**BIOINF101 - Aufgabe 5**

*Human T-cell leukemia virus type I (NC\_001436)*

Genom

GGCTCGCATCTCTCCTTCACGCGCCCGCCGCCTTACCTGAGGCCGCCATCCACGCCGGTTGAGTCGCGTT

CTGCCGCCTCCCGCCTGTGGTGCCTCCTGAACTACGTCCGCCGTCTAGGTAAGTTTAGAGCTCAGGTCGA

GACCGGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGAC

CCTGCTTGCTCAACTCTACGTCTTTGTTTCGTTTTCTGTTCTGCGCCGTTACAGATCGAAAGTTCCACCC

CTTTCCCTTTCATTCACGACTGACTGCCGGCTTGGCCCACGGCCAAGTACCGGCAACTCTGCTGGCTCGG

AGCCAGCGACAGCCCATTCTATAGCACTCTCCAGGAGAGAAATTTAGTACACAGTTGGGGGCTCGTCCGG

GATTCGAGCGCCCCTTTATTCCCTAGGCAATGGGCCAAATCTTTTCCCGTAGCGCTAGCCCTATTCCGCG

GCCGCCCCGGGGGCTGGCCGCTCATCACTGGCTTAACTTCCTCCAGGCGGCATATCGCCTAGAACCCGGT

CCCTCCAGTTACGATTTCCACCAGTTAAAAAAATTTCTTAAAATAGCTTTAGAAACACCGGTCTGGATCT

GCCCCATTAACTACTCCCTCCTAGCCAGCCTACTCCCAAAAGGATACCCCGGCCGGGTGAATGAAATTTT

ACACATACTCATCCAAACCCAAGCCCAGATCCCGTCCCGCCCCGCGCCGCCGCCGCCGTCATCCTCCACC

CACGACCCCCCGGATTCTGACCCACAAATCCCCCCTCCCTATGTTGAGCCTACAGCCCCCCAAGTCCTTC

CAGTCATGCACCCACATGGTGCCCCTCCCAACCACCGCCCATGGCAAATGAAAGACCTACAGGCCATTAA

GCAAGAAGTCTCCCAAGCGGCCCCTGGAAGCCCCCAGTTTATGCAGACCATCCGGCTTGCGGTGCAGCAG

TTTGACCCCACTGCCAAAGACCTCCAAGACCTCCTGCAGTACCTTTGCTCCTCCCTCGTGGCTTCCCTCC

ATCACCAGCAGCTAGATAGCCTTATATCAGAGGCCGAAACTCGAGGTATTACAGGTTATAACCCCTTAGC

CGGTCCCCTCCGTGTCCAAGCCAACAATCCACAACAACAAGGATTAAGGCGAGAATACCAGCAACTCTGG

CTCGCCGCCTTCGCCGCCCTGCCAGGGAGTGCCAAAGACCCTTCCTGGGCCTCTATCCTCCAAGGCCTGG

AGGAGCCTTACCACGCCTTCGTAGAACGCCTCAACATAGCTCTTGACAATGGGCTGCCAGAAGGCACGCC

CAAAGACCCCATTTTACGTTCCTTAGCCTACTCTAATGCAAACAAAGAATGCCAAAAATTACTACAGGCC

CGAGGGCACACTAATAGCCCTCTAGGAGATATGTTGCGGGCTTGTCAGGCCTGGACCCCCAAAGACAAAA

CCAAAGTGTTAGTTGTCCAGCCTAAAAAACCCCCCCCAAATCAGCCGTGCTTCCGGTGCGGGAAAGCAGG

CCACTGGAGTCGGGACTGCACTCAGCCTCGTCCTCCCCCTGGGCCATGCCCCCTATGTCAAGATCCAACT

CACTGGAAGCGAGACTGCCCCCGCCTAAAGCCCACTATCCCAGAACCAGAGCCAGAGGAGGATGCCCTCC

TATTAGATCTCCCCGCCGACATCCCACACCCAAAAAACTCCATAGGGGGGGAGGTTTAACCTCCCCCCCC

ACATTACAGCAAGTCCTTCCTAACCAAGACCCAACATCTATTCTGCCAGTTATACCGTTAGATCCCGCCC

GTCGGCCCGTAATTAAAGCCCAGATTGACACCCAGACCAGCCACCCAAAGACTATCGAAGCTCTACTAGA

TACAGGAGCAGACATGACAGTCCTTCCGATAGCCTTGTTCTCAAGTAATACTCCCCTCAAAAACACATCC

GTGTTAGGGGCAGGGGGCCAAACCCAAGATCACTTTAAGCTCACCTCCCTTCCTGTGCTAATACGCCTCC

CTTTCCGGACGACGCCTATTGTTTTAACATCTTGCCTAGTTGATACCAAAAACAACTGGGCCATCATAGG

TCGTGATGCCTTACAACAATGCCAAGGCGTCCTGTACCTCCCTGAGGCAAAAAGGCCGCCTGTAATCTTG

CCAATACAGGCGCCAGCTGTCCTTGGGCTAGAACACCTCCCAAGGCCCCCCGAAATCAGCCAGTTCCCTT

TAAACCAGAACGCCTCCAGGCCTTGCAACACTTGGTCCGGAAGGCCCTGGAGGCAGGCCATATCGAACCC

TACACCGGGCCAGGAAATAACCCAGTATTCCCAGTTAAAAAAGCCAATGGAACCTGGCGATTCATCCACG

ACCTGCGGGCCACTAACTCTCTAACCATAGATCTCTCATCATCTTCCCCCGGGCCCCCTGACTTGTCCAG

CCTGCCAACTACACTAGCCCACTTACAAACTATAGACCTTAAAGACGCCTTTTTCCAAATCCCCCTACCT

AAACAGTTCCAGCCCTACTTTGCTTTCACTGTCCCACAGCAGTGTAACTACGGCCCCGGCACTAGATACG

CCTGGAGAGTACTACCCCAAGGGTTTAAAAATAGTCCCACCCTGTTCGAAATGCAGCTGGCCCATATCCT

GCAGCCCATTCGGCAAGCCTTCCCCCAATGCACTATTCTTCAGTACATGGATGACATTCTCCTGGCAAGC

CCCTCCCATGCGGACCTGCAACTACTCTCAGAGGCCACAATGGCTTCCCTAATCTCCCATGGGTTGCCTG

TGTCCGAAAACAAAACCCAGCAAACCCCTGGAACAATTAAGTTCCTAGGGCAAATAATTTCACCTAATCA

CCTCACTTATGATGCAGTCCCCAAGGTACCTATACGGTCCCGCTGGGCGCTACCTGAACTTCAAGCCCTA

CTTGGCGAGATTCAGTGGGTCTCCAAAGGAACTCCTACCTTACGCCAGCCCCTTCACAGTCTCTACTGTG

CCTTACAAAGGCATACTGATCCCCGAGACCAAATATATTTAAATCCTTCTCAAGTTCAATCATTAGTGCA

GCTGCGGCAGGCCCTGTCACAGAACTGCCGCAGTAGACTAGTCCAAACCCTGCCCCTCCTAGGGGCTATT

ATGCTGACCCTCACTGGCACCACCACTGTGGTGTTCCAGTCCAAGCAGCAGTGGCCACTTGTCTGGCTAC

ATGCCCCCCTACCCCACACTAGCCAGTGCCCCTGGGGGCAGCTACTTGCCTCAGCTGTGTTATTACTCGA

CAAATACACCTTGCAATCCTATGGACTACTCTGCCAAACCATACATCATAACATCTCCACCCAAACCTTC

AACCAATTCATTCAAACATCTGACCACCCCAGTGTTCCTATCTTACTCCACCACAGTCACCGATTCAAAA

ATTTAGGTGCCCAGACTGGAGAACTTTGGAACACTTTTCTTAAAACAACTGCCCCATTGGCTCCTGTGAA

AGCCCTTATGCCAGTGTTTACTCTTTCCCCTGTGATCATAAACACCGCCCCTTGCCTGTTTTCAGACGGA

TCCACCTCCCAGGCAGCCTATATTCTCTGGGACAAGCATATATTGTCACAAAGATCATTCCCCCTTCCGC

CACCGCACAAGTCGGCCCAACGGGCCGAACTTCTCGGACTTTTGCATGGCCTCTCCAGCGCCCGTTCGTG

