Supplementary Dataset 2

Proviral sequences generated as part of the study using the medaka_consensus program (https://nanoporetech.github.io/medaka/index.html) with manual correction guided by mapping reads to consensus and inspection in IGV

pBLV344
Bos-1053_BLV_consensus
Bos-1439_BLV_consensus
HIV_U1_CHRX
HIV_U1_CHR2
02006_consensus, consensus sequence of the proviruses from patient 02006
02006_chr10_sc, sequence of the provirus with ID chr10:41845251-42110306*
02006_cen, sequence of the provirus with ID chr13:16283554-17732159*
BosT_ERV_chr11_APOB

*these two consensus sequences were generated using a limited number of Nanopore reads and as a consequence there is a higher change that they may contain errors.

>pBLV344 8721 bp

TGTATGAAAGATCATGCCGGCCTAGGCGCCGCCACCGCCCCGTAAACCAGACAGAGACGTCAGCTGCCAGAAAAGCTGGTGACGGCAGCT

GGTGGCTAGAATCCCCGTACCTCCCCAACTTCCCCTTTCCCGAAAAATCCACACCCCGAGCTGCTGACCTCACCTGCTGATAAAACAATA

AAATGCCGGCCCTGTCGAGTTAGCGGCACCAGAAGCGTTCTCCTCCTGAGACCCTCGTGCTCAGCTCTCGGTCCTGAGCTCTCTTGCTCC

CGAGACCTTCTGGTCGGCTATCCGGCAGCGGTCAGGTAAGGCAAACCACGGTTTGGAGGGTGGTTCTCGGCTGAGACCACCGCGAGCTCT

ATCTCCGGTCCTCTGACCGTCTCCACGTGGACTCTCTCTCTTGCCTCCTGACCCCGCGCTCCAAGGGCGTCTGGCTTGCACCCGCGCTTG

TTTCCTGTCTTACTTTCTGTTTCTCGCGGCCCGCGCTCTCTCCTTCGGCGCCCTCTAGCGGCCAGGAGAGACCGGCAAACAATTGGGGGC

TCGTCCGGGATTGATCACCCCGGAACCCTAACAATCCTCTGGACCCACCCCCTCGGCGGCGTTTTGGGTCTTTCCTTTAAATTATATCAT

GGGAAATTCCCCCTCCTATAACCCCCCCGCTGGTATCTCCCCCTCAGACTGGCTCAACCTTTTGCAAAGCGCGCAAAGGCTCAATCCGCG

ACCCTCTCCTAGCGATTTTACCGATTTAAAAAATTACATCCATTGGTTTCATAAGACCCAGAAAAAACCATGGACTTTCACTTCTGGTGG

CCCCGCCTCATGCCCACCCGGGAAATTCGGCCGGGTTCCCCTTGTCTTGGCCACCCTAAACGAAGTGCTCTCAAACGATGAGGGCGCCCC

GGGTGCATCGGCCCCAGAAGAACAACCCCCCCCTTATGACCCCCCCGCCGTTTTGCCAATCATATCTGAAGGGAATCGCAACCGCCATCG

CGCTTGGGCACTCCGAGAATTACAAGATATTAAAAAAGAAATTGAAAATAAGGCACCGGGTTCGCAAGTATGGATACAAACACTACGACT

TGCAATCTTGCAGGCCGACCCTACTCCTGCTGACCTAGAACAACTTTGCCAATATATTGCTTCCCCGGTCGATCAAACGGCCCACATGAC

CAGCCTAACGGCAGCAATAGCAGCCGCTGAAGCGGCCAACACCCTTCAGGGTTTTAATCCCCAAA
ACGGGACCCTGACCCAACAATCAGC

TCAGCCCAACGCCGGGGATCTTAGAAGTCAATATCAAAACCTTTGGCTTCAGGCCTGGAAAAATCT
