPHAHFNIFQRPTISSNSHA
IFRQIVYRAEKPNRTNLGPGLLAPSHELYDYVRASIRGGRCYPTYIGILEEPLYVYDI
CGMYASALTHPMPWGTPLNPYERALAVREWQMTLDDPATISYFDKDLLPGIFTIDADP
PDEFMLDPLPPFCSRKGGRLCWTNEPLRGEVATSVDLITLHNRGWRIRIVPDELTTIF
PEWKCVAREYVQLNITAKERADKEKNQTMRSIAKLLSNALYGSFATKLDNKKIVFSDQ
MDESLMKGISAGTVNIKSSSFLETDNLSAEVMPAFEREYLPQQLALLDSDPEDSEDEQ
RPAPFYTPPAGTPGHVAYTYKPITFLDVEEGDMCLHTVEKVDPLVDNDRYPSHVASFV
LAWTRAFVSEWAGFLYEEDRGTPLEDRPIKSVYGDTDSLFVTQRGHELMETKGKKRIK
KHGGKLVFDPDEPDLTWLVECETVCASCGADAYSPESIFLAPKLYALKCIYCPACHKT
SKGKLRAKGHAAEALNYELMVNCYLADMQGANRQRFSTSRMSLKRTLASAQPGAHPFT
VTETTLTRTLRPWKDRTLAALDAHRLIPYSRSRPNPRNEEVCWIEMP"

intron complement(5282..5559)
/note="IVa2"

gene 5872..32340
/gene="L5"
/locus_tag="HAdVB1_gp06"
/db_xref="GeneID:6870519"
gene 5872..27717
/gene="L4"
/locus_tag="HAdVB1_gp07"
/db_xref="GeneID:6870520"
gene 5872..21943
/gene="L3"
/locus_tag="HAdVB1_gp08"
/db_xref="GeneID:6870521"
gene 5872..17501
/gene="L2"
/locus_tag="HAdVB1_gp09"
/db_xref="GeneID:6870522"
gene 5872..13835
/gene="L1"
/locus_tag="HAdVB1_gp10"
/db_xref="GeneID:6870523"
regulatory 5872..5877
/regulatory_class="TATA_box"
/note="L"
intron 5944..6961
/note="between L1 and L2 leaders"
intron 7034..9476
/note="between L2 and L3 leaders"
intron 7034..7802
/gene="L1"
/locus_tag="HAdVB1_gp10"
/note="between L2 and i leaders; precedes protein 13.6K CDS"
CDS join(7829..8227,9477..9497)
/gene="L1"
/locus_tag="HAdVB1_gp10"
/function="possibly RNA metabolism and transport"
/note="i-leader protein; similar CDSs are present in some other human mastadenoviruses, but their conservation may merely be a secondary effect of conservation of DNA polymerase encoded on the opposing strand"
/codon_start=1
/product="protein 13.6K"
/protein_id="YP_002213771.1"
/db_xref="GeneID:6870523"

/translation="MRANREELDLLPPVGGMAVDVMEVELPATRRAFMLVLVQTAAVL
AAIHGMHLMNELYLTSFDEKFQWKIEAWRLYLALYYVVCIGMTIFCLDGGHADEPSRE
ASPDLGAAGAELEDESAQAGAVQGPETLRSQVLRART"

CDS complement(join(8422..10389,13848..13856))
/gene="E2B"
/locus_tag="HAdVB1_gp04"
/function="DNA replication"
/note="covalently attached to genome 5'-termini;
genus-common"
/codon_start=1
/product="terminal protein precursor pTP"
/protein_id="YP_002213832.1"
/db_xref="GeneID:6870517"
/translation="MALSVNDCARLTGQTVPTMNYFMPLRNIWNRVREFPRASTTASG
ITWMSRYIYGYHRLMLEDLAPGAPATERWPLYRQPPPHFLIGYQYLVRTCNDYIFDTR
AYSRLKYTELVRPGHQTVNWSVMANCSYTINTGAYHRFVDFDDFQATLTQVQQAILAE
RVVADLALVQPMRGFGITRMHGRAGEEEVPVERLMQDYYKDLARCQDNAWGMAHRLRI
QQAGPKDLVLLATIRRLKTAYFNFITSSIVSPSQEEEGEERENPDRASSRPRPQETVL
SLPCDCDWLDAFVERFSDPVDLETIRSLRGVPTGQLIKCIISAVSLPNGEPPSHHFRE
MRGGVFTLRPRENGRAVTETMRRRRGEVIERFIDRLPVRRRRRRVPPPPAAPPEEEEM
LVEEEEIEEEIPGAFEREVRTTIAELIRLLEEELTVSARNSQFFNFAVDFYEAMERLE
ALGDVSEMPLRRWIMYFFVTEHIATTLNYLFQRLCNYAVFTRHVELNLAQVVMRARDP
VGAVVYSRVWNEAGMNAFSQLIGRISNDLAATVERAGRGDLQEEEIEQFMAEIAYQDN
SGDVQEILRQAAVNDTEIDSVELSFRFKLTGPVAFTQRRQIQDVNRRVVAHASLLRAQ
YQNLPARGADVPLPAMPPGPEPPLPPGARPRHRF"

intron complement(8624..13847)
/gene="E2B"
/locus_tag="HAdVB1_gp04"

intron 9564..31367
/gene="L5"
/locus_tag="HAdVB1_gp06"
/note="precedes fiber CDS"

intron 9564..26549
/gene="L4"
/locus_tag="HAdVB1_gp07"
/note="precedes capsid protein precursor pVIII CDS"

intron 9564..25774
/gene="L4"
/locus_tag="HAdVB1_gp07"
/note="precedes protein 33K and encapsidation protein 22K
CDSs"

intron 9564..23570
/gene="L4"
/locus_tag="HAdVB1_gp07"

```
                /note="precedes hexon assembly protein 100K CDS"
intron          9564..21274
                /gene="L3"
                /locus_tag="HAdVB1_gp08"
                /note="precedes protease CDS"
intron          9564..18369
                /gene="L3"
                /locus_tag="HAdVB1_gp08"
                /note="precedes hexon CDS"
intron          9564..17550
                /gene="L3"
                /locus_tag="HAdVB1_gp08"
                /note="precedes capsid protein precursor pVI CDS"
intron          9564..17247
                /gene="L2"
                /locus_tag="HAdVB1_gp09"
                /note="precedes core protein precursor pX CDS"
intron          9564..16163
                /gene="L2"
                /locus_tag="HAdVB1_gp09"
                /note="precedes core protein V CDS"
intron          9564..15526
                /gene="L2"
                /locus_tag="HAdVB1_gp09"
                /note="precedes core protein precursor pVII CDS"
intron          9564..13900
                /gene="L2"
                /locus_tag="HAdVB1_gp09"
                /note="precedes penton base CDS"
intron          9564..12050
                /gene="L1"
                /locus_tag="HAdVB1_gp10"
                /note="precedes capsid protein precursor pIIIa CDS"
intron          9564..10868
                /gene="L1"
                /locus_tag="HAdVB1_gp10"
                /note="precedes encapsidation protein 52K CDS"
intron          complement(10390..13847)
                /gene="E2B"
                /locus_tag="HAdVB1_gp04"
gene            10422..10592
                /gene="VAI"
                /locus_tag="HAdVB1_gs01"
                /db_xref="GeneID:6870524"
```

misc_RNA 10422..10592
/gene="VAI"
/locus_tag="HAdVB1_gs01"
/product="virus-associated RNA I"
/db_xref="GeneID:6870524"
gene 10666..10838
/gene="VAII"
/locus_tag="HAdVB1_gs02"
/db_xref="GeneID:6870525"
misc_RNA 10666..10838
/gene="VAII"
/locus_tag="HAdVB1_gs02"
/product="virus-associated RNA II"
/db_xref="GeneID:6870525"
CDS 10869..12026
/gene="L1"
/locus_tag="HAdVB1_gp10"
/function="DNA encapsidation"
/function="capsid morphogenesis"
/note="genus-common"
/codon_start=1
/product="encapsidation protein 52K"
/protein_id="YP_002213772.1"
/db_xref="GeneID:6870523"
/translation="MHPVLRQMRPQQQAPSQQQPQKALLAPATTAAAAVSGAGQPAYD
LDLEEGEGLARLGAPSPERHPRVQLKKDSREAYVPQQNLFRDRSGEEPEEMRASRFNA
GRELRHGLDRRRVLQDEDFEVDEVTGISPARAHVAAANLVSAYEQTVKEERNFQKSFN
NHVRTLIAREEVTLGLMHLWDLMEAIAQNPTSKPLTAQLFLVVQHSRDNEAFREALLN
ITEPDGRWLYDLINILQSIIVQERSLGLAEKVAAINYSVLSLGKYYARKIYKTPYVPI
DKEVKIDGFYMRMTLKVLTLSDDLGVYRNDRMHRAVSASRRRELSDRELMHSLQRALT
GAGTDGENYFDMGADLQWQPSRRAMEAAGCELPYIEEVDEVEDEEGEYLED"
CDS 12051..13817
/gene="L1"
/locus_tag="HAdVB1_gp10"
/function="capsid morphogenesis"
/note="genus-common"
/codon_start=1
/product="capsid protein precursor pIIIa"
/protein_id="YP_002213773.1"
/db_xref="GeneID:6870523"
/translation="MEQQQAPDPAMRAALQSQPSGINSSDDWTQAMQRIMALTTRNPE
AFRQQPQANRLSAILEAVVPSRSNPTHEKVLAIVNALVENKAIRPDEAGLVYNALLER
VARYNSSNVQTNLDRMVTDVREAVSQRERFQRDANLGSLVALNAFLSTQPANVPRGQQ
DYTNFLSALRLMVAEVPQSEVYQSGPDYFFQTSRQGLQTVNLTQAFKNLKGLWGVHAP

VGDRATVSSLLTPNSRLLLLLVSPFTDSGSIDRNSYLGYLLNLYREAIGQSQVDEQTY
QEITQVSRALGQEDTGSLEATLNFLLTNRSQKIPPQYALTAEEERILRYVQQSVGLFL
MQEGATPTAALDMTARNMEPSMYASNRPFINKLLDYLHRAAAMNSDYFTNAILNPHWL
PPPGFYTGEYDMPDPNDGFLWDDVDSSIFSPPPGYTVWKKEGGDRRHSSVSLSGTAGA
AATVPEAASPFPSLPFSLNSVRSSELGRITRPRLMGEDEYLNDSLLRPEREKNFPNNG
IESLVDKMSRWKTYAQDHRDEPRILGATSRPIRRRQRHDRQMGLVWDDEDSADDSSVL
DLGGRGGGNPFAHLRPRFGRML"

regulatory 13830..13835
/regulatory_class="polyA_signal_sequence"
/gene="L1"
/locus_tag="HAdVB1_gp10"

CDS 13905..15539
/gene="L2"
/locus_tag="HAdVB1_gp09"
/function="capsid morphogenesis"
/function="cell attachment"
/function="cell entry"
/note="capsid protein III; genus-common"
/codon_start=1
/product="penton base"
/protein_id="YP_002213774.1"
/db_xref="GeneID:6870522"
/translation="MRRRAVLGGAVVYPEGPPPSYESVMQQQAAMIQPPLEAPFVPPR
YLAPTEGRNSIRYSELSPLYDTTKLYLVDNKSADIASLNYQNDHSNFLTTVVQNNDFT
PTEASTQTINFDERSRWGGQLKTIMHTNMPNVNEYMFSNKFKARVMVSRKAPEGVTVN
DTYDHKEDILKYEWFEFILPEGNFSATMTIDLMNNAIIDNYLEIGRQNGVLESDIGVK
FDTRNFRLGWDPETKLIMPGVYTYEAFHPDIVLLPGCGVDFTESRLSNLLGIRKRHPF
QEGFKIMYEDLEGGNIPALLDVTAYEESKKDTTTETTTLAVAEETSEDDDITRGDTYI
TEKQKREAAAAEVKKELKIQPLEKDSKSRSYNVLEDKINTAYRSWYLSYNYGNPEKGI
RSWTLLTTSDVTCGAEQVYWSLPDMMQDPVTFRSTRQVNNYPVVGAELMPVFSKSFYN
EQAVYSQQLRQATSLTHVFNRFPENQILIRPPAPTITTVSENVPALTDHGTLPLRSSI
RGVQRVTVTDARRRTCPYVYKALGIVAPRVLSSRTF"

intron complement(14047..26546)
/gene="E2B"
/locus_tag="HAdVB1_gp04"
/note="precedes DNA polymerase and terminal protein
precursor pTP CDSs"

CDS 15551..16129
/gene="L2"
/locus_tag="HAdVB1_gp09"
/function="capsid morphogenesis"
/note="genus-common"
/codon_start=1
/product="core protein precursor pVII"

/protein_id="YP_002213775.1"
/db_xref="GeneID:6870522"
/translation="MSILISPSNNTGWGLYAPTKMYGGARKRSTQHPVRVRGHFRAPW
GALKGRTRTRTTVDDVIDQVVADARNYTPTAPTSTVDAVIDSVVADARAYARRKSRRR
RIARRHRATPAMRAAKAILRRAKRVGRRAMLRAARRAASGASSGRSRRRAATAAAAAI
ANMAQPRRGNVYWVRDATTGQRVPVRTRPPRT"