GCGCTGTCTCAACATATTTCTAGACTCCAAGTATCTTTATCATTACCTTCGGACCCTTGCCCTAGGCACC

TTCCAAGGCAGGTCCTCTCAGGCCCCCTTTCAGGCCCTCCTGCCCCGCTTACTATCGCGTAAGGTCGTCT

ATTTGCACCACGTTCGCAGCCATACCAATCTACCTGATCCCATCTCCAGGCTCAACGCTCTCACAGATGC

CCTACTAATCACCCCTGTCCTGCAGCTCTCTCCTGCAGACCTACACAGTTTCACCCATTGCGGACAGACG

GCCCTCACACTGCAAGGGGCAACCACAACTGAGGCCTCCAATATCCTGCGCTCTTGCCACGCCTGCCGCA

AAAATAACCCACAACATCAGATGCCTCAAGGACACATCCGCCGTGGCCTACTCCCTAACCACATCTGGCA

AGGCGACATTACCCATTTCAAATATAAAAATACACTGTATCGCCTTCATGTATGGGTAGACACCTTTTCA

GGAGCCATCTCAGCTACCCAAAAGAGAAAAGAAACAAGCTCAGAAGCTATTTCCTCTTTGCTCCAGGCCA

TTGCCTATCTAGGCAAGCCTAGCTACATAAACACAGACAATGGCCCTGCCTATATTTCCCAAGACTTCCT

CAATATGTGTACCTCCCTTGCTATTCGCCATACTACCCATGTCCCCTACAATCCAACCAGCTCCGGACTT

GTAGAACGCTCTAATGGCATTCTTAAAACCCTATTATATAAGTACTTTACTGACAAACCCGACCTACCTA

TGGATAATGCTCTATCCATAGCCCTATGGACAATCAACCACCTAAATGTATTAACCAACTGCCACAAAAC

CCGATGGCAGCTTCACCACTCCCCCCGACTCCAGCCGATCCCAGAGACACATTCCCTCAGCAATAAACAA

ACCCATTGGTATTATTTCAAGCTTCCTGGTCTTAATAGCCGCCAGTGGAAAGGACCACAGGAGGCTCTTC

AAGAAGCTGCCGGCGCTGCTCTCATCCCGGTAAGCGCTAGTTCTGCCCAGTGGATCCCGTGGAGGCTCCT

CAAGCGAGCTGCATGCCCAAGACCCGTCGGAGGCCCCGCCGATCCCAAAGAAAAAGACCACCAACACCAT

GGGTAAGTTTCTCGCCACTTTGATTTTATTCTTCCAGTTCTGCCCCCTCATCCTCGGTGATTACAGCCCC

AGCTGCTGTACTCTCACAGTTGGAGTCTCCTCATACCACTCTAAACCCTGCAATCCTGCCCAGCCAGTTT

GTTCATGGACCCTCGACCTGCTGGCCCTTTCAGCAGATCAGGCCCTACAGCCACCCTGCCCTAATCTAGT

AAGTTACTCCAGCTACCATGCCACCTATTCCCTATATCTATTCCCTCATTGGATCAAAAAGCCAAACCGA

AATGGCGGAGGCTATTATTCAGCCTCTTATTCAGACCCTTGTTCCTTAAAATGCCCATACCTAGGGTGCC

AATCATGGACCTGCCCCTATACAGGAGCCGTCTCCAGCCCCTACTGGAAATTTCAGCAAGATGTCAATTT

TACTCAAGAAGTTTCACACCTCAATATTAATCTCCATTTTTCAAAATGCGGTTTTTCCTTCTCCCTTCTA

GTCGACGCTCCAGGATATGACCCCATCTGGTTCCTTAATACCGAACCCAGCCAACTGCCTCCCACCGCCC

CTCCTCTACTCTCCCACTCTAACCTAGACCATATCCTCGAGCCCTCTATACCATGGAAATCAAAACTCCT

GACTCTTGTCCAGTTAACCCTACAAAGCACTAATTATACTTGCATTGTCTGTATCGATCGTGCCAGCCTA

TCCACTTGGCACGTCCTATACTCTCCCAACGTCTCTGTTCCATCCCCTTCTTCTACCCCCCTCCTTTACC

CATCGTTAGCGCTTCCAGCCCCCCACCTGACGTTACCATTTAACTGGACCCACTGCTTTGACCCCCAGAT

TCAAGCTATAGTCTCCTCCCCCTGTCATAACTCCCTCATCCTGCCCCCCTTTTCCTTGTCACCTGTTCCC

ACGCTAGGATCCCGCTCCCGCCGAGCAGTACCGGTGGCGGTCTGGCTTGTCTCCGCCCTGGCCATGGGAG

CCGGAGTGGCTGGCAGGATTACCGGCTCCATGTCCCTCGCCTCAGGAAAGAGCCTCCTACATGAGGTGGA

CAAAGATATTTCCCAATTAACTCAAGCAATAGTCAAAAACCACAAAAATCTGCTCAAAATTGCACAGTAT

GCTGCCCAGAACAGACGAGGCCTTGATCTCCTGTTCTGGGAGCAAGGAGGATTATGCAAAGCATTACAAG

AACAGTGCTGTTTTCTAAATATTACTAATTCCCATGTCTCAATACTACAAGAGAGACCCCCCCTTGAAAA

TCGAGTCCTGACTGGCTGGGGCCTTAACTGGGACCTTGGCCTCTCACAGTGGGCTCGAGAAGCCTTACAA

ACTGGAATCACCCTTGTCGCGCTACTCCTTCTTGTTATCCTTGCAGGACCATGCATCCTCCGTCAGCTAC

GACACCTCCCCTCGCGCGTCAGATACCCCCATTACTCTCTTATAAACCCTGAGTCATCCCTGTAAACCAA

GCACACAATTATTGCAACCACATCGCCTCCAGCCTCCCCTGCCAATAATTAACCTCTCCCATCAAATCCT

CCTTCTCCTGCAGCAACCTCCTCCGTTCAGCCTCCAAGGACTCCACCTCGCCTTCCAACTGTCTAGTATA

GCCATCAACCCCCAACTCCTGCATTTTTTCTTTCCTAGCACTATGCTGTTTCGCCTTCTCAGCCCCTTGT

CTCCACTTGCGCTCACGGCGCTCCTGCTCTTCCTGCTTTCTCCGGGCGAAGTCAGCGGCCTTCTCCTCCG

CCCGCTTCCTGCGCCGTGCCTTCTCCTCTTCCTTCCTTTTCAAATACTCAGCAATCTGCTTTTCCTCCTC

TTTCTCCCGCTCTTTTTTTCGCTTCCTCTTCTCCTCAGCCCGTCGCTGCCGATCACGATGCGTTTCCCCG

CGAGGTGGCGCTTTCCCCCCTGGAGGGCCCCGTCGCAGCCGGCCGCGGCTTTCCTCTTCTAGAGATAGCA

AACCGTCAAGCACAGTTTCCTCCTCCTCCTTGTCCTTTAACTCTTCCTCCAAGGATAATAGCCCGTCCAC

CAATTCCTCCACCAGCAGGTCCTCCGGGCATGGAACAGGCAAACATCGAAACAGCCCTACGGATACAAAG

TTAACCATGCTTATTATCAGCCCACTTCCCAGGGTTTGGACAGAGTCTTCTTTTCGGATACCCAGTCTAC

GTGTTTGGAGACTGTGTACAAGGCGACTGGTGCCCCATCTCTGGGGGACTATGTTCGGCCCGCCTACATC

GTCACGCCCTACTGGCCACCTGTCCAGAGCATCAGATCACCTGGGACCCCATCGATGGACGCGTTATCGG

CTCAGCTCTACAGTTCCTTATCCCTCGACTCCCCTCCTTCCCCACCCAGAGAACCTCTAAGACCCTTAAG

GTCCTTACCCCGCCAATCACTCATACAACCCCCAACATTCCACCCTCCTTCCTCCAGGCCATGCGCAAAT

ACTCCCCCTTCCGAAATGGATACATGGAACCCACCCTTGGGCAGCACCTCCCAACCCTGTCTTTTCCAGA

CCCCGGACTCCGGCCCCAAAACCTGTACACCCTCTGGGGAGGCTCCGTTGTCTGCATGTACCTCTACCAG

CTTTCCCCCCCCATCACCTGGCCCCTCCTGCCCCATGTGATTTTTTGCCACCCCGGCCAGCTCGGGGCCT