CCCTACTCGTCCTTCAGTACAACC

CTGGTCCACCATCGTCCAAGGCCCCGCCGAGAGCTATGTAGAGTTTGTCAACCGGTTACAAATTTC
ATTAGCTGACAACCTTCCCGACGG

AGTCCCTAAAGAACCCATTATTGACTCCCTTAGCTATGCTAATGCTAACAAAGAATGCCAACAAATTT
TGCAGGGGCGGGGCCTAGTGGC

CGCCCCGGTGGGACAAAAACTGCAGGCTTGTGCACATTGGGCCCCCAAGATTAAACAGCCTGCAA
TCCTCGTCCACACCCCAGGGCCCAA

GATGCCCGGGCCTCGGCAACCGGCCCCCAAAAGGCCCCCCCGGGACCATGCTATCGATGCCTC
AAAGAAGGCCATTGGGCCCGGGACTG

TCCCACCAAGACCACCGGCCCCCCTCCGGGACCTTGTCCCATATGCAAAGATCCTTCCCATTGGAA
ACGAGACTGTCCAACCCTCAAATC

AAAAAACTAATAGAGGGGGGACTTAGCGCCCCCCAAACCGTAACCCCTATAACAGATCCTCTTAGT
GAGGCTGAATTGGAATGCTTACTT

TCTATTCCTCTGGCTCGCAGCCGTCCCTCCGTGGCTGTATACCTGTCTGGCCCCTGGCTGCAGCC
CTCTCAGAATCAAGCCCTTATGCTC

GTGGACACCGGGGCTGAAAATACGGTCCTCCCACAAAATTGGCTGGTTCGAGATTACCCACGGAT
CCCCGCCGCAGTGCTCGGAGCGGGG

GGAGTCTCCCGGAACAGATACAATTGGCTACAAGGCCCTCTGACCCTGGCTCTAAAACCAGAGGG
TCCCTTTATCACCATCCCAAAAATT

TTAGTTGACACTTTCGATAAATGGCAAATTTTAGGACGGGACGTCCTCTCCCGCCTACAGGCCTCTA
TCTCCATACCTGAGGAGGTACGC

CCCCCCATGGTAGGCGTCCTAGATGCCCCCCCGAGCCACATTGGATTAGAACATTTGCCCGCCCC
ACCTGAGGTACCTCAATTCCCTTTA

AACTAGAACGCCTCCAGGCCCTTCAAGACCTGGTCCATCGCTCTCTGGAGGCAGGTTATATCTCCC
CCTGGGACGGGCCAGGCAATAATC

CAGTATTCCCGGTACGGAAACCAAATGGCACCTGGAGGTTTGTGCATGATCTACGAGCTACAAATG
CTCTTACAAAGCCCATCCCGGCGC

TCTCCCCCGGACCGCCAGACCTTACCGCTATCCCTACACACCTTCCACATATCATTTGCCTAGATCT
CAAAGATGCCTTCTTCCAGATTC

CAGTCGAAGACCGCTTCCGCTCCTATTTTGCTTTTACCCTCCCTACCCCCGGGGGACTCCAACCTC
ATAGACGCTTTGCCTGGCGGGTCC

TACCTCAAGGCTTCATTAATAGCCCGGCTCTTTTCGAACGGGCACTACAGGAACCCCTTCGCCAAG
TTTCCGCCGCCTTCTCCCAGTCTC

TTCTGGTGTCCTATATGGACGATATCCTTATCGCTTCGCCTACAGAAGAACAACGGTCACAATGTTA
TCAAGCCCTGGCTGCCCGCCTCC

GGGACCTAGGGTTTCAGGTGGCGTCTGAAAAGACTCGCCAGACGCCTTCGCCCGTCCCCTTCCTG
GGACAAATGGTCCATGACCAGATTG

TCACCTATCAGTCCCTACCTACCTTGCAGATCTCATCCCCAATTTCTCTTCACCAATTACAGGCGGT
CTTGGGAGACCTCCAGTGGGTCT

CCAGGGGCACACCTACTACCCGCCGACCCCTGCAACTTCTCTACTCTTCCCTTAAAGGCATCGATG
ACCCTAGGGCCACCATCCAGCTTT

CCCCGGAACAGCTACAAGGCATTGCAGAGCTTCGACAAGCCCTGTCCCATAACGCAAGATCTAGAT
ATAACGAGCAAGAACCCCTGCTGG

CCTACATACACCTAACCCGGGCGGGGTCCACCCTGGTACTCTTCCAAAAGGGCGCTCAATTTCCCC