CDS 16172..17221
/gene="L2"
/locus_tag="HAdVB1_gp09"
/function="capsid morphogenesis"
/note="genus-specific"
/codon_start=1
/product="core protein V"
/protein_id="YP_002213776.1"
/db_xref="GeneID:6870522"
/translation="MSKRKYKEEMLQVIAPEIYGPPVKDEKKPRKIKRVKNNKKEEDD
DDGLVEFVREFAPRRRVQWRGRKVRQVLRPGTTVVFTPGERSSTTFKRSYDEVYGDDD
ILEQAADRLDEFAYGKRTRSSPKEEAVSIPLDHGNPTPSLKPVTLQQVLPVPARRGVK
REGEDLYPTMQLMVPKRQRLEDVLEKMKVDADIQPEVKVRPIKEVAPGLGVQTFDIKI
PTESMEVQTEPAKPTTSIEVQTEPWMPAPVAAPSTTRRSRRKYGPASLLMPNYALHPS
IIPTPGYRGTRYYRNRSNTSRRRKPPASRTRRRRRRNTASKLTPAALVRRVYRDGRAE
PLTLPRARYHPSITT"

CDS 17250..17477
/gene="L2"
/locus_tag="HAdVB1_gp09"
/function="capsid morphogenesis"
/note="putative type 2 membrane protein; product Mu;
genus-common"
/codon_start=1
/product="core protein precursor pX"
/protein_id="YP_002213777.1"
/db_xref="GeneID:6870522"
/translation="MALTCRLRVPITGYRGRNSRRRRMLGRGMRRHRRRRAISRRLGG
GFLPALIPIIAAAIGAIPGIASVAVQASQRH"

regulatory 17496..17501
/regulatory_class="polyA_signal_sequence"
/gene="L2"
/locus_tag="HAdVB1_gp09"

CDS 17553..18305
/gene="L3"
/locus_tag="HAdVB1_gp08"
/function="capsid morphogenesis"
/note="cleaved peptide stimulates protease; genus-common"
/codon_start=1

/product="capsid protein precursor pVI"
/protein_id="YP_002213778.1"
/db_xref="GeneID:6870521"
/translation="MEDINFSSLAPRHGTRPYMGTWSDIGTSQLNGGAFNWSSIWSGL
KNFGSTVKTYGNKAWNSSTGQALRNKLKEQNFQQKVVDGIASGINGVVDLANQAVQKQ
INSRLDPPPATPGEMEVEEELPPLEKRGDKRPRPELEQTLVTRADEPPSYEEAVKLGM
PTTRPVAHMATGVMKPSQSHRPATLDLPPPPASAAPVPKPVATRKPTAVQPVAVARPR
PGGTPRPKANWQSTLNSIVGLGVQSVKRRRCY"
CDS 18418..21252
/gene="L3"
/locus_tag="HAdVB1_gp08"
/function="capsid morphogenesis"
/note="capsid protein II; genus-common"
/codon_start=1
/product="hexon"
/protein_id="YP_002213779.1"
/db_xref="GeneID:6870521"
/translation="MATPSMMPQWAYMHIAGQDASEYLSPGLVQFARATDTYFSMGNK
FRNPTVAPTHDVTTDRSQRLMLRFVPVDREDNTYSYKVRYTLAVGDNRVLDMASTFFD
IRGVLDRGPSFKPYSGTAYNSLAPKGAPNTSQWIVTTNGDNAVTTTTNTFGIASMKGD
NITKEGLQIGKDITTTEGEEKPIYADKTYQPEPQVGEESWTDTDGTNEKFGGRALKPA
TNMKPCYGSFARPTNIKGGQAKNRKVKPTTEGGVETEEPDIDMEFFDGRDAVAGALAP
EIVLYTENVNLETPDSHVVYKPETSNNSHANLGQQAMPNRPNYIGFRDNFVGLMYYNS
TGNMGVLAGQASQLNAVVDLQDRNTELSYQLLLDSLGDRTRYFSMWNQAVDSYDPDVR
IIENHGIEDELPNYCFPLNGIGPGHTYQGIKVKTDDTNGWEKDANVAPANEITIGNNL
AMEINIQANLWRSFLYSNVALYLPDVYKYTPPNITLPTNTNTYEYMNGRVVSPSLVDS
YINIGARWSLDPMDNVNPFNHHRNAGLRYRSMLLGNGRYVPFHIQVPQKFFAVKNLLL
LPGSYTYEWNFRKDVNMVLQSSLGNDLRTDGATISFTSINLYATFFPMAHNTASTLEA
MLRNDTNDQSFNDYLSAANMLYPIPANATNIPISIPSRNWAAFRGWSFTRLKTKETPS
LGSGFDPYFVYSGSIPYLDGTFYLNHTFKKVSIMFDSSVSWPGNDRLLSPNEFEIKRT
VDGEGYNVAQCNMTKDWFLVQMLANYNIGYQGFYIPEGYKDRMYSFFRNFQPMSRQVV
DEVNYTDYKAVTLPYQHNNSGFVGYLAPTMRQGEPYPANYPYPLIGTTAVKSVTQKKF
LCDRTMWRIPFSSNFMSMGALTDLGQNMLYANSAHALDMTFEVDPMDEPTLLYLLFEV
FDVVRVHQPHRGVIEAVYLRTPFSAGNATT"
CDS 21289..21918
/gene="L3"
/locus_tag="HAdVB1_gp08"
/EC_number="3.4.22.39"
/function="capsid morphogenesis"
/note="capsid protein 23K; genus-common"
/codon_start=1
/product="protease"
/protein_id="YP_002213780.1"
/db_xref="GeneID:6870521"

/translation="MTCGSGNGSSEQELKAIVRDLGCGPYFLGTFDKRFPGFMAPDKL
ACAIVNTAGRETGGEHWLAFGWNPRYNTCYLFDPFGFSDERLKQIYQFEYEGLLRRSA
LATKDRCITLEKSTQSVQGPRSAACGLFCCMFLHAFVHWPDRPMDGNPTMKLLTGVSN
SMLQSPQVQPTLRRNQEVLYRFLNTHSSYFRSHRARIERATAFDRMDMQ"

regulatory 21938..21943
/regulatory_class="polyA_signal_sequence"
/gene="L3"
/locus_tag="HAdVB1_gp08"

gene complement(21950..26643)
/gene="E2A"
/locus_tag="HAdVB1_gp11"
/db_xref="GeneID:6870526"

gene complement(21950..25561)
/gene="E2A-L"
/locus_tag="HAdVB1_gp12"
/db_xref="GeneID:6870527"

regulatory complement(21950..21955)
/regulatory_class="polyA_signal_sequence"
/note="E2A, E2A-L"

CDS complement(22004..23557)
/gene="E2A"
/locus_tag="HAdVB1_gp11"
/function="DNA replication"
/note="DBP; genus-common; DBP family"
/codon_start=1
/product="single-stranded DNA-binding protein"
/protein_id="YP_002213833.1"
/db_xref="GeneID:6870526"
/translation="MASRGGNQSSDRHREHTPERGMSSASHPPPRPDRSPSQSPPPLP
PKRNTYRRVGSGSSTESQVVLVSETSQSSLSPERSNSPPPIPPKKKPRKTKHVPLQDI
SQDSEEEREQAQLVAVGFSYPPVRIIEKDGKRSVEKIDKNDPIAKGATFIAVRNPLSL
PLVSAWEKGMEVMAVLMERYRLDNDLRTSFKLMPEQHEQYKRICHQYVNEEHRGIPLT
FTSMKTLTTMMGRFMQGLVHYYSEIAHNNWECTGCALWAHGCTDYEGKVKCLHGTIMI
QKDHIIEMDVASENGQRAMKENPDRAKITQNRWGRNVVQLANNDARCCVNDANCTTNQ
FSSKSCGMFYTEGSKAQEAFKQYEAFMKAVYPGITPDQARMMLIPIHCDCNHKPGCAP
VMGRQTCKMTPFGMANAEDLDVTTISDPTVLASVRHPALIVFQCCNPVYRNSRVQNAG
PNCDFKISAPDLLGAMQLTRKLWQDTFPEIPVPKLVIPEFKWQNRLQFRNVSLPAGHS
DSRQNPFDF"

intron complement(23567..24278)
/note="E2A, E2A-L; precedes single-stranded DNA-binding
protein CDS"

CDS 23588..26074
/gene="L4"
/locus_tag="HAdVB1_gp07"

/function="capsid morphogenesis"
/function="translational regulation"
/note="genus-common"
/codon_start=1
/product="hexon assembly protein 100K"
/protein_id="YP_002213782.1"
/db_xref="GeneID:6870520"
/translation="METQPSLPTSLQAPSHLAPSSDEEESLTTPPPSPATTTSTLDDE
EEVDAAQEMQAQDNVKTEEIEADVEQDPGYVTPAEHQEELKRFLDREDDDRPEHQADG
VYQEAGIRDHVADYLTGLGGEDVLLKHLARQSIIVKDALLDLTEVPISVEELSRAYEL
NLFSPQVPPKRQPNGTCEANPRLNFYPAFTIPEVLATYHIFFKNQKIPVSCRANRTRA
DALLNLGPGARLPDIASLEEVPKIFEGLGSDETRAANALQQGENDMDEHHSALVELEG
DNARIAVLKRSIVVTHFAYPAVNLPPKVMSAVMDHLLIKRARPLSENQNMQDPDASDE
GKPVVSDEQLSRWLGTNSPRDLEERRKLMMAVVLVTVELECLRRFFTDPETLRKLEEN
LHYTFRHGFVRQACKISNVELTNLVSYMGILHENRLGQSVLHTTLKGEARRDYIRDCV
YLYLCHTWQTGMGVWQQCLEEQNLKELDKLLQRSLKALWTGFDERTVASDLADIIFPE
RLRVTLRNGLPDFMSQSMLNNFRSFILERSGILPATCCALPSDFVPLTYRECPPPLWS
HCYLFRLANYLSYHSDVIEDVSGDGLLECHCRCNLCTPHRSLACNPQLLSETQIIGTF
ELQGPRSKGEGSSPGQSLKLTPGLWTSAYLRKFHPEDYHPYEIRFYEDQSHPPKVELS
ACVITQGAILAQLQAIQKSRQEFLLKKGNGVYLDPQTGEELNTRFPQDVPSPRKQEVE
GAAVAPRGYEGRLGQSGRGGDGRLGQPGRGGGQPGGRQFGGGRRGGRGGGRSNRRQTV
VIGGGDKQVPRQQHGYHLRSGSGGPAAAQQ"

intron complement(24356..26546)
/note="E2A, E2B"

intron complement(24356..25455)
/gene="E2A-L"
/locus_tag="HAdVB1_gp12"

regulatory complement(25556..25561)
/regulatory_class="TATA_box"
/gene="E2A-L"
/locus_tag="HAdVB1_gp12"

CDS join(25776..26124,26294..26649)
/gene="L4"
/locus_tag="HAdVB1_gp07"
/function="RNA splicing"
/note="genus-common"
/codon_start=1
/product="protein 33K"
/protein_id="YP_002213783.1"
/db_xref="GeneID:6870520"
/translation="MSHRRGSKKLKVQLSPPEDMKEDWDSQAEEEMEDWDSQAEEVDS
LEEDSLEEEDEEAEEVEEATAAKQLSSAAETSKSPDSSTATISAPGRGAQRRPNSRWD
ETGRFPNPTTASKTAPTAVSKSRQPRQKNTAATTVTRKPAVRKSTTSAPGGGLRITAN
EPAQTRELRNRIFPTLYAIFQQSRGQEQELKVKNRSLRSLTRSCLYHKSEDQLQRTLE

DAEALFNKYCALTLKE"

CDS 25776..26375

/gene="L4"

/locus_tag="HAdVB1_gp07"

/function="DNA encapsidation"

/function="capsid morphogenesis"

/function="transcriptional control"

/note="genus-common"

/codon_start=1

/product="encapsidation protein 22K"

/protein_id="YP_002213784.1"

/db_xref="GeneID:6870520"

/translation="MSHRRGSKKLKVQLSPPEDMKEDWDSQAEEEMEDWDSQAEEVDS
LEEDSLEEEDEEAEEVEEATAAKQLSSAAETSKSPDSSTATISAPGRGAQRRPNSRWD
ETGRFPNPTTASKTGKKERQGYKSWRGHKNAIISCLHECGGNISFTRRYLLFHHGVNF
PRNILHYYRHLHSPYCSQQVPATPTEKYSSDNGDQKTSS"

intron 26125..26293

/gene="L4"

/locus_tag="HAdVB1_gp07"

regulatory complement(26638..26643)

/regulatory_class="TATA_box"

/gene="E2A"

/locus_tag="HAdVB1_gp11"

/note="E2A, E2B; nominal"

CDS 26720..27403

/gene="L4"

/locus_tag="HAdVB1_gp07"

/function="capsid morphogenesis"

/note="genus-common"

/codon_start=1

/product="capsid protein precursor pVIII"

/protein_id="YP_002213785.1"

/db_xref="GeneID:6870520"