TCCTCACCAATGTTCCCTACAAACGAATAGAAAAACTCCTCTATAAAATTTCCCTTACCACAGGGGCCCT

AATAATTCTACCCGAGGACTGTTTGCCCACCACCCTTTTCCAGCCTGCTAGGGCACCCGTCACGCTGACA

GCCTGGCAAAACGGCCTCCTTCCGTTCCACTCAACCCTCACCACTCCAGGCCTTATTTGGACATTTACCG

ATGGCACGCCTATGATTTCCGGGCCCTGCCCTAAAGATGGCCAGCCATCTTTAGTACTACAGTCCTCCTC

CTTTATATTTCACAAATTTCAAACCAAGGCCTACCACCCCTCATTTCTACTCTCACACGGCCTCATACAG

TACTCTTCCTTTCATAATTTGCATCTCCTATTTGAAGAATACACCAACATCCCCATTTCTCTACTTTTTA

ACGAAAAAGAGGCAGATGACAATGACCATGAGCCCCAAATATCCCCCGGGGGCTTAGAGCCTCTCAGTGA

AAAACATTTCCGTGAAACAGAAGTCTGAGAAGGTCAGGGCCCAGAATAAGGCTCTGACGTCTCCCCCCGG

AGGACAGCTCAGCACCAGCTCAGGCTAGGCCCTGACGTGTCCCCCTAAAGACAAATCATAAGCTCAGACC

TCCGGGAAGCCACCGGGAACCACCCATTTCCTCCCCATGTTTGTCAAGCCGTCCTCAGGCGTTGACGACA

ACCCCTCACCTCAAAAAACTTTTCATGGCACGCATACGGCTCAATAAAATAACAGGAGTCTATAAAAGCG

TGGGGACAGTTCAGGAGGGGGCTCGCATCTCTCCTTCACGCGCCCGCCGCCTTACCTGAGGCCGCCATCC

ACGCCGGTTGAGTCGCGTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAACTACGTCCGCCGTCTAGGTA

AGTTTAGAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGC

TCTCCACGCTTTGCCTGACCCTGCTTGCTCAACTCTA

Peptidsequenz

5'3' Frame 1

GSHLSFTRPPPYLRPPSTPVESRSAASRLWCLLNYVRRLGKFRAQVETGPLSGAPLEPT-TQPALHALPDPACSTLRLCFVFCSAPLQIESSTPFPFIHD-LPAWPTAKYRQLCWLGASDSPFYSTLQERNLVHSWGLVRDSSAPLFPRQWAKSFPVALALFRGRPGGWPLITGLTSSRRHIA-NPVPPVTISTS-KNFLK-L-KHRSGSAPLTTPS-PAYSQKDTPAG-MKFYTYSSKPKPRSRPAPRRRRRHPPPTTPRILTHKSPLPMLSLQPPKSFQSCTHMVPLPTTAHGK-KTYRPLSKKSPKRPLEAPSLCRPSGLRCSSLTPLPKTSKTSCSTFAPPSWLPSITSS-IALYQRPKLEVLQVITP-PVPSVSKPTIHNNKD-GENTSNSGSPPSPPCQGVPKTLPGPLSSKAWRSLTTPS-NAST-LLTMGCQKARPKTPFYVP-PTLMQTKNAKNYYRPEGTLIAL-EICCGLVRPGPPKTKPKC-LSSLKNPPQISRASGAGKQATGVGTALSLVLPLGHAPYVKIQLTGSETAPA-SPLSQNQSQRRMPSY-ISPPTSHTQKTP-GGRFNLPPHITASPS-PRPNIYSASYTVRSRPSARN-SPD-HPDQPPKDYRSSTRYRSRHDSPSDSLVLK-YSPQKHIRVRGRGPNPRSL-AHLPSCANTPPFPDDAYCFNILPS-YQKQLGHHRS-CLTTMPRRPVPP-GKKAACNLANTGASCPWARTPPKAPRNQPVPFKPERLQALQHLVRKALEAGHIEPYTGPGNNPVFPVKKANGTWRFIHDLRATNSLTIDLSSSSPGPPDLSSLPTTLAHLQTIDLKDAFFQIPLPKQFQPYFAFTVPQQCNYGPGTRYAWRVLPQGFKNSPTLFEMQLAHILQPIRQAFPQCTILQYMDDILLASPSHADLQLLSEATMASLISHGLPVSENKTQQTPGTIKFLGQIISPNHLTYDAVPKVPIRSRWALPELQALLGEIQWVSKGTPTLRQPLHSLYCALQRHTDPRDQIYLNPSQVQSLVQLRQALSQNCRSRLVQTLPLLGAIMLTLTGTTTVVFQSKQQWPLVWLHAPLPHTSQCPWGQLLASAVLLLDKYTLQSYGLLCQTIHHNISTQTFNQFIQTSDHPSVPILLHHSHRFKNLGAQTGELWNTFLKTTAPLAPVKALMPVFTLSPVIINTAPCLFSDGSTSQAAYILWDKHILSQRSFPLPPPHKSAQRAELLGLLHGLSSARSWRCLNIFLDSKYLYHYLRTLALGTFQGRSSQAPFQALLPRLLSRKVVYLHHVRSHTNLPDPISRLNALTDALLITPVLQLSPADLHSFTHCGQTALTLQGATTTEASNILRSCHACRKNNPQHQMPQGHIRRGLLPNHIWQGDITHFKYKNTLYRLHVWVDTFSGAISATQKRKETSSEAISSLLQAIAYLGKPSYINTDNGPAYISQDFLNMCTSLAIRHTTHVPYNPTSSGLVERSNGILKTLLYKYFTDKPDLPMDNALSIALWTINHLNVLTNCHKTRWQLHHSPRLQPIPETHSLSNKQTHWYYFKLPGLNSRQWKGPQEALQEAAGAALIPVSASSAQWIPWRLLKRAACPRPVGGPADPKEKDHQHHG-VSRHFDFILPVLPPHPR-LQPQLLYSHSWSLLIPL-TLQSCPASLFMDPRPAGPFSRSGPTATLP-SSKLLQLPCHLFPISIPSLDQKAKPKWRRLLFSLLFRPLFLKMPIPRVPIMDLPLYRSRLQPLLEISARCQFYSRSFTPQY-SPFFKMRFFLLPSSRRSRI-PHLVP-YRTQPTASHRPSSTLPL-PRPYPRALYTMEIKTPDSCPVNPTKH-LYLHCLYRSCQPIHLARPILSQRLCSIPFFYPPPLPIVSASSPPPDVTI-LDPLL-PPDSSYSLLPLS-LPHPAPLFLVTCSHARIPLPPSSTGGGLACLRPGHGSRSGWQDYRLHVPRLRKEPPT-GGQRYFPINSSNSQKPQKSAQNCTVCCPEQTRP-SPVLGARRIMQSITRTVLFSKYY-FPCLNTTRETPP-KSSPDWLGP-LGPWPLTVGSRSLTNWNHPCRATPSCYPCRTMHPPSATTPPLARQIPPLLSYKP-VIPVNQAHNYCNHIASSLPCQ-LTSPIKSSFSCSNLLRSASKDSTSPSNCLV-PSTPNSCIFSFLALCCFAFSAPCLHLRSRRSCSSCFLRAKSAAFSSARFLRRAFSSSFLFKYSAICFSSSFSRSFFRFLFSSARRCRSRCVSPRGGAFPPGGPRRSRPRLSSSRDSKPSSTVSSSSLSFNSSSKDNSPSTNSSTSRSSGHGTGKHRNSPTDTKLTMLIISPLPRVWTESSFRIPSLRVWRLCTRRLVPHLWGTMFGPPTSSRPTGHLSRASDHLGPHRWTRYRLSSTVPYPSTPLLPHPENL-DP-GPYPANHSYNPQHSTLLPPGHAQILPLPKWIHGTHPWAAPPNPVFSRPRTPAPKPVHPLGRLRCLHVPLPAFPPHHLAPPAPCDFLPPRPARGLPHQCSLQTNRKTPL-NFPYHRGPNNSTRGLFAHHPFPAC-GTRHADSLAKRPPSVPLNPHHSRPYLDIYRWHAYDFRALP-RWPAIFSTTVLLLYISQISNQGLPPLISTLTRPHTVLFLS-FASPI-RIHQHPHFSTF-RKRGR-Q-P-APNIPRGLRASQ-KTFP-NRSLRRSGPRIRL-RLPPEDSSAPAQARP-RVPLKTNHKLRPPGSHREPPISSPCLSSRPQALTTTPHLKKLFMARIRLNKITGVYKSVGTVQEGARISPSRARRLT-GRHPRRLSRVLPPPACGAS-TTSAV-VSLELRSRPGLCPALPWSLPRLSRLSTLCLTLLAQL