TGGCCTACTTTCAGACCCCCTTGA

CTGACAACCAAGCCTCACCTTGGGGCCTCCTTCTCCTGCTGGGATGCCAATACCTGCAGACTCAG
GCCTTAAGCTCTTATGCCAAGCCCA

TACTCAAATACTATCACAATCTTCCTAAAACCTCTCTCGACAATTGGATTCAATCATCTGAGGACCCT
CGAGTTCAGGAGTTGTTGCGAT

TGTGGCCCCAGATTTCCTCTCAGGGAATACAGCCCCCGGGCCCCTGGAAGACCTTGATCACCAGG
GCAGAGGTTTTTTTGACGCCCCAGT

TCTCTCCTGAACCGATTCCTGCGGCCCTTTGCCTCTTTAGTGACGGGGCTACAGGACGAGGAGCAT
ATTGCCTGTGGAAAGACCACCTTT

TGGACTTTCAGGCCGTTCCGGCTCCAGAGTCCGCCCAAAAGGGAGAACTAGCAGGTCTCTTGGCG
GGCTTAGCAGCCGCCCCGCCTGAAC

CTTTAAATATATGGGTAGATTCCAAATACCTATACTCCTTGCTCAGAACCCTAGTTCTGGGAGCTTG
GCTTCAACCTGACCCCGTACCCT

CCTATGCCCTCCTATACAAAAGCCTCCTCCGACATCCAGCAATCTTTGTTGGTCATGTCCGGAGCC
ACTCCTCAGCATCCCACCCTATTG

CTTCCCTGAACAATTATGTAGATCAACTGCTCCCCTTAGAAACTCCAGAGCAATGGCATAAGCTCAC
CCACTGCAACTCTCGGGCCTTGT

CTCGATGGCCGAACCCACGTATTTCGGCCTGGGATCCCCGTTCCCCCGCTACGCTATGTGAAACC
TGTCAAAAGCTCAATCCAACTGGAG

GTGGAAAGATGCGAACTATTCAGAGAGGGTGGGCCCCGAATCATATTTGGCAGGCCGATATAACC
CATTATAAATACAAACAGTTCACCT

ACGCTTTGCACGTGTTTGTAGATACTTACTCTGGAGCTACTCATGCCTCAGCAAAGCGAGGGCTCA
CCACTCAAATGACCATTGAGGGCC

TTCTGGAGGCCATAGTACATCTGGGTCGTCCAAAAAAGCTAAACACTGACCAAGGCGCAAACTACA
CCTCCAAAACCTTTGTCAGGTTTT

GCCAGCAGTTCGGAGTTTCCCTTTCTCATCACGTTCCCTACAACCCCACAAGTTCAGGGTTGGTAG
AACGGACAAATGGACTGCTCAAAC

TTCTTTTGTCTAAATATCACCTAGACGAACCCCACCTTCCCATGACTCAGGCCCTTTCTCGAGCCCT
CTGGACTCACAATCAGATTAACC

TCCTGCCAATTCTAAAGACCAGATGGGAGTTACACCATTCACCCCTACTTGCTGTCATTTCAGAGGG
CGGAGAAACACCCAAGGGCTCTG

ATAAACTCTTTTTGTACAAGCTCCCCGGGCAAAACAATCGTCGGTGGCTAGGACCACTCCCGGCCC
TAGTCGAAGCCTCGGGAGGCGCCC

TCCTGGCTACTAACCCCCCCGTGTGGGTTCCCTGGCGTTTGCTAAAAGCCTTCAAATGCCCAAAGAACGACGGTCCCGAAGACGCCCACA

ACCGATCATCAGATGGGTAAGTCTCACTCTTACTCTCCTCGCTCTCTGTCAGCCCATCCAGACTTGGAGATGCTCCCTGTCCCTAGGAAA

TCAACAATGGATGACAACATATAACCAAGAGGCAAAATTTTCCATCGCCATTGACCAAATACTAGAGGCTCATAATCAATCGCCTTTCTG

TCCCAGGTCTCCCAGATACACCTTGGACTTTGTAAATGGTTATCCTAAGATCTATTGGCCCCCCCCACAAGGGCGACGCCGGTTTGGAGC

CAGGGCCATGGTCACATATGATTGCGAGCCCCGATGCCCTTATGTGGGGGCAGATCACTTCGACTGCCCCCACTGGGACAATGCTTCCCA

GGCCGATCAAGGGTCCTTTTATGTCAATCATCAGATTTTATTCCTGCATCTCAAACAATGTCATGGAATTTTCACTCTAACCTGGGAAAT