/translation="MSKEIPTPYMWSYQPQMGLAAGASQDYSTRMNWLSAGPSMISRV
NDIRAYRNQLLLEQSALTTTPRQHLNPRNWPATLVYQENPAPTTVLLPRDAQAEVQMT
NAGVQLAGGSALCRHRPQQSIKRLVIRGRGIQLNDESVSSSLGLRPDGVFQIAGCGRS
SFTPRQAVLTLESSSSQPRSGGIGTLQFVEEFTPSVYFNPFSGSPGQYPDEFIPNFDA
ISESVDGYD"

gene 27085..31186

/gene="E3"

/locus_tag="HAdVB1_gp13"

/db_xref="GeneID:6870528"

regulatory 27085..27090

/regulatory_class="TATA_box"

```
                /gene="E3"
                /locus_tag="HAdVB1_gp13"
     intron      27137..30714
                /gene="E3"
                /locus_tag="HAdVB1_gp13"
                /note="precedes control protein E3 14.7K CDS"
     intron      27137..30326
                /gene="E3"
                /locus_tag="HAdVB1_gp13"
                /note="precedes membrane protein E3 RID-beta CDS"
     intron      27137..30043
                /gene="E3"
                /locus_tag="HAdVB1_gp13"
                /note="precedes membrane protein E3 RID-alpha CDS"
     intron      27137..29781
                /gene="E3"
                /locus_tag="HAdVB1_gp13"
                /note="precedes membrane glycoprotein E3 CR1-delta CDS"
     intron      27137..29201
                /gene="E3"
                /locus_tag="HAdVB1_gp13"
                /note="precedes membrane glycoprotein E3 CR1-gamma CDS"
     intron      27137..28648
                /gene="E3"
                /locus_tag="HAdVB1_gp13"
                /note="precedes membrane glycoprotein E3 CR1-beta CDS"
     intron      27137..28076
                /gene="E3"
                /locus_tag="HAdVB1_gp13"
                /note="precedes membrane glycoprotein E3 gp19K CDS"
     intron      27137..27600
                /gene="E3"
                /locus_tag="HAdVB1_gp13"
                /note="precedes membrane glycoprotein E3 CR1-alpha CDS"
     CDS         27403..27723
                /gene="E3"
                /locus_tag="HAdVB1_gp13"
                /function="unknown"
                /note="genus-specific; 12.5K family"
                /codon_start=1
                /product="control protein E3 12.5K"
                /protein_id="YP_002213786.1"
                /db_xref="GeneID:6870528"
                /translation="MSNGGAAELARLRHLDHCRRFRCFARELTEFIYFELSEEHPQGP
```

AHGVRITIEGGIDSRLHRIFSQRPVLIERDQGNTTISIYCICNHPGLHESLCCLVCAE
FNKN"

CDS 27677..28117
/gene="E3"
/locus_tag="HAdVB1_gp13"
/function="unknown"
/note="0 CR1 domains; genus-specific; MP family"
/codon_start=1
/product="membrane glycoprotein E3 CR1-alpha"
/protein_id="YP_002213787.1"
/db_xref="GeneID:6870528"
/translation="MKAFAVLFVLSLIKTELRPSYGLPLLQSGLYNTNQIFQKTQTLP
PFIQDSNSTLPAPSTTNLPETNKLASHLQHRLSRSLLSANTTTPKTGGELRGLPTDDP
WVVAGFVTLGVVAGGLVLILCYLYTPCCAYLVILCCWFKKWGPY"

regulatory 27712..27717
/regulatory_class="polyA_signal_sequence"
/gene="L4"
/locus_tag="HAdVB1_gp07"

CDS 28102..28620
/gene="E3"
/locus_tag="HAdVB1_gp13"
/function="immune regulation"
/note="binds to MHC-I and inhibits transport;
genus-specific; MP family"
/codon_start=1
/product="membrane glycoprotein E3 gp19K"
/protein_id="YP_002213788.1"
/db_xref="GeneID:6870528"
/translation="MGAILVVLALLSLLGLGSANLNPLDHDPCLDFDPENCTLTFAPD
TSRLCGVLIKCGWDCRSVEITHNNKTWNNTLSTTWEPGVPQWYTVSVRGPDGSIRISN
NTFIFSEMCDLAMFMSRQYDLWPPSKENIVAFSIAYCLVTCIITAIICVCIHLLIVIR
PRQSNEEKEKMP"

CDS 28650..29189
/gene="E3"
/locus_tag="HAdVB1_gp13"
/function="unknown"
/note="1 CR1 domain; genus-specific; MP family"
/codon_start=1
/product="membrane glycoprotein E3 CR1-beta"
/protein_id="YP_002213789.1"
/db_xref="GeneID:6870528"
/translation="MASVTALIIASIVTVAHGQTIVHITLGHNHTLVGPPITSEVIWT
KLGSVDYFDIICNKTKPIFVICNRQNLTLINVSKIYNGYYYGYDRSSSQYKNYLVRIT
QPKLTVPTMTIIKMANKALENFTSPTTPNEKNIPNSMIAIIAAVALGMALIIICMLLY

ACYYKKFQHKQDPLLNFNI"

CDS 29202..29771

/gene="E3"

/locus_tag="HAdVB1_gp13"

/function="unknown"

/note="1 CR1 domain; genus-specific; MP family"

/codon_start=1

/product="membrane glycoprotein E3 CR1-gamma"

/protein_id="YP_002213790.1"

/db_xref="GeneID:6870528"

/translation="MISTTIFIITSLAAVTYGRSHLTVPVGSTCTLQGPQEGYVTWWR
IYDNGGFARPCDQPGTKFSCNGRDLTIINITSNEQGFYYGTNYKNSLDYNIIVVPATT
SAPRKSTFSSSSAKASTIPKTASAMLKLPKIALSNSTAAPNTIPKSTIGIITAVVVGL
MIIFLCIMYYACCYRKHEQKGDALLNFDI"

CDS 29786..30019

/gene="E3"

/locus_tag="HAdVB1_gp13"

/function="unknown"

/note="0 CR1 domains; MP family"

/codon_start=1

/product="membrane protein E3 CR1-delta"

/protein_id="YP_002213791.1"

/db_xref="GeneID:6870528"

/translation="MILFQSNTTTSYAYTNIQPKYAMQLEITILIVIGILILSVILYF
IFCRQIPNVHRNSKRRPIYSPMISRPHMALNEI"

CDS 30061..30336

/gene="E3"

/locus_tag="HAdVB1_gp13"

/function="apoptosis"

/note="genus-specific; MP family"

/codon_start=1

/product="membrane protein E3 RID-alpha"

/protein_id="YP_002213792.1"

/db_xref="GeneID:6870528"

/translation="MIPRNFFFTILICAFNVCATFTAVATASPDCIGPFASYALFAFV
TCICVCSIVCLVINFFQLVDWIFVRIAYLRHHPEYRNQNVAALLRLI"

CDS 30341..30745

/gene="E3"

/locus_tag="HAdVB1_gp13"

/function="apoptosis"

/note="genus-specific; MP family"

/codon_start=1

/product="membrane protein E3 RID-beta"

/protein_id="YP_002213793.1"

/db_xref="GeneID:6870528"
/translation="MQAMLPVILILLLPCIPLASTATRATPEQLRKCKFQQPWSFLDC
YHEKSDFPTYWIVIVGIINILSCTFFSITIYPTFNFGWNSPNALGYPQEPDEHIPLQH
IQQPLALVQYENEPQPSLPPAISYFNLTGGDD"
CDS 30738..31148
/gene="E3"
/locus_tag="HAdVB1_gp13"
/function="apoptosis"
/note="genus-specific; 12.5K family"
/codon_start=1
/product="control protein E3 14.7K"
/protein_id="YP_002213794.1"
/db_xref="GeneID:6870528"
/translation="MTDPIATSSTAAKELLDMDGRASEQRLIQLRIRQQQERAVKELR
DAIGIHQCKKGIFCLVKQSKISYEITATDHRLSYELGPQRQKFTCMVGINPIVITQQS
GDTKGCIQCSCDSTECIYTLLKTLCGLRDLLPMN"
regulatory 31181..31186
/regulatory_class="polyA_signal_sequence"
/gene="E3"
/locus_tag="HAdVB1_gp13"
gene complement(<31196..31353)
/gene="U"
/locus_tag="HAdVB1_gp14"
/db_xref="GeneID:6870529"
CDS complement(<31196..31353)
/gene="U"
/locus_tag="HAdVB1_gp14"
/function="unknown"
/note="exon encoding C terminus unidentified;
genus-common"
/codon_start=1
/product="protein U"
/protein_id="YP_002213834.1"
/db_xref="GeneID:6870529"
/translation="MKIVGKEEEIKFDIPFKVWRKYATLRRLEYQSWEEGEVLLVESF
DKDCIADF"
CDS 31368..32327
/gene="L5"
/locus_tag="HAdVB1_gp06"
/function="cell attachment"
/note="capsid protein IV; genus-common; fiber family"
/codon_start=1
/product="fiber"
/protein_id="YP_002213796.1"

/db_xref="GeneID:6870519"
/translation="MAKRARLSTSFNPVYPYEDESSSQHPFINPGFISPDGFTQSPNG
VLSLKCVNPLTTASGSLQLKVGSGLTVDTTDGSLEENIKVNTPLTKSNHSINLPIGNG
LQIEQNKLCSKLGNGLTFDSSNSIALKNNTLWTGPKPEANCIIEYGKQNPDSKLTLIL
VKNGGIVNGYVTLMGASDYVNTLFKNKNVSINVELYFDATGHILPDSSSLKTDLELKY
KQTADFSARGFMPSTTAYPFVLPNAGTHNENYIFGQCYYKASDGALFPLEVTVMLNKR
LPDSRTSYVMTFLWSLNAGLAPETTQATLITSPFTFSYIREDD"
regulatory 32335..32340
/regulatory_class="polyA_signal_sequence"
/gene="L5"
/locus_tag="HAdVB1_gp06"
gene complement(32352..35040)
/gene="E4"
/locus_tag="HAdVB1_gp15"
/db_xref="GeneID:6870530"
regulatory complement(32352..32357)
/regulatory_class="polyA_signal_sequence"
/gene="E4"
/locus_tag="HAdVB1_gp15"
CDS complement(join(32368..32619,33342..33515))
/gene="E4"
/locus_tag="HAdVB1_gp15"
/function="gene regulation"
/function="cell cycle regulation"
/note="genus-specific"
/codon_start=1
/product="control protein E4orf6/7"
/protein_id="YP_002213835.1"
/db_xref="GeneID:6870530"
/translation="MSGSNSIMTRLRARSTSCARHHPYTRAQLPRCEENETRASMTED
HPLLPDCDTMTMHSMTVIQTPESHPQQLDCESALKDYPDAFLCITDPRLARSETVWNV
ESKTMSISNGIQMFKAVRGERLVYSVKWEGGGKITTRIL"
CDS complement(32616..33515)
/gene="E4"
/locus_tag="HAdVB1_gp15"
/function="apoptosis"
/function="DNA damage response"
/function="RNA metabolism and transport"
/function="transformation"
/note="E4orf6; genus-specific; 34K family"
/codon_start=1
/product="control protein E4 34K"
/protein_id="YP_002213836.1"
/db_xref="GeneID:6870530"

/translation="MSGSNSIMTRLRARSTSCARHHPYTRAQLPRCEENETRASMTED
HPLLPDCDTMTMHSVSCVRGLPCSASFTVLQEFPIPWDMFLNPEELKIMRRCMHLCLC
CATIDIFQSQVIHGRENWVLHCHCNQQGSLQCMAGGAVLAVWFRKVILGCMINQRCPW
YRQIVNMHMPKEIMYVGSVFLRERHLIYIKLWYDGHAGAIISDMSFGWSAFNYGLLNN
IVIMCCTYCSNLSEIRMRCCAHRTRKLMLRAIKIMLQDTADPDPINSSRTERRRQRLL
VGLMRHNRPIPFSDYDSHRSSSR"

intron complement(32620..33341)
/gene="E4"
/locus_tag="HAdVB1_gp15"

CDS complement(33418..33786)
/gene="E4"
/locus_tag="HAdVB1_gp15"
/function="RNA splicing"
/note="genus-specific"
/codon_start=1
/product="control protein E4orf4"
/protein_id="YP_002213837.1"
/db_xref="GeneID:6870530"
/translation="MILPSLPSPLLLETQSSCIAWLGFAYATVDDFLRTIKYDGVLIT
TEASILLTNLRVWLYLNYQTEHTKRQDRRRRRVCCARAAFCYTKYENVRKQLHYDKVA
RTLHIMRQTSPLHQSPITTM"

intron complement(33545..34960)
/gene="E4"
/locus_tag="HAdVB1_gp15"
/note="precedes control protein E4 34K CDS"

CDS complement(33795..34148)
/gene="E4"
/locus_tag="HAdVB1_gp15"
/function="DNA damage response"
/function="transformation"
/note="genus-specific"
/codon_start=1
/product="control protein E4orf3"
/protein_id="YP_002213838.1"
/db_xref="GeneID:6870530"
/translation="MRVCLRMVVEGALRDLFVMCGLDLPQELTRIIQGWKAENYLGMV
QECNMMFEELENAPSFGILLFLDVRVEALLEATVEHLENRISFDLAVLFHQHSGGDRC
HLRDLQFEVLRDRLE"

intron complement(33881..34960)
/gene="E4"
/locus_tag="HAdVB1_gp15"
/note="precedes control protein E4orf4 CDS"

CDS complement(34145..34534)
/gene="E4"