5'3' Frame 2

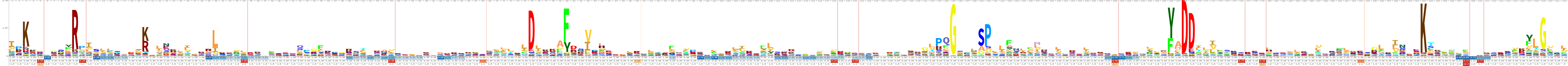
ARISPSRARRLT-GRHPRRLSRVLPPPACGAS-TTSAV-VSLELRSRPGLCPALPWSLPRLSRLSTLCLTLLAQLYVFVSFSVLRRYRSKVPPLSLSFTTDCRLGPRPSTGNSAGSEPATAHSIALSRREI-YTVGGSSGIRAPLYSLGNGPNLFP-R-PYSAAAPGAGRSSLA-LPPGGISPRTRSLQLRFPPVKKIS-NSFRNTGLDLPH-LLPPSQPTPKRIPRPGE-NFTHTHPNPSPDPVPPRAAAAVILHPRPPGF-PTNPPSLC-AYSPPSPSSHAPTWCPSQPPPMANERPTGH-ARSLPSGPWKPPVYADHPACGAAV-PHCQRPPRPPAVPLLLPRGFPPSPAAR-PYIRGRNSRYYRL-PLSRSPPCPSQQSTTTRIKARIPATLARRLRRPARECQRPFLGLYPPRPGGALPRLRRTPQHSS-QWAARRHAQRPHFTFLSLL-CKQRMPKITTGPRAH--PSRRYVAGLSGLDPQRQNQSVSCPA-KTPPKSAVLPVRESRPLESGLHSASSSPWAMPPMSRSNSLEARLPPPKAHYPRTRARGGCPPIRSPRRHPTPKKLHRGGGLTSPPTLQQVLPNQDPTSILPVIPLDPARRPVIKAQIDTQTSHPKTIEALLDTGADMTVLPIALFSSNTPLKNTSVLGAGGQTQDHFKLTSLPVLIRLPFRTTPIVLTSCLVDTKNNWAIIGRDALQQCQGVLYLPEAKRPPVILPIQAPAVLGLEHLPRPPEISQFPLNQNASRPCNTWSGRPWRQAISNPTPGQEITQYSQLKKPMEPGDSSTTCGPLTL-P-ISHHLPPGPLTCPACQLH-PTYKL-TLKTPFSKSPYLNSSSPTLLSLSHSSVTTAPALDTPGEYYPKGLKIVPPCSKCSWPISCSPFGKPSPNALFFSTWMTFSWQAPPMRTCNYSQRPQWLP-SPMGCLCPKTKPSKPLEQLSS-GK-FHLITSLMMQSPRYLYGPAGRYLNFKPYLARFSGSPKELLPYASPFTVSTVPYKGILIPETKYI-ILLKFNH-CSCGRPCHRTAAVD-SKPCPS-GLLC-PSLAPPLWCSSPSSSGHLSGYMPPYPTLASAPGGSYLPQLCYYSTNTPCNPMDYSAKPYIITSPPKPSTNSFKHLTTPVFLSYSTTVTDSKI-VPRLENFGTLFLKQLPHWLL-KPLCQCLLFPL-S-TPPLACFQTDPPPRQPIFSGTSIYCHKDHSPFRHRTSRPNGPNFSDFCMASPAPVRGAVSTYF-TPSIFIITFGPLP-APSKAGPLRPPFRPSCPAYYRVRSSICTTFAAIPIYLIPSPGSTLSQMPY-SPLSCSSLLQTYTVSPIADRRPSHCKGQPQLRPPISCALATPAAKITHNIRCLKDTSAVAYSLTTSGKATLPISNIKIHCIAFMYG-TPFQEPSQLPKREKKQAQKLFPLCSRPLPI-ASLAT-TQTMALPIFPKTSSICVPPLLFAILPMSPTIQPAPDL-NALMAFLKPYYISTLLTNPTYLWIMLYP-PYGQSTT-MY-PTATKPDGSFTTPPDSSRSQRHIPSAINKPIGIISSFLVLIAASGKDHRRLFKKLPALLSSR-ALVLPSGSRGGSSSELHAQDPSEAPPIPKKKTTNTMGKFLATLILFFQFCPLILGDYSPSCCTLTVGVSSYHSKPCNPAQPVCSWTLDLLALSADQALQPPCPNLVSYSSYHATYSLYLFPHWIKKPNRNGGGYYSASYSDPCSLKCPYLGCQSWTCPYTGAVSSPYWKFQQDVNFTQEVSHLNINLHFSKCGFSFSLLVDAPGYDPIWFLNTEPSQLPPTAPPLLSHSNLDHILEPSIPWKSKLLTLVQLTLQSTNYTCIVCIDRASLSTWHVLYSPNVSVPSPSSTPLLYPSLALPAPHLTLPFNWTHCFDPQIQAIVSSPCHNSLILPPFSLSPVPTLGSRSRRAVPVAVWLVSALAMGAGVAGRITGSMSLASGKSLLHEVDKDISQLTQAIVKNHKNLLKIAQYAAQNRRGLDLLFWEQGGLCKALQEQCCFLNITNSHVSILQERPPLENRVLTGWGLNWDLGLSQWAREALQTGITLVALLLLVILAGPCILRQLRHLPSRVRYPHYSLINPESSL-TKHTIIATTSPPASPANN-PLPSNPPSPAATSSVQPPRTPPRLPTV-YSHQPPTPAFFLS-HYAVSPSQPLVSTCAHGAPALPAFSGRSQRPSPPPASCAVPSPLPSFSNTQQSAFPPLSPALFFASSSPQPVAADHDAFPREVALSPLEGPVAAGRGFPLLEIANRQAQFPPPPCPLTLPPRIIARPPIPPPAGPPGMEQANIETALRIQS-PCLLSAHFPGFGQSLLFGYPVYVFGDCVQGDWCPISGGLCSARLHRHALLATCPEHQITWDPIDGRVIGSALQFLIPRLPSFPTQRTSKTLKVLTPPITHTTPNIPPSFLQAMRKYSPFRNGYMEPTLGQHLPTLSFPDPGLRPQNLYTLWGGSVVCMYLYQLSPPITWPLLPHVIFCHPGQLGAFLTNVPYKRIEKLLYKISLTTGALIILPEDCLPTTLFQPARAPVTLTAWQNGLLPFHSTLTTPGLIWTFTDGTPMISGPCPKDGQPSLVLQSSSFIFHKFQTKAYHPSFLLSHGLIQYSSFHNLHLLFEEYTNIPISLLFNEKEADDNDHEPQISPGGLEPLSEKHFRETEV-EGQGPE-GSDVSPRRTAQHQLRLGPDVSP-RQIISSDLREATGNHPFPPHVCQAVLRR-RQPLTSKNFSWHAYGSIK-QESIKAWGQFRRGLASLLHAPAALPEAAIHAG-VAFCRLPPVVPPELRPPSR-V-SSGRDRAFVRRSLGAYLDSAGSPRFA-PCLLNS