ATGGGGATATGATCCCCTGATCACCTTTTCTTTACATAAAATCCCTGATCCCCCTCAACCCGACTTCCCTCAGCTGAACAGTGACTGGGT

TCCCTCTGTCAGGTCATGGGCCCTGCTTTTAAATCAAACGGCACGGGCCTTCCCAGACTGTGCTATATGTTGGGAACCTTCCCCTCCCTG

GGCTCCCGAAATATTAGTATATAACAAAACCATCTCCAACTCTGGACCCGGTCTCGCCCTCCCGGACGCCCAAATCTTCTGGGTCAACAC

GTCCTTGTTTAACACCACCCAAGGATGGCACCACCCTTCCCAGAGGTTGTTGTTCAACGTTTCTCAAGGCAACGCCTTATTATTACCCCC

TATCTCCCTGGTTAATCTCTCTACGGCTTCCTCCGCCCCTCCTACCCGGGTCAGACGCAGTCCTGCCGCAGCCCTGACCTTGGGCCTAGC

CCTGTCAGTGGGGCTCACTGGAATTAATGTAGCCGTGTCCGCCCTTAGCCATCAGAGACTCACCTCCCTGATCCACGTTCTGGAGCAAGA

TCAGCAACGCTTGATCACAGCAATTAACCAGACCCACTATAATTTGCTTAATGTGGCCTCTGTGGTCGCCCAGAACCGACGGGGGCTAGA

TTGGTTGTACATCCGGCTGGGTTTTCAAAGCCTATGTCCCACGATCAATGAACCTTGCTGTTTCCTGCGCATTCAAAATGACTCCATTAT

CCGCCTCGGTGATCTCCAGCCTCTCTCGCAAAGAGTCTCTACAGACTGGCAATGGCCCTGGAATTGGGATCTGGGGCTCACCGCCTGGGT

GCGAGAAACCATTCATTCTGTTCTAAGCCTATTCCTATTAGCCCTTTTTTTGCTCTTCTTGGCCCCCTGCCTGATAAAATGCTTGACCTC

TCGCCTTTTAAAACTCCTCCGGCAGGCTCCCCACTTCCCTGAAATCTCCTTCCCCCCTAAACCCGATTCTGATTATCAGGCCTTGCTACC

ATCCGCGCCAGAGATCTACTCTCACCTCTCCCCCACCAAACCCGATTACATCAACCTTCGACCCTGCCCTTGACACCCCCATGTTTCACG

CACCCTCAGGCTGTGGTGGGGCACTGGCTTAGTGGAATAGTCAGTGTACCATCACAAGCCTCTTCTTGCTGCCAGCGCCGAGTTCGAACA

CAGCCCTACCCTGAGCCTCTCTGAGTGCATGACTGAGTGTAGCGCAGAGAGATTGTCGCTTCTGCGTGTCACTCAGTCATTTTTTATAGC

CGATTGGGGTTCGCGCCCTCCCGTTGCCTGTGACACGGTTAAGACCTCTCTCACTTCTGCTTCACC
ATCCCCCTGCCAGCGTTGGTCTAG

TGGAAAGAACTAACGCTGACGGGGGCGATTTCTTGCAGCTGTGCTAAGCGAGAGGCTCTGGTGCT
GGGGATAAGATGCGGCCCCTAGCAC

CACAGTCTCTGCGCCTTTTGGGTTCGAATCTTCCCCATGCAGCTTCCGCTTTTTACGCCCTGTTGCA
CACCCTTTCTAGAGATACCTGAA

AATCTCAGCTCGCACCCCAAGGAAGGTTGTGGCTCAGAGGTTAAAATAGCTCGGACCGCAACCTC
CCTTTCTTTTTATTCCACCCTCGCA

AGGCCCCGGGTTCTAGACCCCCTAACGGAGGTTCAAAATTTCCTCTACTAGGGGGTGCTCAGGTC
CAAGTGTGCACAACATCTCTTCCAA

AAGGTCCTGATGAACATCTTCCCATGTAACAAGCCCCAGCAGAGACATTCCAGCCACATCCAGCAG
CATTTGGGCCGCCTTCTCTAACAG

TGCCCATAAAGTCCCTTCTGTTTCCACAACGGCTGCCTCTGCATCTTCTATTTCCACCTCGGCACCG
ACTCCCCCGCCGAGCCCTTCAAG

CTCTTCGGGATCCATTACCTGATAACGACAAAATTATTTCTTGTCTTTTAAGCAAGTGTTGTTGGTTG
GGGGCCCCACTCTCTACATGCC

TGCCCGGCCCTGGTTTTGTCCAATGATGTCACCATCGATGCCTGGTGCCCCCTCTGCGGGCCCCA