/locus_tag="HAdVB1_gp15"
/function="unknown"
/note="genus-specific"
/codon_start=1
/product="control protein E4orf2"
/protein_id="YP_002213839.1"
/db_xref="GeneID:6870530"
/translation="MLERSAVTYCICIPTVLNVSLENFSFIQFLQETLPDYISSNFDD
ITGGSQYAYSNLVFAGANWGGLRLYCTVASPAIVPGGLLAKQFGDNMREYLQLELREE
LRAKGMSIDDAILNLLQVTQEQDILSL"

intron complement(34178..34960)
/gene="E4"
/locus_tag="HAdVB1_gp15"
/note="precedes control protein E4orf3 CDS"

intron complement(34553..34960)
/gene="E4"
/locus_tag="HAdVB1_gp15"
/note="precedes control protein E4orf2 CDS"

CDS complement(34576..34953)
/gene="E4"
/locus_tag="HAdVB1_gp15"
/function="transformation"
/note="genus-specific; DURP family"
/codon_start=1
/product="control protein E4orf1"
/protein_id="YP_002213840.1"
/db_xref="GeneID:6870530"
/translation="MADEALYVYLEGPGATLPEQQQRNNYIFYSPVPFTLYPRGVALL
YLKLSIIIPRGYVGCFFSLTDANMSGLYASSRIIHAGHREELSVLLFNHNDRFYEGRA
GDPVACLVMERLIFPPVRQATMI"

regulatory complement(35035..35040)
/regulatory_class="TATA_box"
/gene="E4"
/locus_tag="HAdVB1_gp15"

repeat_region 35208..35343
/standard_name="ITR"
/rpt_type=inverted

ORIGIN

```
     1 ctatctatat aatatacctt atagatggaa tggtgccaac atgtaaatga ggtaatttaa
    61 aaaagtgcgc gctgtgtggt gattggctgc ggggttaacg gctaaaaggg gcggcgcgac
   121 cgtgggaaaa tgacgtgact tatgtgggag gagttatgtt gcaagttatt acggtaaatg
   181 tgacgtaaaa cgaggtgtgg tttgaacacg gaagtagaca gttttcccac gcttactgac
   241 aggatatgag gtagttttgg gcggatgcaa gtgaaaattc tccattttcg cgcgaaaact
   301 aaatgaggaa gtgaatttct gagtcatttc gcggttatgc cagggtggag tatttgccga
```

```
 361 gggccgagta gactttgacc gtttacgtgg aggtttcgat taccgtgttt ttcacctaaa
 421 tttccgcgta cggtgtcaaa gttctgtgtt tttacgtagg tgtcagctga tcgctagggt
 481 atttaaacct gacgagttcc gtcaagaggc cactcttgag tgccagcgag aagagttttc
 541 tcctccgcgc cgcaagtcag ttctgcgctt tgaaaatgag acacctgcgc ttcctgccac
 601 aggaggttat ctccagtgag accgggatcg aaatactgga gtttgtggta aataccctaa
 661 tgggagacga cccggaaccg ccagtgcagc ctttcgatcc acctacgctg cacgatctgt
 721 atgatttaga gatagacggg ccggaggatc ccaatgagga agctgtgaat gggtttttta
 781 ctgattctat gctgctagct gctgatgaag gattggacat aaaccctcct cctgagacac
 841 ttgttacccc aggggtggtt gtggaaagcg gcataggtgg gaaaaaattg cctgatctgg
 901 gagcagctga aatggacttg cgttgttatg aagagggttt tcctcccagt gatgatgaag
 961 atggggaaac tgagcagtcc atccataccg cagtaaatga gggagtaaaa gctgccagcg
1021 atgtttttaa gttggactgt ccggagctgc ctggacatgg ctgtaagtct tgtgaatttc
1081 acaggaataa cactggaatg aaagaactat tgtgctcgct ttgctatatg agaatgcact
1141 gccactttat ttacagtaag tgtatttaag tgaaatttaa aggaatagtg tagctattta
1201 ataactgttg aatggtagat ttatgttttt ttcttgcgat tttttgtagg tcctgtgtct
1261 gatgatgagt caccttctcc tgattcaact acctcacctc ctgaaattca ggcgcccgca
1321 cctgcaaacg tatgcaagcc cattcctgtg aagcctaagc ctgggaaacg ccctgctgtg
1381 gataagcttg aggacttgtt ggagggtggg gatggacctt tggaccttag tacccggaaa
1441 ctgccaaggc aatgagtgcc ctgcagctgt gtttatttaa tgtgacgtca tgtaataaaa
1501 ttatgtcagc tgctgagtgt tttattactt cttgggtggg gacttggata tataagtagg
1561 agcagatctg tgtggttagc tcacagcaac ctgctgccat ccatggaggt ttgggctatc
1621 ttggaagacc tcagacagac taagctactg ctagaaaacg cctcggacgg agtctctggc
1681 ctttggagat tctggttcgg tggtgatcta gctaggctag tgtttaggat aaaacaggac
1741 tacagggaag aatttgaaaa gttattggac gatagtccgg gacttttga agctcttaac
1801 ttgggtcatc aggctcattt taaggagaag gttttatcag ttttagattt ttctactcct
1861 ggtagaactg ctgctgctgt agcttttctt acttttatat tggataaatg gatccgccaa
1921 actcacttca gcaagggata cgttttggat ttcatagcag cagctttgtg gagaacatgg
1981 aaggctcgca ggatgaggac aatcttagat tactggccag tgcagcctct gggagtagca
2041 gggatactga gacacccacc gaccatgcca gcggttctgc aggaggagca gcaggaggac
2101 aatccgagag ccggcctgga ccctccggtg gaggagtagc tgacctgttt cctgaactgc
2161 gacgggtgct tactaggtct acgaccagtg gacagaacag gggaattaag agggagagga
2221 atcctagtgg gaataattca agaaccgagt tggctttaag tttaatgagc cgcaggcgtc
2281 ctgaaactgt ttggtggcat gaggttcaga gcgaaggcag ggatgaagtt tcaatattgc
2341 aggagaaata ttcactagaa caacttaaga cctgttggtt ggaacctgag gatgattggg
2401 aggtggccat taggaattat gctaagatat ctctgaggcc tgataaacaa tatagaatta
2461 ctaagaagat taatattaga aatgcatgct acatatcagg gaatggggca gaggttataa
2521 tagatacaca agataaagca gtttttagat gttgtatgat gggtatgtgg ccaggggttg
2581 tcggcatgga agcagtaaca cttatgaata ttaggtttaa aggggatggg tataatggca
2641 ttgtatttat ggctaacact aagctgattc tacatggttg tagctttttt gggtttaata
2701 atacgtgtgt agaagcttgg gggcaagtta gtgtgagggg ttgtagtttt tatgcatgct
2761 ggattgcaac atcaggtagg gtcaagagtc agttgtctgt gaagaaatgc atgtttgaga
2821 gatgtaatct tggcatactg aatgaaggtg aagcaagggt ccgccactgc gcagctacag
2881 aaactggctg cttcattcta ataaagggaa atgccagtgt gaagcataat atgatctgtg
2941 gacattcgga tgagaggcct tatcagatgc tgacctgcgc tggtggacat tgcaatattc
```

3001 ttgctaccgt gcatatcgtt tcacatgcac gcaaaaaatg gcctgtattt gaacataatg
3061 tgattaccaa gtgcaccatg cacataggtg gtcgcagggg aatgtttatg ccttaccagt
3121 gtaacatgaa tcatgtgaag gtaatgttgg aaccagatgc cttttccaga gtgagcttaa
3181 caggaatctt tgatatgaat attcaactat ggaagatcct gagatatgat gacactaaac
3241 caagggtgcg cgcatgcgaa tgcggaggca agcatgctag attccagccg gtgtgcgtgg
3301 atgtgactga agacttgaga cccgatcatt tggtgcttgc ctgcactgga gcggagttcg
3361 gttctagtgg tgaagaaact gactaaagtg agtagtgggg caagatgtgg atggggactt
3421 tcaggttggt aaggtgggca gattgggtaa attttgttaa tttctgtctt gcagctgcca
3481 tgagtggaag cgcttctttt gagggggggag tatttagccc ttatctgacg ggcaggctcc
3541 caccatgggc aggagttcgt cagaatgtca tgggatccac tgtggatggg agacccgtcc
3601 agcccgccaa ttcctcaacg ctgacctatg ccactttgag ttcgtcacca ttggatgcag
3661 ctgcagccgc cgccgctact gctgccgcca acaccatcct tggaatgggc tattatggaa
3721 gcatcgttgc caattccagt tcctctaata acccttcaac cctggctgag gacaagctac
3781 ttgttctctt ggcgcagctc gaggccttaa cccaacgctt aggcgaactg tctaagcagg
3841 tggcccagtt gcgtgagcaa actgagtctg ctgttgccac agcaaagtct aaataaagat
3901 ctcaaatcaa taaataaaga aatacttgat ataaaacaaa tgaatgttta tttgattttt
3961 cgcgcgcggt atgccctgga ccatcggtct cgatcattga gaacgcggtg gatcttttcc
4021 agtaccctgt aaaggtggga ttgaatgttt agatacatgg gcattagtcc gtctcggggg
4081 tggagatagc tccattgaag agcctcttgc tccggggtag tgttataaat cacccagtca
4141 tagcaaggtc ggagtgcatg gtgttgcaca atatctttta ggagcagact aattgcaacg
4201 gggaggccct tagtgtaggt gtttacaaat ctgttgagct gggacgggtg catccggggt
4261 gaaattatat gcattttgga ctggatcttg aggttggcaa tgttgccgcc tagatcccgt
4321 ctcgggttca tattgtgcag gaccaccaag acagtgtatc cggtgcactt gggaaattta
4381 tcatgcagct tagagggaaa agcatgaaaa aatttggaga cgcctttgtg acccccccaga
4441 ttctccatgc actcatccat aatgatagcg atggggccgt gggcagcggc acgggcgaac
4501 acgttccggg ggtctgaaac atcatagtta tgctcctgag tcaggtcatc ataagccatt
4561 ttaataaact ttgggcggag ggtgccagat tgggggatga aagttccctc tggcccggga
4621 gcatagtttc cctcacatat ttgcatttcc caggctttca gttcagaggg ggggatcatg
4681 tccacctgcg gggctataaa aaataccgtt tctggagccg gggtgattaa ctgggacgag
4741 agcaaattcc taagcagctg agacttgccg cacccagtgg gaccgtaaat gaccccaatt
4801 acgggttgca gatggtagtt tagggaacga cagctgccgt cctcccggag caggggggcc
4861 acttcgttca tcatttccct tacatggata ttttcccgca ccaagtccgt taggaggcgc
4921 tctcccccaa gggatagaag ctcctggagc gaggagaagt ttttcagcgg cttcagcccg
4981 tcagccatgg gcattttgga aagagtctgt tgcaagagct cgagccggtc ccagagctcg
5041 gtgatgtgct ctatggcatc tcgatccagc agacctcctc gtttcgcggg ttgggacggc
5101 tcctggagta gggaatcaga cgatgggcgt ccagcgctgc cagggtccga tccttccatg
5161 gtcgcagcgt ccgagtcagg gttgtttccg tcacggtgaa ggggtgcgcg cctggttggg
5221 cgcttgcgag ggtgcgcttc agactcatcc tgctggtcga gaaccgctgc cgattggcgc
5281 cctgcatgtc ggccaggtag cagtttacca tgagttcgta gttgagcgcc tcggccgcgt
5341 ggcctttggc acggagctta cctttggaag ttttatggca ggcggggcag tagatacatt
5401 tgagggcata caacttgggc gcgaggaaaa tggattcggg ggagtatgca tccgcaccgc
5461 aggaggcgca gacggtttcg cactccacga gccaggtcag atccggctca tcggggtcaa
5521 aaacaagttt tccgccatgt tttttgatgc gtttcttacc tttggtttcc atgagttcgt
5581 gtccccgctg ggtgacaaag aggctgtccg tgtccccgta gaccgacttt atgggcctgt