5'3' Frame 3

LASLLHAPAALPEAAIHAG-VAFCRLPPVVPPELRPPSR-V-SSGRDRAFVRRSLGAYLDSAGSPRFA-PCLLNSTSLFRFLFCAVTDRKFHPFPFHSRLTAGLAHGQVPATLLARSQRQPIL-HSPGEKFSTQLGARPGFERPFIP-AMGQIFSRSASPIPRPPRGLAAHHWLNFLQAAYRLEPGPSSYDFHQLKKFLKIALETPVWICPINYSLLASLLPKGYPGRVNEILHILIQTQAQIPSRPAPPPPSSSTHDPPDSDPQIPPPYVEPTAPQVLPVMHPHGAPPNHRPWQMKDLQAIKQEVSQAAPGSPQFMQTIRLAVQQFDPTAKDLQDLLQYLCSSLVASLHHQQLDSLISEAETRGITGYNPLAGPLRVQANNPQQQGLRREYQQLWLAAFAALPGSAKDPSWASILQGLEEPYHAFVERLNIALDNGLPEGTPKDPILRSLAYSNANKECQKLLQARGHTNSPLGDMLRACQAWTPKDKTKVLVVQPKKPPPNQPCFRCGKAGHWSRDCTQPRPPPGPCPLCQDPTHWKRDCPRLKPTIPEPEPEEDALLLDLPADIPHPKNSIGGEV-PPPPHYSKSFLTKTQHLFCQLYR-IPPVGP-LKPRLTPRPATQRLSKLY-IQEQT-QSFR-PCSQVILPSKTHPC-GQGAKPKITLSSPPFLC-YASLSGRRLLF-HLA-LIPKTTGPS-VVMPYNNAKASCTSLRQKGRL-SCQYRRQLSLG-NTSQGPPKSASSL-TRTPPGLATLGPEGPGGRPYRTLHRARK-PSIPS-KSQWNLAIHPRPAGH-LSNHRSLIIFPRAP-LVQPANYTSPLTNYRP-RRLFPNPPT-TVPALLCFHCPTAV-LRPRH-IRLESTTPRV-K-SHPVRNAAGPYPAAHSASLPPMHYSSVHG-HSPGKPLPCGPATTLRGHNGFPNLPWVACVRKQNPANPWNN-VPRANNFT-SPHL-CSPQGTYTVPLGAT-TSSPTWRDSVGLQRNSYLTPAPSQSLLCLTKAY-SPRPNIFKSFSSSIISAAAAGPVTELPQ-TSPNPAPPRGYYADPHWHHHCGVPVQAAVATCLATCPPTPH-PVPLGAATCLSCVITRQIHLAILWTTLPNHTS-HLHPNLQPIHSNI-PPQCSYLTPPQSPIQKFRCPDWRTLEHFS-NNCPIGSCESPYASVYSFPCDHKHRPLPVFRRIHLPGSLYSLGQAYIVTKIIPPSATAQVGPTGRTSRTFAWPLQRPFVALSQHISRLQVSLSLPSDPCPRHLPRQVLSGPLSGPPAPLTIA-GRLFAPRSQPYQST-SHLQAQRSHRCPTNHPCPAALSCRPTQFHPLRTDGPHTARGNHN-GLQYPALLPRLPQK-PTTSDASRTHPPWPTP-PHLARRHYPFQI-KYTVSPSCMGRHLFRSHLSYPKEKRNKLRSYFLFAPGHCLSRQA-LHKHRQWPCLYFPRLPQYVYLPCYSPYYPCPLQSNQLRTCRTL-WHS-NPII-VLY-QTRPTYG-CSIHSPMDNQPPKCINQLPQNPMAASPLPPTPADPRDTFPQQ-TNPLVLFQASWS--PPVERTTGGSSRSCRRCSHPGKR-FCPVDPVEAPQASCMPKTRRRPRRSQRKRPPTPWVSFSPL-FYSSSSAPSSSVITAPAAVLSQLESPHTTLNPAILPSQFVHGPSTCWPFQQIRPYSHPALI--VTPATMPPIPYIYSLIGSKSQTEMAEAIIQPLIQTLVP-NAHT-GANHGPAPIQEPSPAPTGNFSKMSILLKKFHTSILISIFQNAVFPSPF-STLQDMTPSGSLIPNPANCLPPPLLYSPTLT-TISSSPLYHGNQNS-LLSS-PYKALIILALSVSIVPAYPLGTSYTLPTSLFHPLLLPPSFTHR-RFQPPT-RYHLTGPTALTPRFKL-SPPPVITPSSCPPFPCHLFPR-DPAPAEQYRWRSGLSPPWPWEPEWLAGLPAPCPSPQERASYMRWTKIFPN-LKQ-SKTTKICSKLHSMLPRTDEALISCSGSKEDYAKHYKNSAVF-ILLIPMSQYYKRDPPLKIES-LAGALTGTLASHSGLEKPYKLESPLSRYSFLLSLQDHASSVSYDTSPRASDTPITLL-TLSHPCKPSTQLLQPHRLQPPLPIINLSHQILLLLQQPPPFSLQGLHLAFQLSSIAINPQLLHFFFPSTMLFRLLSPLSPLALTALLLFLLSPGEVSGLLLRPLPAPCLLLFLPFQILSNLLFLLFLPLFFSLPLLLSPSLPITMRFPARWRFPPWRAPSQPAAAFLF-R-QTVKHSFLLLLVL-LFLQG--PVHQFLHQQVLRAWNRQTSKQPYGYKVNHAYYQPTSQGLDRVFFSDTQSTCLETVYKATGAPSLGDYVRPAYIVTPYWPPVQSIRSPGTPSMDALSAQLYSSLSLDSPPSPPREPLRPLRSLPRQSLIQPPTFHPPSSRPCANTPPSEMDTWNPPLGSTSQPCLFQTPDSGPKTCTPSGEAPLSACTSTSFPPPSPGPSCPM-FFATPASSGPSSPMFPTNE-KNSSIKFPLPQGP--FYPRTVCPPPFSSLLGHPSR-QPGKTASFRSTQPSPLQALFGHLPMARL-FPGPALKMASHL-YYSPPPLYFTNFKPRPTTPHFYSHTASYSTLPFIICISYLKNTPTSPFLYFLTKKRQMTMTMSPKYPPGA-SLSVKNISVKQKSEKVRAQNKALTSPPGGQLSTSSG-ALTCPPKDKS-AQTSGKPPGTTHFLPMFVKPSSGVDDNPSPQKTFHGTHTAQ-NNRSL-KRGDSSGGGSHLSFTRPPPYLRPPSTPVESRSAASRLWCLLNYVRRLGKFRAQVETGPLSGAPLEPT-TQPALHALPDPACSTL

HiddenMarkovModel Logo

*Reverse transcriptase (RNA-dependent DNA polymerase)*

<http://pfam.xfam.org/family/PF00078.26#tabview=tab4> 

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*Human immunodeficiency virus 1 (NC\_001802.1)*

Genom

GGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCC

TCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGA

TCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACCTGAAAGCGAA

AGGGAAACCAGAGGAGCTCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGG

CGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCG

TCAGTATTAAGCGGGGGAGAATTAGATCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAAT

ATAAATTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGA

AACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTT

AGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGG

AAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAAGCACAGCAAGCAGCAGCTGACAC

AGGACACAGCAATCAGGTCAGCCAAAATTACCCTATAGTGCAGAACATCCAGGGGCAAATGGTACATCAG

GCCATATCACCTAGAACTTTAAATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTGA

TACCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCACAAGATTTAAACACCATGCTAAACACAGTGGG

GGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCTGCAGAATGGGATAGAGTG

CATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAA

CTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCTATCCCAGTAGGAGAAATTTA

TAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATA

AGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAACTCTAAGAGCCGAGCAAG

CTTCACAGGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCGAACCCAGATTGTAAGAC

TATTTTAAAAGCATTGGGACCAGCGGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGA

CCCGGCCATAAGGCAAGAGTTTTGGCTGAAGCAATGAGCCAAGTAACAAATTCAGCTACCATAATGATGC

AGAGAGGCAATTTTAGGAACCAAAGAAAGATTGTTAAGTGTTTCAATTGTGGCAAAGAAGGGCACACAGC

CAGAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGAT

TGTACTGAGAGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCCTACAAGGGAAGGCCAGGGAATTTTC

TTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTCTGGGGTAGAGACAACAACTCC

CCCTCAGAAGCAGGAGCCGATAGACAAGGAACTGTATCCTTTAACTTCCCTCAGGTCACTCTTTGGCAAC

GACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATAC

AGTATTAGAAGAAATGAGTTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATC

AAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGTGGACATAAAGCTATAGGTACAGTATTAGTAG

GACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCCAT

TAGCCCTATTGAGACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCA

TTGACAGAAGAAAAAATAAAAGCATTAGTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAA

AAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAG

AAAATTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCA

CATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTC

CCTTAGATGAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAACAATGAGACACCAGGGAT

TAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACA

AAAATCTTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATG

TAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTG

GGGACTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCAT

CCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGA

AGTTAGTGGGGAAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAACT

CCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAA

AACAGAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAA

TACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGG

AAAATATGCAAGAATGAGGGGTGCCCACACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAAAAATA

ACCACAGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAAACTGCCCATACAAAAGGAAACATGGG

AAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAGTGGGAGTTTGTTAATACCCCTCCCTT

AGTGAAATTATGGTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCA

GCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCC

TAACTGACACAACAAATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGA

AGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGATCAAAGTGAATCA

GAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAGGAAAAGGTCTATCTGGCATGGGTACCAGCAC

ACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGCTGGAATCAGGAAAGTACTATTTTT

AGATGGAATAGATAAGGCCCAAGATGAACATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGAT

TTTAACCTGCCACCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAG

CCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAAAAGT

TATCCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGAAGTTATTCCAGCAGAAACAGGGCAG

GAAACAGCATATTTTCTTTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAATACATACTGACAATGGCA

GCAATTTCACCGGTGCTACGGTTAGGGCCGCCTGTTGGTGGGCGGGAATCAAGCAGGAATTTGGAATTCC

CTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAAGAAAATTATAGGACAGGTA

AGAGATCAGGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAG

GGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGA

ATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAAATCCACTTTGG

AAAGGACCAGCAAAGCTCCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATAAAAG

TAGTGCCAAGAAGAAAAGCAAAGATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGC

AAGTAGACAGGATGAGGATTAGAACATGGAAAAGTTTAGTAAAACACCATATGTATGTTTCAGGGAAAGC

TAGGGGATGGTTTTATAGACATCACTATGAAAGCCCTCATCCAAGAATAAGTTCAGAAGTACACATCCCA

CTAGGGGATGCTAGATTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCATTTGG

GTCAGGGAGTCTCCATAGAATGGAGGAAAAAGAGATATAGCACACAAGTAGACCCTGAACTAGCAGACCA

ACTAATTCATCTGTATTACTTTGACTGTTTTTCAGACTCTGCTATAAGAAAGGCCTTATTAGGACACATA

GTTAGCCCTAGGTGTGAATATCAAGCAGGACATAACAAGGTAGGATCTCTACAATACTTGGCACTAGCAG

CATTAATAACACCAAAAAAGATAAAGCCACCTTTGCCTAGTGTTACGAAACTGACAGAGGATAGATGGAA

CAAGCCCCAGAAGACCAAGGGCCACAGAGGGAGCCACACAATGAATGGACACTAGAGCTTTTAGAGGAGC

TTAAGAATGAAGCTGTTAGACATTTTCCTAGGATTTGGCTCCATGGCTTAGGGCAACATATCTATGAAAC

TTATGGGGATACTTGGGCAGGAGTGGAAGCCATAATAAGAATTCTGCAACAACTGCTGTTTATCCATTTT

CAGAATTGGGTGTCGACATAGCAGAATAGGCGTTACTCGACAGAGGAGAGCAAGAAATGGAGCCAGTAGA

TCCTAGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAA

AAGTGTTGCTTTCATTGCCAAGTTTGTTTCATAACAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGC

GGAGACAGCGACGAAGAGCTCATCAGAACAGTCAGACTCATCAAGCTTCTCTATCAAAGCAGTAAGTAGT

ACATGTAATGCAACCTATACCAATAGTAGCAATAGTAGCATTAGTAGTAGCAATAATAATAGCAATAGTT

GTGTGGTCCATAGTAATCATAGAATATAGGAAAATATTAAGACAAAGAAAAATAGACAGGTTAATTGATA

GACTAATAGAAAGAGCAGAAGACAGTGGCAATGAGAGTGAAGGAGAAATATCAGCACTTGTGGAGATGGG

GGTGGAGATGGGGCACCATGCTCCTTGGGATGTTGATGATCTGTAGTGCTACAGAAAAATTGTGGGTCAC

AGTCTATTATGGGGTACCTGTGTGGAAGGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCA

TATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAG

TAGTATTGGTAAATGTGACAGAAAATTTTAACATGTGGAAAAATGACATGGTAGAACAGATGCATGAGGA

TATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCACTCTGTGTTAGTTTAAAG

TGCACTGATTTGAAGAATGATACTAATACCAATAGTAGTAGCGGGAGAATGATAATGGAGAAAGGAGAGA

TAAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTA

TAAACTTGATATAATACCAATAGATAATGATACTACCAGCTATAAGTTGACAAGTTGTAACACCTCAGTC

ATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTG

CGATTCTAAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACAAATGTCAGCACAGTACAATG

TACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTA

GTAATTAGATCTGTCAATTTCACGGACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAA

TTAATTGTACAAGACCCAACAACAATACAAGAAAAAGAATCCGTATCCAGAGAGGACCAGGGAGAGCATT

TGTTACAATAGGAAAAATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATAAC

ACTTTAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAAT

CCTCAGGAGGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTC

AACACAACTGTTTAATAGTACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACTGAAGGA

AGTGACACAATCACCCTCCCATGCAGAATAAAACAAATTATAAACATGTGGCAGAAAGTAGGAAAAGCAA

TGTATGCCCCTCCCATCAGTGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGA

TGGTGGTAATAGCAACAATGAGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGA

AGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAA

GAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGG

AAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAG

CAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCA

AGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGG

TTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAGTAATAAATCTCTGGAA

CAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACT

CCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGC

AAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGA

GGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCAC

CATTATCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGG

TGGAGAGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCCTTGGCACTTATCTGGGACGATCTG

CGGAGCCTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAAC

TTCTGGGACGCAGGGGGTGGGAAGCCCTCAAATATTGGTGGAATCTCCTACAGTATTGGAGTCAGGAACT

AAAGAATAGTGCTGTTAGCTTGCTCAATGCCACAGCCATAGCAGTAGCTGAGGGGACAGATAGGGTTATA

GAAGTAGTACAAGGAGCTTGTAGAGCTATTCGCCACATACCTAGAAGAATAAGACAGGGCTTGGAAAGGA

TTTTGCTATAAGATGGGTGGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCCTACTGTAAGGGAAAGAA

TGAGACGAGCTGAGCCAGCAGCAGATAGGGTGGGAGCAGCATCTCGAGACCTGGAAAAACATGGAGCAAT

CACAAGTAGCAATACAGCAGCTACCAATGCTGCTTGTGCCTGGCTAGAAGCACAAGAGGAGGAGGAGGTG

GGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACT

TTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTG

GATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCA

CTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGATAGAAGAGGCCAATAAAGGAG

AGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTG

GAGGTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGC

TGACATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTG

GGGAGTGGCGAGCCCTCAGATCCTGCATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAG

ACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCC

TTGAGTGCTTC

Peptidsequenz

5'3' Frame 1

GLSG-TRSEPGSSLAN-GTHCLSLNKACLECFK-CVPVCCVTLVTRDPSDPFSQCGKSLAVAPEQGPESERETRGALSTQDSAC-SAHGKRRGAATGEYAKNFD-RRLEGERWVRERQY-AGEN-IDGKKFG-GQGERKNIN-NI-YGQAGS-NDSQLILAC-KHQKAVDKYWDSYNHPFRQDQKNLDHYIIQ-QPSIVCIKG-R-KTPRKL-TR-RKSKTKVRKKHSKQQLTQDTAIRSAKITL-CRTSRGKWYIRPYHLEL-MHG-K--KRRLSAQK-YPCFQHYQKEPPHKI-TPC-TQWGDIKQPCKC-KRPSMRKLQNGIECIQCMQGLLHQAR-ENQGEVT-QELLVPFRNK-DG-QIIHLSQ-EKFIKDG-SWD-IK--ECIALPAFWT-DKDQRNPLETM-TGSIKL-EPSKLHRR-KIG-QKPCWSKMRTQIVRLF-KHWDQRLH-KK--QHVRE-EDPAIRQEFWLKQ-AK-QIQLP--CREAILGTKERLLSVSIVAKKGTQPEIAGPLGKRAVGNVERKDTK-KIVLRDRLIF-GRSGLPTREGQGIFFRADQSQQPHQKRASGLG-RQQLPLRSRSR-TRNCIL-LPSGHSLATTPRHNKDRGATKGSSIRYRSR-YSIRRNEFARKMETKNDRGNWRFYQSKTV-SDTHRNLWT-SYRYSISRTYTCQHNWKKSVDSDWLHFKFSH-PY-DCTSKIKARNGWPKS-TMAIDRRKNKSISRNLYRDGKGRENFKNWA-KSIQYSSICHKEKRQY-MEKISRFQRT--ENSRLLGSSIRNTTSRRVKKEKISNSTGCG-CIFFSSLR-RLQEVYCIYHT-YKQ-DTRD-ISVQCASTGMERITSNIPK-HDKNLRAF-KTKSRHSYLSIHG-FVCRI-LRNRAA-NKNRGAETTSVEVGTYHTRQKTSERTSIPLDGL-TPS--MDSTAYSAARKRQLDCQ-HTEVSGEIELGKSDLPRD-SKAIM-TP-RNQSTNRSNTTNRRSRARTGRKQRDSKRTSTWSVL-PIKRLNSRNTEAGARPMDISNLSRAI-KSENRKICKNEGCPH--CKTINRGSAKNNHRKHSNMGKDS-I-TAHTKGNMGNMVDRVLASHLDS-VGVC-YPSLSEIMVPVRERTHSRSRNLLCRWGS-QGD-IRKSRICY--RKTKSCHPN-HNKSED-VTSNLSSFAGFGIRSKHSNRLTICIRNHSSTTRSK-IRVSQSNNRAVNKKGKGLSGMGTSTQRNWRK-TSR-ISQCWNQESTIFRWNR-GPR-T-EISQ-LESNG--F-PATCSSKRNSSQL--MSAKRRSHAWTSRL-SRNMATRLYTFRRKSYPGSSSCSQWIYRSRSYSSRNRAGNSIFSFKISRKMASKNNTY-QWQQFHRCYG-GRLLVGGNQAGIWNSLQSPKSRSSRIYE-RIKENYRTGKRSG-TS-DSSTNGSIHPQF-KKRGDWGVQCRGKNSRHNSNRHTN-RITKTNYKNSKFSGLLQGQQKSTLERTSKAPLER-RGSSNTR---HKSSAKKKSKDH-GLWKTDGR--LCGK-TG-GLEHGKV--NTICMFQGKLGDGFIDITMKALIQE-VQKYTSH-GMLDW--QHIGVCIQEKETGIWVRESP-NGGKRDIAHK-TLN-QTN-FICITLTVFQTLL-ERPY-DT-LALGVNIKQDITR-DLYNTWH-QH--HQKR-SHLCLVLRN-QRIDGTSPRRPRATEGATQ-MDTRAFRGA-E-SC-TFS-DLAPWLRATYL-NLWGYLGRSGSHNKNSATTAVYPFSELGVDIAE-ALLDRGEQEMEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRQRRRAHQNSQTHQASLSKQ-VVHVMQPIPIVAIVALVVAIIIAIVVWSIVIIEYRKILRQRKIDRLIDRLIERAEDSGNESEGEISALVEMGVEMGHHAPWDVDDL-CYRKIVGHSLLWGTCVEGSNHHSILCIRC-SI-YRGT-CLGHTCLCTHRPQPTRSSIGKCDRKF-HVEK-HGRTDA-GYNQFMGSKPKAMCKINPTLC-FKVH-FEE-Y-YQ---RENDNGERRDKKLLFQYQHKHKR-GAERICIFL-T-YNTNR--YYQL-VDKL-HLSHYTGLSKGIL-ANSHTLLCPGWFCDSKM---DVQWNRTMYKCQHSTMYTWN-ASSINSTAVKWQSSRRRGSN-ICQFHGQC-NHNSTAEHICRN-LYKTQQQYKKKNPYPERTRESICYNRKNRKYETSTL-H--SKME-HFKTDS-QIKRTIWK--NNNL-AILRRGPRNCNAQF-LWRGIFLL-FNTTV--YLV--YLEY-RVK-H-RK-HNHPPMQNKTNYKHVAESRKSNVCPSHQWTN-MFIKYYRAAINKRWW--QQ-VRDLQTWRRRYEGQLEK-II-I-SSKN-TIRSSTHQGKEKSGAERKKSSGNRSFVPWVLGSSRKHYGRSLNDADGTGQTIIVWYSAAAEQFAEGY-GATASVATHSLGHQAAPGKNPGCGKIPKGSTAPGDLGLLWKTHLHHCCALEC-LE--ISGTDLESHDLDGVGQRN-QLHKLNTLLN-RIAKPARKE-TRIIGIR-MGKFVELV-HNKLAVVYKIIHNDSRRLGRFKNSFCCTFYSE-S-AGIFTIIVSDPPPNPEGTRQARRNRRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCNEDCGTSGTQGVGSPQILVESPTVLESGTKE-CC-LAQCHSHSSS-GDR-GYRSSTRSL-SYSPHT-KNKTGLGKDFAIRWVASGQKVV-LDGLL-GKE-DELSQQQIGWEQHLETWKNMEQSQVAIQQLPMLLVPG-KHKRRRRWVFQSHLRYL-DQ-LTRQL-ILATF-KKRGDWKG-FTPKEDKISLICGSTTHKATSLISRTTHQGQGSDIH-PLDGATS-YQLSQIR-KRPIKERTPACYTL-ACMGWMTRREKC-SGGLTAA-HFITWPESCIRSTSRTADIELATRDFPLGTFQGGVAWAGLGSGEPSDPAYKQLLFACTGSLWLDQI-AWELSG-LGNPLLKPQ-SLP-VL

5'3' Frame 2

VSLVRPDLSLGALWLTREPTA-ASIKLALSASSSVCPSVV-LW-LEIPQTLLVSVENL-QWRPNRDLKAKGKPEELSRRRTRLAEARTARGEGRRLVSTPKILTSGG-KERDGCESVSIKRGRIRSMGKNSVKARGKEKI-IKTYSMGKQGARTIRS-SWPVRNIRRL-TNTGTATTIPSDRIRRT-III-YSSNPLLCASKDRDKRHQGSFRQDRGRAKQK-EKSTASSS-HRTQQSGQPKLPYSAEHPGANGTSGHIT-NFKCMGKSSRREGFQPRSDTHVFSIIRRSHPTRFKHHAKHSGGTSSSHANVKRDHQ-GSCRMG-SASSACRAYCTRPDERTKGK-HSRNY-YPSGTNRMDDK-STYPSRRNL-KMDNPGIK-NSKNV-PYQHSGHKTRTKGTL-RLCRPVL-NSKSRASFTGGKKLDDRNLVGPKCEPRL-DYFKSIGTSGYTRRNDDSMSGSRRTRP-GKSFG-SNEPSNKFSYHNDAERQF-EPKKDC-VFQLWQRRAHSQKLQGP-EKGLLEMWKGRTPNERLY-ETG-FFREDLAFLQGKAREFSSEQTRANSPTRRELQVWGRDNNSPSEAGADRQGTVSFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMSLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTNRGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQDEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED-NMEKFSKTPYVCFRES-GMVL-TSL-KPSSKNKFRSTHPTRGC-IGNNNILGSAYRRKRLAFGSGSLHRMEEKEI-HTSRP-TSRPTNSSVLL-LFFRLCYKKGLIRTHS-P-V-ISSRT-QGRISTILGTSSINNTKKDKATFA-CYETDRG-MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIWLHGLGQHIYETYGDTWAGVEAIIRILQQLLFIHFQNWVST-QNRRYSTEESKKWSQ-ILD-SPGSIQEVSLKLLVPIAIVKSVAFIAKFVS-QKP-ASPMAGRSGDSDEELIRTVRLIKLLYQSSK-YM-CNLYQ--Q--H---Q---Q-LCGP--S-NIGKY-DKEK-TG-LID--KEQKTVAMRVKEKYQHLWRWGWRWGTMLLGMLMICSATEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLVNVTENFNMWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSSGRMIMEKGEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDNDTTSYKLTSCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSVNFTDNAKTIIVQLNTSVEINCTRPNNNTRKRIRIQRGPGRAFVTIGKIGNMRQAHCNISRAKWNNTLKQIASKLREQFGNNKTIIFKQSSGGDPEIVTHSFNCGGEFFYCNSTQLFNSTWFNSTWSTEGSNNTEGSDTITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNITGLLLTRDGGNSNNESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIGALFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEQIWNHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLPTPRGPDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRRGWEALKYWWNLLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGLERILL-DGWQVVKK-CDWMAYCKGKNETS-ASSR-GGSSISRPGKTWSNHK-QYSSYQCCLCLARSTRGGGGGFSSHTSGTFKTNDLQGSCRS-PLFKRKGGTGRANSLPKKTRYP-SVDLPHTRLLP-LAELHTRARGQISTDLWMVLQASTS-AR-DRRGQ-RREHQLVTPCEPAWDG-PGERSVRVEV-QPPSISSRGPRAASGVLQELLTSSLLQGTFRWGLSREAWPGRDWGVASPQILHISSCFLPVLGLSG-TRSEPGSSLAN-GTHCLSLNKACLECF