TGAACGACTCCAATTCGAAAGGATC

GACACCACGCTCACCTGCGAGACCCACCGTATCACCTGGACCGCCGATGGACGACCTTTCGGCCT
CAATGGAACGTTGTTCCCTCGACTG

CATGTCTCCGAGACCCGCCCCCAAGGGCCCCGACGACTCTGGATCAACTGCCCCCTTCCGGCCGT
TCGCGCTCAGCCCGGCCCGGTTTCA

CTTTCCCCCTTCGAGCAGTCCCCCTTCCAGCCCTACCAATGCCAATTGCCCTCGGCCTCTAGCGAT
GGTTGCCCCATCATCGGGCACGGC

CTTCTTCCCTGGAACAGCTTAGTAACGCATCCTGTCCTCGGAAAAGTCCTTACATTAAATCAAATGG
CCAATTTTTCCTTACTCCCCCCC

TTCGATACCCTCCTTGTGGACCCCCTCCGGCTGTCCGTCTTTGCCCCGGACACTAGGGGAGCCAT
ACGTTATCTCTCCACCCTTTTGACG

CTATGCCCAGCTACTTGTATTCTACCCCTAGGCGAGCCCTTCTCTCCTAATGTCCCCATATGCCGCT
TTCCCCGGGACACCAATGAACCT

CCCCTTTCAGAATTCGAGCTGCCCCTTATCCAAACGCCCGGCCTGTCTTGGTCTGTCCCCGCGATC
GACCTATTCCTAACCGGTCCCCCT

TCCCCATGCGACCGGTTACACGTGTGGTCCAGTCCTCAGGCCTTACAACGCTTCCTCCATGACCCC
ACGCTCACCTGGTCAGAATTGGTT

GCTAGCGGGAAACTAAGACTTGATTCACCCTTAAAATTACAGCTGTTAGAAAATGAATGGCTCTCCC
GCCTTTTTTGAGGGGGAGTCATT

TGTATGAAAGATCATGCCGGCCTAGGCGCCGCCACCGCCCCGTAAACCAGACAGAGACGTCAGCT
GCCAGAAAAGCTGGTGACGGCAGCT

GGTGGCTAGAATCCCCGTACCTCCCCAACTTCCCCTTTCCCGAAAAATCCACACCCCGAGCTGCTG
ACCTCACCTGCTGATAAAACAATA

AAATGCCGGCCCTGTCGAGTTAGCGGCACCAGAAGCGTTCTCCTCCTGAGACCCTCGTGCTCAGC
TCTCGGTCCTGAGCTCTCTTGCTCC

CGAGACCTTCTGGTCGGCTATCCGGCAGCGGTCAGGTAAGGCAAACCACGGTTTGGAGGGTGGTT
CTCGGCTGAGACCACCGCGAGCTCT

ATCTCCGGTCCTCTGACCGTCTCCACGTGGACTCTCTCTCTTGCCTCCTGACCCCGCGCTCCAAGG
GCGTCTGGCTTGCACCCGCGCTTG

TTTCCTGTCTTACTTTCTGTTTCTCGCGGCCCGCGCTCTCTCCTTCGGCGCCCTCTAGCGGCCAGG
AGAGACCGGCAAACA

>Bos-1053_BLV_consensus 8720 bp

TGTATGAAAGATCATGCCGACCTAGGCGCCGCCACCGCCCCGTAAACCAGACAGAGACGTCAGCT
GCCAGAAAAGCTGGTGACGGCAGCT

GGTGGCTAGAATCCCCGTACCTCCCCAACTTCCCCTTTCCCGAAAAATCCACACCCTGAGCTGCTG
ACCTCACCTGCTGATAAACTAATA

AAATGCCGGCCCTGTCGAGTTAGCGGCACCAGAAGCGTTCTTCTCCTGAGACCCTCGTGCTCAGC
TCTCGGTCCTGAGCTCTCTTGCTCC

CGAGACCTTCTGGTCGGCTATCCGGCAGCGGTCAGGTAAGGCAAACCACGGTTTGGAGGGTGGTT
CTCGGCTGAGACCACCGCGAGCTCT

ATCTCCGGTCCTCTGACCGTCTCCACGTGGACTCTCTCCTTTGCCTCCTGACCCCGCGCTCCAAGG
GCGTCTGGCTTGCACCCGCGTTTG

TTTCCTGTCTTACTTTCTGTTTCTCGCGGCCCGCGCTCTCTCCTTCGGCGCCCTCTAGCGGCCAGG
AGAGACCGGCAAACAATTGGGGGC

TCGTCCGGGATTGATCACCCCGGAACCCTAACAACTCTCTGGACCCACCCCCTCGGCGGCATTTT