5641 cctcgagcgg agtgcctcgg tcctcttcgt agaggaaccc agcccactct gatacaaaag
5701 cgcgtgtcca ggccagcaca aaggaggcca cgtgggaggg gtagcggtcg ttgtcaacca
5761 gggggtccac cttctctacg gtatgtaaac acatgtcccc ctcctccaca tccaagaatg
5821 tgattggctt gtaagtgtag gccacgtgac caggggtccc cgccgggggg gtataaaagg
5881 gggcgggcct ctgttcgtcc tcactgtctt ccggatcgct gtccaggagc gccagctgtt
5941 ggggtaggta ttccctctcg aaggcgggca tgacctctgc actcaggttg tcagtttcta
6001 ggaacgagga ggatttgata ttgacagtac cagcagagat gcctttcata agactctcgt
6061 ccatctggtc agaaaacaca atcttcttgt tgtccagctt ggtggcaaat gatccataga
6121 gggcattgga tagaagcttg gcgatggagc gcatggtttg gttcttttcc ttgtccgcgc
6181 gctccttggc ggtgatgtta agctggacgt actcgcgcgc cacacatttc cattcaggaa
6241 agatggttgt cagttcatcc ggaactattc tgattcgcca tcccctattg tgcagggtta
6301 tcagatccac actggtggcc acctcgcctc ggaggggctc attggtccag cagagtcgac
6361 ctccttttct tgaacagaaa ggggggaggg ggtctagcat gaactcatca ggggggtccg
6421 catctatggt aaatattccc ggtagcaaat ctttgtcaaa atagctgatg gtggcgggat
6481 catccaaggt catctgccat tctcgaactg ccagcgcgcg ctcatagggg ttaagagggg
6541 tgccccaggg catggggtgg gtgagcgcgg aggcatacat gccacagata tcgtagacat
6601 agaggggctc ttcgaggatg ccgatgtaag tgggataaca tcgcccccct ctgatgcttg
6661 ctcgcacata gtcatagagt tcatgtgagg gggcaagaag acccgggccc agattggtgc
6721 ggttgggttt ttccgccctg taaacgatct ggcgaaagat ggcatgggaa ttggaagaga
6781 tagtaggtct ctggaatatg ttaaaatggg catgaggtaa gcctacagag tcccttatga
6841 agtgggcata tgactcttgc agcttggcta ccagctcggc ggtgatgagt acatccaggg
6901 cacagtagtc gagagtttcc tggatgatgt cataacgcgg ttggcttttc ttttcccaca
6961 gctcgcggtt gagaaggtat tcttcgtgat ccttccagta ctcttcgagg ggaaacccgt
7021 ctttttctgc acggtaagag cccaacatgt agaactgatt gactgccttg tagggacagc
7081 atcccttctc cactgggaga gagtatgctt gggctgcatt gcgcagcgag gtatgagtga
7141 gggcaaaagt gtccctgacc atgactttga ggaattgata cttgaagtcg atgtcatcac
7201 aggccccctg ttcccagagt tggaagtcca cccgcttctt gtaggcgggg ttgggcaaag
7261 cgaaagtaac atcattgaag aggatcttgc cggccctggg catgaaattt cgggtgattt
7321 tgaaaggctg aggaacctct gctcggttat tgataacctg agcggccaag acgatctcat
7381 caaagccatt gatgttgtgc cccactatgt acagttctaa gaatcgaggg gtgcccctga
7441 catgaggcag cttcttgagt tcttcaaaag tgagatctgt agggtcagtg agagcatagt
7501 gttcgagggc ccattcgtgc acgtgagggt tcgctttaag gaaggaggac cagaggtcca
7561 ctgccagtgc tgtttgtaac tggtcccggt actgacgaaa atgctgtccg actgccatct
7621 tttctggggt gacgcaatag aaggtttggg ggtcctgccg ccagcgatcc cacttgagtt
7681 ttatggcgag gtcataggcg atgttgacga gccgctggtc tccagagagt ttcatgacca
7741 gcatgaaggg gattagctgc ttgccaaagg accccatcca ggtgtaggtt tccacatcgt
7801 aggtgagaaa gagcctttct gtgcgaggat gagagccaat cgggaagaac tggatctcct
7861 gccaccagtt ggaggaatgg ctgttgatgt gatggaagta gaactccctg cgacgcgccg
7921 agcattcatg cttgtgcttg tacagacggc cgcagtactc gcagcgattc acgggatgca
7981 ccttatgaat gagttgtacc tgacttcctt tgacgagaaa tttcagtgga aaattgaggc
8041 ctggcgcttg tacctcgcgc tttactatgt tgtctgcatc ggcatgacca tcttctgtct
8101 cgatggtggt catgctgacg agccctcgcg ggaggcaagt ccagacctcg gcgcggcagg
8161 ggcggagctc gaggacgaga gcgcgcaggc cggagctgtc cagggtcctg agacgctgcg
8221 gagtcaggtt agtaggcagt gtcaggagat taacttgcat gatcttttgg agggcgtgag

8281 ggaggttcag atagtacttg atctcaacgg gtccgttggt ggagatgtcg atggcttgca
8341 gggttccgtg tcccttgggc gctaccaccg tgcccttgtt tttcattttg gacggcggtg
8401 gctctgttgc ttcttgcatg tttagaagcg gtgtcgaggg cgcgcaccgg gcggcagggg
8461 cggctcggga cccggcggca tggctggcag tggtacgtcg gcgccgcgcg cgggtaggtt
8521 ctggtactgc gccctgagaa gactcgcatg cgcgacgacg cggcggttga catcctggat
8581 ctgacgcctc tgggtgaaag ctaccggccc cgtgagcttg aacctgaaag agagttcaac
8641 agaatcaatc tcggtatcgt tgacggcggc ttgcctaagg atttcttgca cgtcgccaga
8701 gttgtcctgg taggcgatct cggccatgaa ctgctcgatc tcttcctctt gaagatctcc
8761 gcggcccgct ctctcgacgg tggccgcgag gtcgttggag atgcgcccaa tgagttgaga
8821 gaatgcattc atgcccgcct cgttccagac gcggctgtag accacagccc ccacgggatc
8881 tctcgcgcgc atgaccacct gggcgaggtt gagctccacg tggcgggtga agaccgcata
8941 gttgcatagg cgctggaaaa ggtagttgag tgtggtggcg atgtgctcgg tgacgaagaa
9001 atacatgatc catcgtctca gcggcatctc gctgacatcg cccagcgctt ccaagcgctc
9061 catggcctcg tagaagtcca cggcaaagtt aaaaaactgg gagttacgcg cggacacggt
9121 caactcctct tccagaagac ggataagttc ggcgatggtg gtgcgcacct cgcgctcgaa
9181 agcccctggg atttcttcct caatctcttc ttcttccact aacatctctt cctcttcagg
9241 tggggctgca ggaggagggg gaacgcggcg acgccggcgg cgcacgggca gacggtcgat
9301 gaatctttca atgacctctc cgcggcggcg gcgcatggtc tcggtgacgg cacgaccgtt
9361 ctccctgggt ctcagagtga agacgcctcc gcgcatctcc ctgaagtggt gactgggagg
9421 ctctccgttg ggcagggaca ccgcgctgat tatgcatttt atcaattgcc ccgtaggtac
9481 tccgcgcaag gacctgatcg tctcaagatc cacgggatct gaaaaccttt cgacgaaagc
9541 gtctaaccag tcgcaatcgc aaggtaggct gagcactgtt tcttgcgggc gggggcggct
9601 agacgctcgg tcggggttct ctctttcttc tccttcctcc tcttgggagg gtgagacgat
9661 gctgctggtg atgaaattaa aataggcagt tttgagacgg cggatggtgg cgaggagcac
9721 caggtctttg ggtccggctt gttggatacg caggcgatga gccattcccc aagcattatc
9781 ctgacatctg gccagatctt tatagtagtc ttgcatgagt cgttccacgg gcacttcttc
9841 ttcgcccgct ctgccatgca tgcgagtgat cccgaacccg cgcatgggct ggacaagtgc
9901 caggtccgct acaacccttt cggcgaggat ggcttgctgc acctgggtga gggtggcttg
9961 gaagtcgtca aagtccacga agcggtggta ggccccggtg ttgattgtgt aggagcagtt
10021 ggccatgact gaccagttga ctgtctggtg cccagggcgc acgagctcgg tgtacttgag
10081 gcgcgagtat gcgcgggtgt caaagatgta atcgttgcag gtgcgcacca ggtactggta
10141 gccaatgaga aagtgtggcg gtggctggcg gtacaggggc catcgctctg tagccggggc
10201 tccgggggcg aggtcttcca gcatgaggcg gtggtagccg tagatgtacc tggacatcca
10261 ggtgataccg gaggcggtgg tggatgcacg tgggaactcg cgcacgcggt tccagatgtt
10321 gcgcagcggc atgaagtagt tcatggtagg cacggtctgg ccagtgaggc gcgcgcagtc
10381 attgacgctc tgtagacacg gagaaaacga aagcgatgag cggctcgact ccgtggtctg
10441 ggggaacgtg aacgggttgg gtcgcggtgt accccggttc gagtccaaag ctaagcgatc
10501 acgctcggat cggccggagc cgcggctaac gtggtattgg ctatcccgtc tcgacccagc
10561 cgacgaatat ccagggtacg gagtagagtc gtttttgctg ctttttcctg gacgtgtgcc
10621 attgccacgt caagctttag aacgctcagt tctcgggccg tgagtggctc gcgcccgtag
10681 tctggagaat cagtcgccag ggttgcgttg cggtatgccc cggttggagc ctaagcgcgg
10741 ctcgtatcgg ccggtttccg cgacaagcga gggtttggca gccccgttat ttccaagacc
10801 ccgccagccg acttctccag tttacgggag cgagcccttt ttttttttt tttgtttttg
10861 tcgcccagat gcatccagtg ctgcgacaga tgcgccccca gcaacaggcc ccttctcagc

10921 aacagccaca aaaggctctt cttgctcctg caactactgc agctgcagcc gtgagcggcg
10981 cgggacagcc cgcctatgat ctggacttgg aagagggcga gggattggcg cgcctggggg
11041 ctccatcgcc cgagcggcac ccgcgggtgc aactaaaaaa ggactctcgc gaggcgtacg
11101 tgccccagca gaacctgttc agggacagga gcggcgagga gccagaggag atgcgagcat
11161 ctcgatttaa cgcgggtcgc gagctgcgcc acggtctgga tcgaagacgg gtgctgcaag
11221 acgaggattt tgaggtcgat gaagtcacag ggatcagccc agctagggca catgtggccg
11281 cggccaacct agtctcggcc tacgagcaga ccgtgaagga ggagcgcaac ttccaaaaat
11341 cttttaacaa ccatgtgcgc accctgatcg cccgcgagga agtgaccctg ggtctgatgc
11401 atctgtggga cctgatggag gctatcgccc aaaaccccac tagcaaacca ctgacagctc
11461 agctgtttct ggtggttcaa catagcaggg acaacgaggc attcagggag gcgttgttga
11521 acatcaccga gcctgatggg agatggctgt atgatctgat caacatcctg caaagtatta
11581 tagtgcagga acgtagcctg ggtttggctg agaaagtggc agctatcaac tactcggtct
11641 tgagcctggg caaatactac gctcgcaaga tctacaagac cccctacgta cccatagaca
11701 aggaggtgaa gatagatggg ttttacatgc gcatgactct gaaggtgctg actctgagcg
11761 acgatctggg ggtgtatcgc aatgacagga tgcaccgcgc ggtgagcgcc agcaggaggc
11821 gcgagctgag cgacagagaa cttatgcaca gcttgcaaag agctctaacg ggggccggga
11881 ctgatgggga gaactacttt gacatgggag cggatttgca atggcaaccc agtcgcaggg
11941 ccatggaggc tgcagggtgt gagcttcctt acatagaaga ggtggatgaa gtcgaggacg
12001 aggagggcga gtacttggaa gactgatggc gcgacccgta tttttgctag atggaacagc
12061 agcaggcacc ggaccccgca atgcgggcgg cgctgcagag ccagccgtcc ggcattaact
12121 cctcggacga ttggacccag gccatgcaac gcataatggc gctgacgacc cgcaaccccg
12181 aagcctttag acagcaaccc caggccaacc gcctttctgc catactggag gccgtagtgc
12241 cctcccgctc caaccccacc cacgagaagg tcctggctat cgtgaacgcg ctggtggaga
12301 acaaggccat ccgtcccgat gaggccgggc tggtatacaa tgctctcttg gagcgcgtgg
12361 cccgttacaa cagcagcaac gtgcaaacca acctggaccg gatggtgacc gatgtgcgcg
12421 aggccgtgtc tcagcgcgag cgattccagc gcgacgccaa cttggggtcg ttggtagcgc
12481 taaacgcttt cctcagcacc cagcccgcca acgtgccccg tggtcagcaa gactatacaa
12541 acttttgag tgcattgaga ctcatggtag ctgaggtgcc ccagagcgag gtgtaccagt
12601 ccgggccaga ttacttcttc cagaccagca gacagggctt gcagacagtg aacctgactc
12661 aggctttcaa gaacctgaag ggtctgtggg gagtgcacgc cccagtaggg gatcgcgcga
12721 ccgtgtctag cttgctgact cccaactccc gcctgctgct gctgctggta tccccccttta
12781 ctgacagcgg tagcattgac cgcaactcgt acttgggcta cctgcttaac ctgtatcgcg
12841 aggccatagg acagagccag gtggacgagc agacctatca agaaatcacc caagtgagcc
12901 gcgccctggg tcaggaagac acgggcagtt tggaagccac cctgaacttc ttgctaacca
12961 accggtcaca gaagatccct cctcagtatg cgcttaccgc tgaggaggag cggatcctca
13021 gatacgtgca acagagcgtt ggactgttcc tgatgcagga gggggcgaca cctaccgccg
13081 cgctggacat gacagctcga aacatggagc ccagcatgta tgctagtaac aggcctttca
13141 ttaacaaact gctggactac ctgcacaggg cggccgccat gaactctgat tatttcacca
13201 atgctatcct gaacccacac tggctgcccc cacctggttt ctacactggc gagtacgaca
13261 tgcccgaccc caatgacggg ttcctgtggg acgatgtgga cagcagcata ttttccccgc
13321 ctcccggtta tacagtttgg aagaaggaag ggggcgatag aagacactct tccgtgtcgc
13381 tatccggaac ggctggtgct gccgcgaccg tgcccgaagc tgcaagtcct ttccctagct
13441 tgccctttc actaaacagc gttcgcagca gtgaactggg gagaataacc cgcccgcgct
13501 tgatgggcga ggatgagtac ttgaatgact ctttgctgag gccagagagg gaaaagaact