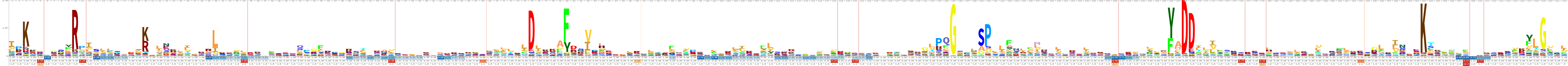
5'3' Frame 3

SLWLDQI-AWELSG-LGNPLLKPQ-SLP-VLQVVCARLLCDSGN-RSLRPF-SVWKISSSGARTGT-KRKGNQRSSLDAGLGLLKRARQEARGGDW-VRQKF-LAEARRREMGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEEQNKSKKKAQQAAADTGHSNQVSQNYPIVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKIVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSYKGRPGNFLQSRPEPTAPPEESFRSGVETTTPPQKQEPIDKELYPLTSLRSLFGNDPSSQ-R-GGN-RKLY-IQEQMIQY-KK-VCQEDGNQK--GELEVLSK-DSMIRYS-KSVDIKL-VQY--DLHLST-LEEIC-LRLVAL-IFPLALLRLYQ-N-SQEWMAQKLNNGH-QKKK-KH--KFVQRWKRKGKFQKLGLKIHTILQYLP-RKKTVLNGEN--ISENLIRELKTSGKFN-EYHIPQG-KRKNQ-QYWMWVMHIFQFP-MKTSGSILHLPYLV-TMRHQGLDISTMCFHRDGKDHQQYSKVA-QKS-SLLENKIQT-LSINTWMICM-DLT-K-GSIEQK-RS-DNIC-GGDLPHQTKNIRKNLHSFGWVMNSILINGQYSL-CCQKKTAGLSMTYRS-WGN-IGQVRFTQGLK-GNYVNSLEEPKH-QK-YH-QKKQS-NWQKTERF-KNQYMECIMTHQKT--QKYRSRGKANGHIKFIKSHLKI-KQENMQE-GVPTLMM-NN-QRQCKK-PQKA--YGERLLNLNCPYKRKHGKHGGQSIGKPPGFLSGSLLIPLP--NYGTS-RKNP--EQKPSM-MGQLTGRLN-EKQDMLLIEEDKKLSP-LTQQIRRLSYKQFI-LCRIRD-K-T--QTHNMH-ESFKHNQIKVNQS-SIK--SS--KRKRSIWHGYQHTKELEEMNK-IN-SVLESGKYYF-ME-IRPKMNMRNITVIGEQWLVILTCHL--QKK--PAVINVS-KEKPCMDK-TVVQEYGN-IVHI-KEKLSW-QFM-PVDI-KQKLFQQKQGRKQHIFF-N-QEDGQ-KQYILTMAAISPVLRLGPPVGGRESSRNLEFPTIPKVKE--NL-IKN-RKL-DR-EIRLNILRQQYKWQYSSTILKEKGGLGGTVQGKE--T--QQTYKLKNYKNKLQKFKIFGFITGTAEIHFGKDQQSSSGKVKGQ--YKIIVT-K-CQEEKQRSLGIMENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPLGDARLVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAGHNKVGSLQYLALAALITPKKIKPPLPSVTKLTEDRWNKPQKTKGHRGSHTMNGH-SF-RSLRMKLLDIFLGFGSMA-GNISMKLMGILGQEWKP--EFCNNCCLSIFRIGCRHSRIGVTRQRRARNGASRS-TRALEASRKSA-NCLYQLLL-KVLLSLPSLFHNKSLRHLLWQEEAETATKSSSEQSDSSSFSIKAVSSTCNATYTNSSNSSISSSNNNSNSCVVHSNHRI-ENIKTKKNRQVN--TNRKSRRQWQ-E-RRNISTCGDGGGDGAPCSLGC--SVVLQKNCGSQSIMGYLCGRKQPPLYFVHQMLKHMIQRYIMFGPHMPVYPQTPTHKK-YW-M-QKILTCGKMTW-NRCMRI-SVYGIKA-SHV-N-PHSVLV-SALI-RMILIPIVVAGE--WRKER-KTALSISAQA-EVRCRKNMHFFINLI-YQ-IMILPAIS-QVVTPQSLHRPVQRYPLSQFPYIIVPRLVLRF-NVIIRRSMEQDHVQMSAQYNVHMELGQ-YQLNCC-MAV-QKKR--LDLSISRTMLKP--YS-THL-KLIVQDPTTIQEKESVSREDQGEHLLQ-EK-EI-DKHIVTLVEQNGITL-NR-LAN-ENNLEIIKQ-SLSNPQEGTQKL-RTVLIVEGNFSTVIQHNCLIVLGLIVLGVLKGQITLKEVTQSPSHAE-NKL-TCGRK-EKQCMPLPSVDKLDVHQILQGCY-QEMVVIATMSPRSSDLEEEI-GTIGEVNYINIK--KLNH-E-HPPRQREEWCREKKEQWE-ELCSLGSWEQQEALWAQPQ-R-RYRPDNYCLV-CSSRTIC-GLLRRNSICCNSQSGASSSSRQESWLWKDT-RINSSWGFGVALENSFAPLLCLGMLVGVINLWNRFGITRPGWSGTEKLTITQA-YTP-LKNRKTSKKRMNKNYWN-INGQVCGIGLT-QIGCGI-NYS----EAW-V-E-FLLYFL--IELGRDIHHYRFRPTSQPRGDPTGPKE-KKKVERETETDPFD--TDPWHLSGTICGACASSATTA-ETYS-L-RGLWNFWDAGGGKPSNIGGISYSIGVRN-RIVLLACSMPQP-Q-LRGQIGL-K-YKELVELFATYLEE-DRAWKGFCYKMGGKWSKSSVIGWPTVRERMRRAEPAADRVGAASRDLEKHGAITSSNTAATNAACAWLEAQEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLWIYHTQGYFPD-QNYTPGPGVRYPLTFGWCYKLVPVEPDKIEEANKGENTSLLHPVSLHGMDDPEREVLEWRFDSRLAFHHVARELHPEYFKNC-HRACYKGLSAGDFPGRRGLGGTGEWRALRSCI-AAAFCLYWVSLVRPDLSLGALWLTREPTA-ASIKLALSA

HiddenMarkovModel Logo

*Reverse transcriptase (RNA-dependent DNA polymerase)*

<http://pfam.xfam.org/family/PF00078.26#tabview=tab4>



*Retroviral aspartyl protease*

<http://pfam.xfam.org/family/PF00077.19#tabview=tab4>