GGGTCTCTCCTTCAAATTATATCAT

GGGAAATTCCCCCTCCTATAACCCCCCCGCTGGTATCTCCCCCTCAGACTGGCTCAACCTTCTGCA
AAGCGCGCAAAGGCTCAATCCGCG

ACCCTCTCCTAGCGATTTTACCGATTTAAAGAATTACATCCATTGGTTTCATAAGACCCAGAAAAAAC
CATGGACTTTCACTTCTGGTGG

CCCCACCTCATGTCCACCCGGGAGATTCGGCCGGGTTCCCCTTGTCTTGGCCACCCTAAACGAAG
TGCTCTCAAACGAAGGGGGCGCCCC

GGGTGCATCGGCCCCAGAAGAACAACCCCCCCCTTATGACCCCCCCGCCGTTTTGCCAATCATAT
CTGAAGGGAATCGCAACCGCCATCG

TGCTTGGGCACTCCGAGAATTACAAGATATCAAAAAGGAAATTGAAAATAAGGCACCGGGTTCGCA
AGTATGGATACAAACACTACGACT

TGCAATCCTGCAGGCCGACCCTACTCCGGCTGACTTAGAACAACTTTGCCAATATATTGCTTCCCC
GGTCGACCAAACGGCCCATATGAC

CAGCCTAACGGCAGCAATTGCCGCCGCTGAAGCGGCCAACACCCTCCAGGGTTTTAACCCCCAAA
ACGGGACCCTAACCCAACAATCAGC

TCAGCCCAACGCCGGGGATCTTAGAAGTCAATATCAAAACCTCTGGCTTCAGGCCTGGAAAAATCTCCCTACTCGTCCTTCAGTACAACC

TTGGTCCACCATCGTCCAAGGCCCCGCCGAAAGCTATGTAGAGTTTGTCAACCGGTTACAAATTTCATTAGCTGACAACCTTCCCGACGG

AGTCCCTAAGGAACCCATTATTGACTCCCTTAGTTATGCAAATGCTAACAAAGAGTGTCAGCAAATTTTGCAAGGGCGAGGCCTAGTGGC

CGCCCCGGTGGGGCAAAAGCTGCAGGCTTGCGCACATTGGGCCCCCAAGATGAAACAGCCTGCAATTCTCGTCCACACCCCAGGGCCCAA

GATGCCCGGGCCTCGGCAACCGGCCCCCAAAAGGCCTCCCCCAGGACCATGCTATCGATGCCTCAAAGAAGGCCATTGGGCCCGGGATTG

TCCTACCAAGGCCACCGGCCCCCCTCCGGGACCTTGCCCCATATGTAAAGATCCTTCCCATTGGAAACGAGACTGTCCAACCCTCAAATC

AAAAAACTAATAGAGGGGGGACTTAGCGCCCCCCAAACCATAACACCCATAACAGATTCTCTTAGTGAGGCCGAATTAGAATGCTTACTT

TCTATTCCTCTGGCTCGCAGCCGTCCCTCCGTGGCTGTATACCTGTCTGGCCCCTGGCTGCAGCCCTCTCAGAATCAAGCCCTCATGCTT

GTGGACACCGGGGCTGAAAATACGGTTCTCCCACAAAATTGGCTGGTTCGGGATTACCCACGGATCCCCGCCGCAGTGCTCGGAGCAGGG

GGAGTCTCCCGGAACAGATACAATTGGCTACAAGGCCCTCTGACCCTGGCTCTAAAACCAGAGGGTCCCTTTATCACCATCCCAAAAATT

TTAGTTGACACTTTCGACAAATGGCAAATTTTAGGACGGGACGTCCTCTCCCGCCTACAGGCCTCTATCTCCATACCTGAGGAAGTACGC

CCCCCTGTGGTAGGCGTCTTGGATGCCCCCCCGAGCCACATTGGATTAGAACATCTGCCCCCTCCACCTGAGGTGCCTCAATTCCCTTTA

AACTAGAACGCCTCCAGGCCCTTCAAGACCTGGTCCATCGCTCTCTGGAGGCAGGTTATATCTCCCCCTGGGACGGGCCAGGCAATAATC

CAGTCTTCCCGGTACGGAAACCAAATGGCGCCTGGAGGTTTGTGCATGACCTACGAGCTACAAATGCTCTTACAAAGCCCATTCCGGCAC

TCTCCCCCGGACCGCCAGACCTTACCGCTATCCCTACACACCTTCCACATATCATTTGCCTAGATCTCAAAGACGCCTTCTTCCAGATTC