13561 tccccaacaa tggaatagag agtctggtgg ataagatgag tagatggaag acctatgcgc
13621 aggatcacag agacgagccc aggatcttgg gggctacaag cagaccgatc cgtagacgcc
13681 agcgccacga caggcagatg ggtcttgtgt gggacgatga ggactctgcc gatgatagca
13741 gcgtgttgga cttgggtgga agaggagggg gcaacccgtt cgctcatctg cgtcccagat
13801 tcgggcgcat gttgtaaaag tgaaagtaaa ataaaaaggc aactcaccaa ggccatggcg
13861 accgagcgtg cgttcgttct tttttgttat ctgtgtctag tacgatgagg agacgagccg
13921 tgctaggcgg agcggtggtg tatccggagg gtcctcctcc ttcttacgag agcgtgatgc
13981 agcaacaggc ggcgatgata cagcccccac tggaggctcc cttcgtaccc ccacggtacc
14041 tggcgcctac ggaagggaga aacagcattc gttactcgga gctgtcgccc ctgtacgata
14101 ccaccaagtt gtatctggtg gacaacaagt cggcggacat cgcctccctg aactatcaga
14161 acgaccacag caacttcctg accacggtgg tgcagaacaa tgactttacc cccacggagg
14221 ctagcaccca gaccatcaac tttgacgagc ggtcgcgatg gggcggtcag ctgaagacca
14281 tcatgcacac caacatgccc aacgtgaacg agtacatgtt cagcaacaag ttcaaggcga
14341 gggtgatggt gtccagaaaa gctcctgaag gtgttacagt aaatgacacc tatgatcata
14401 aagaggatat cttgaagtat gagtggtttg agttcatttt accagaaggc aacttttcag
14461 ccaccatgac gatcgacctg atgaacaatg ccatcattga caactacctg gaaattggca
14521 gacagaatgg agtgctggaa agtgacattg gtgttaagtt tgacactaga aatttcaggc
14581 tcgggtggga ccccgaaact aagttgatta tgccaggtgt ctacacttat gaggcattcc
14641 atcctgacat tgtattgctg cctggttgcg gggtagactt tactgaaagc cgacttagca
14701 acttgcttgg catcaggaag agacatccat tccaggaggg tttcaaaatc atgtatgaag
14761 atcttgaagg gggtaatatt cctgcccttt tggatgtcac tgcctatgag gaaagcaaaa
14821 aggataccac tactgaaaca accacactgg ctgttgcaga ggaaactagt gaagatgatg
14881 atataactag aggagatacc tatataactg aaaaacaaaa acgtgaagct gcagctgctg
14941 aagttaaaaa agagttaaag atccaacctc tagaaaaaga cagcaagagt agaagctaca
15001 atgtcttgga agacaaaatc aacacagcct accgcagttg gtacctgtcc tacaattacg
15061 gtaaccctga gaaaggaata aggtcttgga cactgctcac cacttcagat gtcacctgtg
15121 gggcagagca ggtctactgg tcgctccctg acatgatgca agacccagtc accttccgct
15181 ccacaagaca agtcaacaac tacccagtgg tgggtgcaga gcttatgccc gtcttctcaa
15241 agagtttcta caatgagcaa gccgtgtact ctcagcagct ccgacaggcc acttcgctca
15301 cgcacgtctt caaccgcttc cctgagaacc agatcctcat ccgcccgccg gcgcccacaa
15361 ttaccaccgt cagtgaaaac gttcctgctc tcacagatca cgggaccctg ccgttacgca
15421 gcagtatccg gggagtccag cgcgtgaccg ttactgacgc cagacgccgc acctgtccct
15481 acgtttacaa ggccctgggc atagtcgcgc cgcgcgttct ttcaagccgc actttctaaa
15541 aaaaaaaaaa atgtccattc tcatctcgcc cagtaataat accggttggg gactgtatgc
15601 gcccaccaag atgtatggag gcgcccgcaa gcgctctacc cagcatcctg tgcgcgttcg
15661 cggtcatttc cgcgctccct ggggcgcact caagggtcgt acccgcactc ggaccacggt
15721 cgatgatgtg atcgaccagg tggtcgccga tgctcgtaat tatactccta ctgcgcctac
15781 atctactgtg gatgcagtta ttgacagtgt ggtggcagac gcccgcgcct atgctcgccg
15841 gaagagccga aggaggcgca tcgccaggcg ccacagggct actcccgcca tgcgagccgc
15901 aaaagctatt ctgcggaggg ccaaacgtgt ggggcgaaga gccatgctta gagcggccag
15961 acgcgcggct tcaggtgcca gcagcggcag gtcccgcagg cgcgcggcca cggcggcagc
16021 agcggccatt gccaacatgg cccaaccgcg aagaggcaat gtgtactggg tgcgtgatgc
16081 cactaccggc cagcgcgtgc ccgtgcgcac ccgcccccct cgcacttaga agatactgag
16141 cagtctccga tgttgtgtcc cagcggcaag tatgtccaag cgcaaataca aggaagagat

16201 gctccaggtc atcgcgcctg aaatctacgg tccgccggtg aaggatgaaa aaaagccccg
16261 caaaatcaag cgggtcaaaa ataacaaaaa ggaagaagat gacgatgatg ggctggtgga
16321 gtttgtgcgc gagttcgccc caagacggcg cgtgcagtgg cgcgggcgca aagtgcgtca
16381 agtgctcaga cccgggacca ctgtggtttt tacacctggc gagcgttcca gcactacttt
16441 taaacggtcc tatgatgagg tgtacgggga tgacgatatt cttgagcagg cggcagaccg
16501 ccttgacgag tttgcttatg gcaagcgcac tagatccagt cccaaagagg aggctgtgtc
16561 cattcctttg gatcatggaa atcccacccc cagcctcaaa ccagtcaccc tgcagcaagt
16621 gctgcccgtg cctgcgcgga gaggcgtaaa gcgcgagggt gaggacctat atcccaccat
16681 gcagctaatg gtgcccaagc gccagaggct agaagacgta ctggagaaaa tgaaagtgga
16741 tgccgatatc cagcctgagg tcaaagtaag acctatcaag gaagtggcgc caggtttggg
16801 agtacaaacc ttcgacatca agattcccac cgagtccatg gaagtgcaga ccgaacctgc
16861 aaaacccacc acctcaattg aggtgcaaac ggaaccctgg atgcccgcgc ccgttgccgc
16921 ccccagcacc actcgaagat cacgacgaaa gtacggccca gcaagtctgc taatgcccaa
16981 ctatgctctg cacccatcca tcattcccac tccgggttac agaggcactc gctactatcg
17041 aaaccggagc aacacctctc gccgccgcaa accacctgca agtcgcactc gccgtcgccg
17101 ccgccgcaac actgccagca aattgactcc cgccgccctg gtgcggagag tgtaccgcga
17161 tggtcgcgct gaacctctga cgctgccgcg cgcgcgctac catccaagca tcaccactta
17221 atgactgttg acgctgcctc cttgcagata tggccctcac ttgccgcctt cgcgtcccca
17281 ttactggcta ccgaggaaga aactcgcgcc gtagaaggat gttggggcga gggatgcgcc
17341 gccacagacg aaggcgcgct atcagcagac gattaggggg tggctttttg ccagctctta
17401 tacccatcat cgccgcagcg atcggggcga taccaggcat agcttcagtg gcggttcagg
17461 cctcgcagcg ccactaacat tggaaaaaac ttataaataa aaaatagaat ggactctgac
17521 gctcctggtc ctgtgactat gtttttgtag agatggaaga catcaatttt tcatccctgg
17581 ctccgcgaca cggcacgagg ccgtacatgg gcacctggag cgacatcggc acgagccaac
17641 tgaacggggg cgccttcaat tggagcagta tctggagcgg gcttaaaaat tttggctcga
17701 ccgtaaaaac ctatgggaac aaagcttgga acagcagcac agggcaggct ctgagaaata
17761 agcttaagga acaaaacttc caacagaagg tggtcgatgg gatcgcctct ggtattaacg
17821 gcgtagtgga tctggccaac caggctgtac aaaaacagat aaacagccgc ctggacccgc
17881 cgcccgcaac ccctggtgaa atggaagtgg aggaagaact tcctccgctg gaaaagcggg
17941 gcgacaagcg tccgcgaccc gagctagagc agacgctggt gacgcgcgca gacgagcccc
18001 cttcatatga ggaggcagta aagctcggaa tgcccactac caggcctgta gctcacatgg
18061 ctaccggggt gatgaaacct tctcagtcac atcgacccgc caccttggac ttgcctcctc
18121 cccctgcttc tgcggcgcct gttcccaaac ctgtcgctac cagaaagccc accgccgtac
18181 agcccgtcgc cgtagccaga ccgcgtcctg ggggcacacc gcgcccgaaa gcaaactggc
18241 aaagtactct gaacagcatc gtgggtctgg gcgtgcagag tgtaaagcgc cgtcgctgct
18301 attaattaaa tatggagtag cgcttaactt gcttgtctgt gtgtatgtat catcaccacg
18361 ccgccgcagc agaggagaaa ggaagaggtc gcgcgccgag gctgagttgc tttcaagatg
18421 gccaccccat cgatgatgcc ccaatgggca tacatgcaca tcgccggaca ggatgcttcg
18481 gagtacctca gtccgggtct ggtgcagttc gcccgtgcaa cagacaccta cttcagtatg
18541 gggaacaaat ttagaaaccc cacagtggcg cccacccacg atgtgaccac cgaccgtagc
18601 cagcgcctga tgctgcgctt cgtgcccgtt gaccgggaag acaataccta ctcttacaaa
18661 gttcgctaca cgctggctgt aggcgacaac agagtgcttg acatggccag cacattcttt
18721 gacattcggg gggtgcttga tagaggtcct agcttcaagc catattccgg cacagcttac
18781 aattcactcg ctcctaaggg cgcgcccaat acatctcagt ggatagttac aacaaatggg

18841 gacaatgcag taactaccac cacaaacaca tttggcattg cttccatgaa gggagacaat
18901 attactaaag aaggtttgca aattgggaaa gacattacca ctactgaagg agaagaaaag
18961 cccatttatg ccgataaaac atatcagcca gagcctcaag ttggagaaga atcatggact
19021 gatactgatg gaacaaatga aaagtttggt ggaagagccc ttaaaccagc taccaacatg
19081 aagccatgct acgggtcttt tgcaagacct acaaacataa aaggggggca agctaaaaac
19141 agaaaagtaa aaccaacaac cgaaggaggg gttgaaactg aggaaccaga tattgatatg
19201 gaatttttcg atggtagaga tgctgttgca ggagctttag cgcctgaaat tgtgctttat
19261 acggaaaatg taaatttgga aactccagac agtcatgtgg tatataaacc agaaacgtct
19321 aataactctc atgcaaattt gggtcaacaa gccatgccta acagacccaa ttacattggc
19381 ttcagggata acttcgtagg cctaatgtac tacaacagta ctggaaatat gggagttttg
19441 gctggccaag catcacaact gaatgcagtg gttgacttgc aggacagaaa tactgaactg
19501 tcatatcagc ttttgcttga ttctctggga gacagaacca gatacttcag catgtggaat
19561 caggctgtgg acagttacga tcccgatgtt cgcattattg aaaatcatgg catcgaggat
19621 gaactgccta attactgttt tcctctgaat ggcataggac cagggcacac atatcaaggc
19681 attaaagtta aaaccgatga cactaatgga tgggaaaaag atgctaatgt tgctccagct
19741 aatgaaataa ccataggcaa caacctggct atggaaatta atatccaagc taacctttgg
19801 agaagttttc tgtactctaa tgtggctttg taccttccag atgtttacaa gtacacgcca
19861 cctaacatta ctttgcccac taacaccaac acctatgagt acatgaacgg gcgagtggta
19921 tccccatccc tggttgattc atacatcaac attggcgcca ggtggtctct tgacccaatg
19981 gacaatgtga atccattcaa ccaccaccgc aatgctggtc tgcgctacag gtccatgctt
20041 ctgggaaatg gtcgttatgt gcctttccac atacaagtgc ctcagaaatt ctttgctgtc
20101 aagaacctac ttcttctacc tggctcctac acctacgagt ggaacttccg aaaggatgtg
20161 aacatggtcc tgcaaagttc ccttggaaat gacctcagaa cggatggtgc taccataagt
20221 ttcaccagca tcaatctcta tgccaccttc ttccccatgg ctcacaacac agcttccacc
20281 cttgaagcca tgctgcgcaa cgataccaat gatcagtcat ttaacgacta cctctctgca
20341 gctaacatgc tttaccccat tcctgccaat gcaaccaaca ttccaatttc catcccatct
20401 cgcaactggg cagccttcag gggctggtcc ttcaccagac tcaaaaccaa ggagactcca
20461 tctcttggat cagggttcga tccctacttc gtatattctg gatctattcc ctacctggat
20521 ggcacctttt accttaacca cactttcaag aaggtctcca tcatgtttga ctcctcagtc
20581 agctggcctg gcaatgacag gctgttgagt ccaaatgagt ttgaaatcaa gcgcactgtg
20641 gacggggaag gatacaacgt ggcacaatgc aacatgacca aagactggtt cctggttcag
20701 atgcttgcca attacaacat tggctaccag ggcttttaca tccctgaggg atacaaggat
20761 cgcatgtact ccttttcag aaacttccag cctatgagca ggcaggtggt tgatgaggtt
20821 aattacactg actacaaagc cgtcaccttа ccataccaac acaacaactc tggctttgta
20881 gggtaccttg cacctactat gagacaaggg gaaccttacc cagccaatta tccatacccg
20941 ctcatcggaa ctactgccgt taagagtgtc acccagaaaa agttcctgtg cgacaggacc
21001 atgtggcgca ttcccttctc cagcaacttc atgtccatgg gggcccttac cgacctggga
21061 cagaacatgc tctatgccaa ctcagcccat gcgctggaca tgacttttga ggtggatccc
21121 atggatgagc ccaccctgct ttatcttctt ttcgaagtct tcgacgtggt cagagtgcac
21181 cagccacacc gcggcgtcat cgaggccgtc tacctgcgca caccgttctc ggccggcaac
21241 gccaccacat aagaagcctc ttgcttcttg caagcagcag ctgcagccat gacatgcggg
21301 tccggaaacg gctccagcga gcaagagctc aaagccatcg tccgagacct gggctgcgga
21361 ccctatttcc tgggaacctt tgacaagcgt ttcccggggt tcatggcccc cgacaagctc
21421 gcctgcgcca tagtcaacac tgccggacgc gagacggggg gagagcactg gctggctttt

21481 ggttggaacc cgcgctacaa cacctgctac ctttttgatc cttttgggtt ctcggatgag
21541 cggctcaaac agatttacca gtttgagtac gaggggctcc tgcgtcgcag tgcccttgct
21601 accaaagacc gctgcatcac cctggagaag tctacccaaa gcgtgcaggg tccgcgctca
21661 gccgcctgtg gacttttttg ctgtatgttc cttcatgcct ttgtgcactg gcccgaccgc
21721 cccatggacg gaaaccccac catgaagttg ctgactgggg tgtccaacag catgctccaa
21781 tcaccccaag tccagcccac cctgcgccgc aaccaggagg tgctataccg cttcctaaac
21841 acccactcat cttactttcg ttctcaccgc gcgcgcattg aaagggccac cgcgtttgac
21901 cgtatggata tgcaataagt catgtaaaac cgtgttcaat aaacagcact ttatttttac
21961 atgcactgag gctctggttt tgctcatttg tttcatcatt tactcagaag tcgaatgggt
22021 tctggcggga gtcagagtga cccgcgggca gggatacgtt gcggaactgt aacctgttct
22081 gccacttgaa ctcggggatt accagcttgg gaactggaat ctcgggaaag gtgtcttgcc
22141 acaactttct ggtcagttgc atagcgccaa gcaggtcagg agcagagatc ttgaaatcac
22201 agttggggcc ggcattctgg acacgggagt tgcgatacac tgggttgcaa cactggaaca
22261 ctatcaacgc tgggtgtctt acgcttgcca acacggttgg gtcactgatg gtagtcacat
22321 ccaagtcttc agcattggcc atcccaaagg gggtcatctt acatgtctgc ctgcccatca
22381 cgggagcgca gcctggcttg tggttgcaat cacaatgaat ggggatcagc atcatcctgg
22441 cttggtcggg agttatccct gggtacacag ccttcatgaa ggcttcgtac tgcttaaaag
22501 cttcctgggc cttacttccc tcggtgtaga acatcccaca ggacttgctg gaaaattgat
22561 tagtagtaca gttggcatca ttcacacaac agcgggcatc gttgttggcc aactgaacca
22621 catttctgcc ccagcggttt tgggtgatct tggctctgtc tggattctcc ttcatagcgc
22681 gctgcccgtt ctcgctcgcc acatccatct cgataatgtg gtccttctgg atcatgatag
22741 tgccatgcag gcatttcacc ttgccttcat aatcggtgca tccatgagcc cacagagcgc
22801 acccggtgca ctcccaatta ttgtgggcga tctcagaata ataatgtacc aatccctgca
22861 tgaatcttcc catcattgtt gtcaaggtct tcatgctggt aaatgtcagc gggatgccac
22921 ggtgctcctc gttcacatac tggtggcaga tacgcttgta ttgctcgtgc tgctctggca
22981 tcagcttgaa agaggttctc agatcattat ccagcctgta cctttccatt agcacagcca
23041 tcacttccat gcccttctcc caggcagata ccaggggcag actcaaagga ttcctaacag
23101 caataaaagt agctccttta gctatagggt cattcttgtc gatcttctca acacttctct
23161 tgccatcctt ctcaatgatg cgcaccgggg ggtagctgaa gcccacggcc accaactgag
23221 cctgttctct ttcttcttca ctatcctggc tgatgtcttg cagagggaca tgcttggtct
23281 tcctgggctt cttcttggga gggatcgggg gaggactgtt gctccgctcc ggagacaggg
23341 atgactgcga agtttcgctt accaatacca cctggctctc ggtagaagaa ccggacccca
23401 cacgacggta ggtgttcctc ttcgggggca gaggtggagg cgactgagat gggctgcggt
23461 ctggccttgg aggcggatgg ctggcagagc tcattccgcg ttcgggggtg tgctcccggt
23521 ggcggtcgct tgactgattt cctccgcggc tggccattgt gttctcctag gcagagaaac
23581 aacagacatg gaaactcagc catcactgcc aacatcgctg caagcaccat cacacctcgc
23641 ccccagcagc gacgaggagg agagcttaac caccccacca cccagtcccg ctaccaccac
23701 ctctaccctc gatgatgagg aggaggtcga cgcagcccag gagatgcagg cgcaggataa
23761 tgtgaaaacg gaagagattg aggcagatgt cgagcaggac ccgggctatg tgacgccggc
23821 ggagcaccag gaggagctga aacgctttct agacagagag gatgacgacc gcccagagca
23881 tcaagcagat ggcgtttacc aggaggctgg gatcagggat catgtcgccg actacctcac
23941 cggccttggt ggggaggacg tgctcctcaa acatctagca aggcagtcga tcatagttaa
24001 agacgcattg ctcgatctca ctgaagtgcc catcagtgtg gaagagctta gccgcgccta
24061 cgagctgaac ctcttttcgc ctcaggtacc ccccaagcgg cagccaaacg gcacctgcga

24121 ggccaaccct cgactcaact tctatccagc ttttactatc cccgaagtgt tggccaccta
24181 ccacatcttt ttcaagaacc aaaagattcc agtctcctgc cgcgccaacc gcacccgcgc
24241 cgatgccctg ctcaacttgg gtccgggagc tcgcttacct gatatagctt ccttggaaga
24301 ggttccaaag atctttgagg gtctgggaag tgatgagaca cgggccgcaa atgctctgca
24361 acagggagag aatgacatgg atgaacacca cagcgctctg gtggaactgg agggtgacaa
24421 tgcccggatt gcagtgctca agcgcagtat cgtggtcacc cattttgcct accccgctgt
24481 taacctgccc cccaaagtta tgagcgctgt catggaccat ctgctcatca aacgagcaag
24541 acctctttca gaaaaccaga acatgcagga tccagacgcc tcggacgagg gcaagccggt
24601 agtcagtgac gagcagctat ctcgctggct gggtaccaac tccccccgag atttggaaga
24661 gaggcgcaag cttatgatgg ctgtagtgct agtaactgtg gagctggagt gtctgcgccg
24721 ctttttcacc gaccctgaga ccctgcgcaa gctagaggag aacctgcact acacctttag
24781 acatggcttc gtgcggcagg catgcaagat ctccaacgtg gagcttacca acctggtttc
24841 ttacatgggc attttgcatg agaaccgact agggcagagc gtcctgcaca ccacccttaa
24901 aggggaggcc cgccgtgact acatccgaga ctgtgtctac ctctacctct gccatacctg
24961 gcaaactggt atgggtgtgt ggcaacagtg tttggaagag cagaacctaa aagagctgga
25021 caagctcttg cagagatccc tcaaagccct gtggacaggt tttgatgagc gcaccgtcgc
25081 ctcggacctg gcagacatca tcttccccga gcgtctcagg gttactctgc gaaacggcct
25141 gccagacttt atgagccaga gcatgcttaa caactttcgc tctttcatcc tggaacgctc
25201 cggtatcctg cctgccacct gctgtgcgct gccctccgac tttgtgcctc tcacctaccg
25261 cgagtgccca ccgccgctat ggagccactg ctacctgttc cgcctggcca actacctctc
25321 ctaccactcg gatgttatag aggatgtgag cggagacggc ctgctggaat gccactgccg
25381 ctgcaatctt tgcacacccc accgctccct tgcctgcaac ccccagttgc tgagcgagac
25441 ccagattatc ggcaccttcg agctgcaggg tcccagaagt aaaggcgagg ggtcttctcc
25501 ggggcagagt ttgaaactga caccggggct gtggacctcc gcctacctgc gcaagtttca
25561 ccccgaggac taccatccct atgagatcag gttctatgag gaccaatcac atcctcccaa
25621 agtcgagctc tcagcctgcg tcatcaccca gggagcaatt ctggcccaat tgcaagccat
25681 ccaaaaatct cgccaagaat ttctgctaaa aaagggaaac ggggtctacc ttgaccctca
25741 gaccggtgag gagctcaaca caaggttccc ccaggatgtc ccatcgccga ggaagcaaga
25801 agttgaaggt gcagctgtcg cccccagagg atatgaagga agactgggac agtcaggcag
25861 aggaggagat ggaagattgg gacagccagg cagaggaggt ggacagcctg gaggaagaca
25921 gtttggagga ggaagacgag gaggcagagg aggtggaaga agcaaccgcc gccaaacagt
25981 tgtcatcggc ggcggagaca agcaagtccc cagacagcag cacggctacc atctccgctc
26041 cgggtcgggg ggcccagcgg cggcccaaca gtagatggga cgagaccggg cgattcccaa
26101 acccgaccac cgcttccaag accggtaaga aggagcgaca gggatacaag tcctggcgtg
26161 gacataaaaa cgctatcatc tcctgcttgc atgaatgcgg gggcaacata tccttcaccc
26221 ggcgatacct gctcttccac cacggtgtaa acttcccccg caatatcttg cattactacc
26281 gtcacctcca cagcccctac tgcagtcagc aagtcccggc aaccccgaca gaaaaataca
26341 gcagcgacaa cggtgaccag aaaaccagca gttagaaaat ccacaacaag tgcaccagga
26401 ggaggactga ggatcacagc gaacgagcca gcgcagacca gagagctgag gaaccggatc
26461 tttccaaccc tctatgccat tttccagcag agtcgggggc aagagcagga actgaaagta
26521 aaaaaccgat ctctgcgctc gctcaccaga agttgtttgt atcacaagag cgaagaccaa
26581 cttcagcgca ctctcgagga cgccgaggct ctcttcaaca agtactgcgc gctgactctt
26641 aaagagtagc ccttgcccgc gctcattttg aaaacggcgg gaatcacgtc acccttggca
26701 cctgtccttt gcccttgtca tgagtaaaga gattcccacg ccttacatgt ggagctatca