CAGTCGAAGACCGCTTCCGCTCCTACTTTGCTTTTACCCTCCCTACCCCCGGGGGACTCCAACCTCATAGACGCTTTGCCTGGCGGGTCC

TACCTCAAGGCTTCATTAACAGCCCAGCTCTTTTCGAACGAGCACTACAGGAACCTCTTCGCCAAGTTTCCGCCGCCTTTTCCCAGTCTC

TTCTGGTGTCCTATATGGACGATATCCTTATCGCTTCGCCTACAGAAGAACAGCGGTCACAATGTTATCAAGCCCTGGCTGCCCGCCTCC

GGGACCTAGGGTTTCAGGTGGCATCCGAAAAGACTCGCCAGACGCCTTCGCCCGTCCCCTTCTTGGGACAAATGGTCCATGAGCAGATTG

TCACCTACCAGTCCCTACCTACCTTGCAGATCTCATCCCCAATTTCTCTTCACCAATTACAGGCGGTCTTAGGAGACCTCCAATGGGTCT

CTAGGGGCACACCCACTACCCGCCGGCCCCTGCAACTTCTCTACTCTTCCCTTAAAGGCATCGATG
ACCCTAGGGCCATCATCCAGCTTT

CCCCGGAACAGCTGCAAGGCATTGCAGAGCTTCGACAAGCCCTGTCCCATAACGCAAGATCTAGA
TATAACGAGCAAGAACCCCTGCTAG

CCTACGTACACCTAACCCGGGCGGGGTCCACCCTGGTACTCTTCCAAAAGGGCGCTCAATTTCCC
CTGGCCTACTTCCAGACCCCCTTGA

CTGACAACCAAGCCTCACCTTGGGGCCTCCTTCTCCTGCTGGGATGCCAATACCTGCAGACTCAG
GCCTTAAGCTCGTATGCCAAGCCCA

TACTTAAATATTATCACAATCTTCCTAAAACTTCTCTAGACAATTGGATTCAATCATCTGAGGACCCT
CGAGTCCAGGAGTTGCTGCAAT

TGTGGCCCCAGATTTCCTCTCAGGGAATACAGCCCCCGGGCCCTTGGAAGACCTTAATCACCAGG
GCAGAGGTTTTTTTGACGCCCCAGT

TCTCCCCTGATCCGATTCCTGCGGCCCTTTGCCTCTTTAGTGACGGGGCTAAAGGACGAGGAGCAT
ATTGCTTGTGGAAAGACCACCTTT

TAGACTTTCAGGCCGTTCCGGCTCCAGAATCCGCTCAAAAGGGAGAACTAGCAGGTCTCTTGGCG
GGCTTGGCAGCCGCCCCGCCTGAAC

CTGTAAATATATGGGTAGATTCCAAATACCTGTACTCTTTGCTCAGAACCCTAGTTCTGGGAGCTTG
GCTTCAACCTGACCCCGTACCCT

CCTACGCCCTCCTATATAAAAGCCTCCTCCGACATCCAGCAATCTTTGTTGGCCATGTCCGGAGCC
ACTCTTCAGCATCCCACCCTATTG

CTTCCCTGAACAATTATGTAGATCAACTGCTTCCCTTAGAAACTCCAGAGCAATGGCATAAGCTCAC
CCACTGCAACTCTCGGGCCTTGT

CTCGATGGCCGAACCCACGTATCTCTGCCTGGGACCCCCGTTCCCCCGCTACGCTGTGTGAAACC
TGTCAAAAGCTTAATCCAACTGGAG

GAGGAAAGATGCGAACTATTCAGAGAGGGTGGGCCCCGAACCATATTTGGCAGGCCGATATAACC
CATTATAAATACAAACAGTTCACCT

ACGCTCTGCATGTGTTTGTAGATACTTACTCTGGAGCTACTCATGCCTCGGCGAAGCGTGGGCTCA
CCACTCAAATGACCATTGAGGGCC

TTCTTGAGGCCATAGTGCATCTGGGTCGTCCAAAAAGGCTAAACACTGACCAAGGTGCAAACTACA
CCTCCAAAACCTTTGTCAGGTTTT

GCCAGCAGTTCGGAGTTTCCCTTTCTCATCATGTTCCCTACAACCCCACAAGTTCGGGGTTAGTAG
AACGGACAAATGGACTGCTCAAAC

TTCTTCTATCTAAATATCACCTAGACGAACCCCACCTTCCCATGACTCAGGCCCTTTCTCGAGCCCT
CTGGACTCACAATCAGATTAACC