26761 gccccaaatg gggttggcag caggcgcttc ccaggactac tccacccgca tgaattggct
26821 tagcgccggg ccctcaatga tatcacgggt taatgatata cgagcttatc gaaaccagtt
26881 actcctagaa cagtcagctc tcaccaccac accccgtcaa caccttaatc cccgaaattg
26941 gcccgccacc ctggtgtacc aggaaaatcc cgctcccacc accgtactac ttcctcgaga
27001 cgcccaggcc gaagttcaga tgactaacgc aggtgtacag ctggcgggcg gttccgccct
27061 atgtcgtcac cgacctcaac agagtataaa acgcctggtg attagaggcc gaggtatcca
27121 gctcaacgac gagtcggtta gctcttcgct tggtctgcga ccagacggag tcttccaaat
27181 cgccggctgt gggagatctt ccttcactcc tcgtcaggct gtgctgactt tggagagttc
27241 gtcctcgcag ccccgctcgg gcggcattgg aactctccag tttgtggagg agtttactcc
27301 ctctgtctac ttcaacccct tctccggctc tcctggccag tacccggacg agttcatacc
27361 aaacttcgac gcaatcagcg agtcagtgga tggctatgat tgatgtctaa tggtggtgcg
27421 gctgagctag ctcgactgcg acacctagac cactgccgcc gctttcgctg cttcgcccgg
27481 gaactcaccg agttcatcta cttcgaactc tccgaggagc accctcaggg tccggcccac
27541 ggagtgcgga ttaccatcga agggggaata gactctcgcc tgcatcgcat cttctcccag
27601 cggcccgtgc taattgaacg cgaccaggga aatacaacca tctccatcta ctgcatctgt
27661 aaccacccccg gattgcatga aagcctttgc tgtcttgttt gtgctgagtt taataaaaac
27721 tgagttaaga ccctcctacg gactaccgct tcttcaatca ggactttaca acaccaacca
27781 gatcttccag aagacccaga cccttcctcc tttcatccag gactctaact ctaccttacc
27841 agcaccctcc actactaacc ttcccgaaac aaacaagctt gcatctcatc tgcaacaccg
27901 cctttcacga agccttcttt ctgccaatac taccactccc aaaaccggag gtgagctccg
27961 cggtcttcct actgacgacc cctgggtggt agcgggtttt gtaacgttag gagtagttgc
28021 gggtgggctt gtgctgatcc tttgctacct atacacacct tgctgtgcat atttagtcat
28081 attgtgctgt tggtttaaga aatgggggcc atactagtcg tgcttgcttt actttcgctt
28141 ttgggtctgg gctctgctaa tctcaatcct ctcgatcacg atccatgttt agacttcgac
28201 ccagaaaact gcacacttac ttttgcaccc gacacaagcc gtctctgtgg agttcttatt
28261 aagtgcggat gggactgcag gtccgttgaa attacacata ataacaaaac atggaacaat
28321 accttatcca ccacatggga gccaggagtt ccccagtggt atactgtctc tgtccgaggt
28381 cctgacggtt ccatccgcat tagtaacaac actttcattt tttctgaaat gtgcgatctg
28441 gccatgttca tgagcagaca gtatgaccta tggcctccca gcaaagagaa cattgtggca
28501 ttttccattg cttattgctt ggtaacatgc atcatcactg ctatcatttg tgtgtgcata
28561 cacttgctta tagttattcg ccctagacaa agcaatgagg aaaaagagaa aatgccttaa
28621 cctttttcct catacctttt ctttacagca tggcttctgt tacagctcta attattgcca
28681 gcattgtcac tgtcgctcac gggcaaacaa ttgtccatat taccttagga cataatcaca
28741 ctcttgtagg gcccccaatt acttcagagg ttatttggac caaacttgga agtgttgatt
28801 attttgatat aatttgcaac aaaactaaac caatatttgt aatctgtaac agacaaaatc
28861 tcacgttaat taatgttagc aaaatttata acggttacta ttatggttat gacagatcca
28921 gtagtcaata taaaaattac ttagttcgca taactcagcc caaattaaca gtgccaacta
28981 tgacaataat taaaatggct aataaagcat tagaaaattt tacatcacca acaacaccca
29041 atgaaaaaaa cattccaaat tcaatgattg caattattgc ggcggtggca ttgggaatgg
29101 cactaataat aatatgcatg ctcctatatg cttgttacta taaaaagttt caacataaac
29161 aggatccact actaaatttt aacatttaat tttttataca gatgatttcc actacaattt
29221 ttatcattac tagccttgca gctgtaactt atggccgttc acacctaact gtacctgttg
29281 gctcaacatg tacactacaa ggaccccaag aaggctatgt cacttggtgg agaatatatg
29341 ataatggagg gttcgctaga ccatgtgatc agcctggtac aaaattttca tgcaacggaa

29401 gagacttgac cattattaac ataacatcaa atgagcaagg cttctattat ggaaccaact
29461 ataaaaatag tttagattac aacattattg tagtgccagc caccacttct gctccccgca
29521 aatccacttt ctctagcagc agtgccaaag caagcacaat tcctaaaaca gcttctgcta
29581 tgttaaagct tccaaaaatc gctttaagta attccacagc cgctcccaat acaattccta
29641 aatcaacaat tggcatcatt actgccgtgg tagtgggatt aatgattata tttttgtgta
29701 taatgtacta cgcctgctgc tatagaaaac atgaacaaaa aggtgatgca ttactaaatt
29761 ttgatattta atttttata gaattatgat attgtttcaa tcaaatacca ctacctccta
29821 tgcatacaca aacattcagc ctaaatacgc tatgcaacta gaaatcacaa tactaattgt
29881 aattggaatt cttatactat ctgttattct ttatttata ttctgccgtc aaatacccaa
29941 tgttcataga aattctaaaa gacgtcccat ctattctcct atgattagtc gtccccatat
30001 ggctctgaat gaaatctaag atctttttt ttttctctta cagtatggtg aacatcaatc
30061 atgatcccta gaaatttctt cttcaccata ctcatctgtg cttttaatgt ctgtgctact
30121 ttcacagcag tagccactgc aagcccagac tgtataggac catttgcttc ctatgcactt
30181 tttgccttcg ttacttgcat ctgcgtgtgt agcatagtct gcctggttat taatttttc
30241 caactggtag actggatctt tgtgcgaatt gcctacctac gtcaccatcc cgaataccgc
30301 aatcaaaatg ttgcggcact tcttaggctt atttaaaacc atgcaggcta tgctaccagt
30361 cattttaatt ttgctactac cctgcattcc cctagcttcc accgccactc gcgctacacc
30421 tgaacaactt agaaaatgca aatttcaaca accatggtca tttcttgatt gctaccatga
30481 aaaatctgat tttcccacat actggatagt gattgttgga ataattaaca tactttcatg
30541 taccttttc tcaatcacaa tataccccac atttaatttt gggtggaatt ctcccaatgc
30601 actgggttac ccacaagaac cagatgaaca tattccacta caacacatac aacaaccact
30661 agcactggta cagtatgaaa atgagccaca accttcactg ccccctgcca ttagttactt
30721 caacctaacc ggcggagatg actgacccaa tcgccacatc atccaccgct gccaaggagc
30781 tgctggacat ggacggacgt gcctcagaac agcgactcat ccaactacgc attcgtcagc
30841 agcaggaacg agcagtaaaa gagctaaggg atgccattgg gattcaccag tgcaaaaaag
30901 gcatattctg cttagtaaaa caatccaaaa tctcctacga gatcaccgct actgaccatc
30961 gtctctcata cgagctcggt ccgcagcgac aaaaattcac ctgcatggtg ggaatcaacc
31021 ccatagttat cacccagcag tctggagata ctaagggttg tatccagtgt tcctgtgatt
31081 ccaccgagtg catctacaca ctgctgaaga ccctctgcgg ccttcgagac ctcctaccca
31141 tgaactaatc attgccccta ccttacccaa tcaaaatatt aataaagaca cttacttgaa
31201 atcagcaata cagtctttgt caaaactttc taccagcagc acctcaccct cttcccaact
31261 ctggtactct aaacgtcgga gggtggcata ctttctccac actttgaaag ggatgtcaaa
31321 ttttatttcc tcttctttgc ccacaatctt catttcttta tccccagatg gccaagcgag
31381 ctcggctaag cacttccttc aacccggtgt acccttatga agatgaaagc agctcacaac
31441 acccatttat aaatcctggt ttcatttccc ctgacgggtt cacacaaagt ccaaacgggg
31501 ttttaagtct taaatgtgtt aatccactta ccactgcaag cggctccctc caacttaaag
31561 tgggaagtgg tcttacagta gacactactg atggatcctt agaagaaaac atcaaagtta
31621 acacccccct aacaaagtca aaccattcta taaatttacc aataggaaac ggtttgcaaa
31681 tagaacaaaa caaactttgc agtaaactcg gaaatggtct tacatttgac tcttccaatt
31741 ctattgcact gaaaaataac actttatgga caggtccaaa accagaagcc aactgcataa
31801 ttgaatacgg gaaacaaaac ccagatagca aactaacttt aatccttgta aaaaatggag
31861 gaattgttaa tggatatgta acgctaatgg gagcctcaga ctacgttaac accttattta
31921 aaaacaaaaa tgtctccatt aatgtagaac tatactttga tgccactggt catatattac
31981 cagactcatc ttctcttaaa acagatctag aactaaaata caagcaaacc gctgacttta

32041 gtgcaagagg ttttatgcca agtactacag cgtatccatt tgtccttcct aatgcgggaa
32101 cacataatga aaattatatt tttggtcaat gctactacaa agcaagcgat ggtgcccttt
32161 ttccgttgga agttactgtt atgcttaata aacgcctgcc agatagtcgc acatcctatg
32221 ttatgacttt tttatggtcc ttgaatgctg gtctagctcc agaaactact caggcaaccc
32281 tcataacctc cccatttacc ttttcctata ttagagaaga tgactgacaa caaaaataaa
32341 gttcaacatt ttttattgaa attcctttta cagtattcga gtagttattt tgcctccccc
32401 ttcccattta acagaataca ccaatctctc cccacgcaca gctttaaaca tttggatacc
32461 attagagata gacatagttt tagattccac attccaaaca gtttcagagc gagccaatct
32521 ggggtcagta atacataaaa atgcatcggg atagtctttt aaagcgcttt cacagtccaa
32581 ctgttgcgga tgcgactccg gagtctgaat cacggtcatc tggaagaaga acgatgggaa
32641 tcataatccg aaaacggaat cgggcgattg tgtctcatca aacccacaag caaccgctgt
32701 ctgcgtcgct ccgtgcgact gctgtttatg ggatcggggt ccgcagtgtc ctgaagcatg
32761 attttaatag cccttaacat taactttctg gtgcgatgcg cgcagcaacg cattctgatt
32821 tcacttagat tactacagta ggtacagcac attatcacaa tattgtttaa taaaccataa
32881 ttaaaagcgc tccagccaaa actcatatct gatataatcg cccctgcatg accatcatac
32941 caaagtttaa tataaattaa atgtcgttcc ctcaaaaaca cactacccac atacatgatc
33001 tcttttggca tgtgcatatt aacaatttgt ctgtaccatg gacaacgttg gttaatcatg
33061 caacccaata taaccttccg gaaccacact gccaacaccg ctccccccagc catgcattga
33121 agtgaaccct gctgattaca atgacaatga agaacccaat tctctcgacc atgaatcact
33181 tgagactgaa aaatatctat agtagcacaa caaagacata aatgcatgca tcttctcata
33241 atttttaact cctctggatt taaaaacata tcccaaggaa tgggaaactc ttgcagaaca
33301 gtaaagctgg cagaacaagg aagaccacga acacaactta cactatgcat agtcatagta
33361 tcacaatctg gcaacagcgg gtggtcttca gtcatagaag ctcgggtttc attttcctca
33421 catcgtggta attgggctct ggtgtaaggg tgatgtctgg cgcatgatgt ggagcgtgcg
33481 cgcaaccttg tcataatgga gttgcttcct gacattctcg tattttgtat agcaaaacgc
33541 tgccctggca caacacactc ttcttcgtct tctatcctgc cgcttagtgt gttccgtctg
33601 ataattcaag tacagccaca ctcttaagtt ggtcaaaaga atgctggctt cagttgtaat
33661 caaaactcca tcatatttaa ttgttctaag gaaatcatcc acggtagcat atgcaaatcc
33721 caaccaagca atgcaactgg attgcgtttc aagcagcaga ggagagggaa gagacggaag
33781 aatcatgtta atttttattc caaacgatct cgcagtactt caaattgtag atcgcgcaga
33841 tggcatctat cgcccccact gtgttggtga aaaagcacag ctaaatcaaa agaaatgcga
33901 ttttcaaggt gctcaacggt ggcttccaac aaagcctcca cgcgcacatc caaaaacaaa
33961 agaataccaa aagaaggagc attttctaac tcctcaaaca tcatattaca ttcctgcacc
34021 attcccagat aattttcagc tttccagcct tgaattattc gtgtcagttc ttgtggtaaa
34081 tccaaaccac acattacaaa caggtcccgg agggcgccct ccaccaccat tcttaaacac
34141 accctcataa tgacaaaata tcttgctcct gtgtcacctg tagcaaatta agaatggcat
34201 catcaattga catgcccttg gctctaagtt cttctctaag ttctagttgt agatactctc
34261 tcatattatc accaaactgc ttagccagaa gccccccggg aacaatagca ggggacgcta
34321 cagtgcagta caagcgcaga cctccccaat tggctccagc aaaaacaaga ttagaataag
34381 catactggga accaccagta atatcatcaa agttgctgga aatataatca ggcagagttt
34441 cttgtaaaaa ttgaataaaa gaaaaatttt ccaaagaaac attcaaaacc gttgggatgc
34501 aaatacaata ggttaccgcg ctgcgctcca acattgttag ttttgaatta gtctgcaaaa
34561 taaaagaaac aagcgtcata tcatagtagc ctgtcgaaca ggtggaaaaa tcagtctttc
34621 catcacaaga caagccacag ggtctccagc tcgaccctcg taaaacctgt cattgtgatt

34681 aaacaacagc accgaaagtt cctcgcggtg gccagcatga ataattcttg atgaagcata
34741 caatccagac atgttagcat cagttaaaga gaaaaaacag ccaacatagc ctctgggtat
34801 aattatgctt aattttaagt atagcaaagc cacccctcgc ggatacaaag taaaaggcac
34861 aggagaataa aaaatataat tatttctctg ctgctgttca ggcaacgttg ctcccggtcc
34921 ctctaaatag acatacaaag cctcatcagc catggcttac caggcaaagt acagcgggcg
34981 cacaaagcac aagctctaaa gaagctctaa aaacactctc caacctctcc acaatatata
35041 cacaagccct aaactgacgt aatgggagta aagtgaaaaa aaaataccgc caagcccaac
35101 acacaccccg aaactgcgtc agcaggaaaa agtacagttt cacttccgca ttcccaacaa
35161 gcgtaacttc ctctttctca tggtacgtca catccgatta acttgcaacg tcattttccc
35221 acggtcgcgc cgccccttt agccgttaac cccgcagcca atcaccacac agcgcgcact
35281 tttttaaatt acctcattta catgttggca ccattccatc tataaggtat attatataga
35341 tag