TCCTACCAATTCTAAAGACCAGATGGGAGCTACACCATTCACCCCCACTTGCTGTCATTTCAGAGG
GCGGAGAAACACCCAAGGGCTCTG

ATAAACTCTTTTTGTACAAGCTCCCCGGGCAAAACAATCGTCGGTGGCTAGGACCACTCCCGGCCC
TAGTCGAAGCCTCGGGAGGCGCTC

TCCTGGCTACTGACCCCCCCGTGTGGGTTCCCTGGCGTTTGCTGAAAGCCTTCAAATGCCTAAAGA
ACGACGGTCCCGAAGACGCCCACA

ACCGATCATCAGATGGGTAAGTCTCACTCTCACTCTCCTCGCTCTCTCTCGGCCCATCCAGACTTG
GAGATGCTCCCTGTCCCTAGGAAA

CCAACAATGGATGACAGCATATAACCAAGAGGCAGAATTTTCCATCTCCATTGACCAAATACTAGAG
GCTCATAATCAGTCACCTTTCTG

TGCCAAGTCTCCCAGATACACCTTGGACTCTGTAAATGGCTATCCTAAGATCTACTGGCCCCCCCC
ACAAGGGCGGCGCCGGTTTGGAGC

CAGGGCCATGGTCACATATGATTGCGAGCCCCGATGCCCTTATGTGGGGGCAGATCGCTTCGACT
GCCCCCACTGGGACAATGCCTCCCA

GGCCGATCAAGGATCCTTTTATGTCAATCATCAGATTTTATTCCTGCATCTCAAACAATGTCATGGAA
TTTTCACTCTAACCTGGGAGAT

ATGGGGATATGATCCCCTGATCACCTTTTCTTTACATAAGATCCCTGATCCCCCTCAACCCGACTTT
CCCCAGTTGAACAGTGACTGGGT

TCCCTCTGTCAGATCATGGGCCCTGCTTTTAAATCAAACAGCACGGGCCTTCCCAGACTGTGCTAT
ATGTTGGGAACCTTCCCCTCCCTG

GGCTCCCGAAATATTAGTATATAACAAAACCATCTCCAGCTCTGGACCCGGCCTCGCCCTCCCGGA
CGCCCAAATCTTCTGGGTCAACGC

GTCCTCGTTTAACACCACCCAAGGATGGCACCACCCTTCCCAGAGGTTGTTGTTCAATGTTTCTCAA
GGCAACGCCTTGTTATTACCTCC

TATCTCCCTGGTTAATCTCTCTACGGCTTCCTCCGCCCCTCCTACCCGGGTCAGACGTAGTCCCGT
CGCAGCCCTGACCTTAGGCCTAGC

CCTGTCAGTGGGGCTCACTGGAATTAATGTGGCCGTGTCTGCCCTTAGCCATCAGAGACTCACCTC
CCTGATCCACGTTCTGGAGCAAGA

TCAGCAACGCTTGATCACAGCAATTAACCAGACCCACTATAATTTGCTTAATGTGGCCTCTGTGGTT
GCCCAGAACCGACGGGGGCTTGA

TTGGTTATACATCCGGCTGGGTTTTCAAAGCCTATGTCCCACAATTAATGAGCCTTGCTGTTTCCTG
CGCATTCAAAATGACTCCATTAT

CCGCCTCGGTGATCTCCAGCCTCTCTCGCAAAGAGTCTCTACAGACTGGCAGTGGCCCTGGAATT
GGGATCTGGGGCTCACTGCCTGGGT

GCGAGAAACCATTCATTCTGTTCTAAGCCTGTTCCTATTAGCCCTTTTTTTGCTCTTCCTGGCCCCC
TGCCTGATAAAATGCTTGACCTC

TCGCCTTTTAAAGCTCCTCCGGCAGGCTCCCCACTTCCCTGAAATCTCCTTAACCCCTAAACCCGAT
TCTGATTATCAGGCCTTGCTACC

ATCTGCACCAGAGATCTACTCTCACCTCTCCCCCGTCAAACCCGATTACATCAACCTCCGACCCTG
CCCTTGATACCCCCGCGTTTCACG

CACCCCCAGGCTGTGGTGGTGCACTGGCTTAGTGGAATAGTCAGTGTACCATCACAAGCCTCTTCT
TGCTGCCGGCACCGAGTTCGAACA

CAGCCCTACCCTGAACCTCTCTGAGTGCATGACTGAGTGTAGCGCAGAGAGATTGTCGCTTCTGCG
TGTCGCTCAGTCATTTTTTATAGC

CGATTGGGGTTCGCGCCCTTCCGTTGCCTGTGACACAGATAAGACCTCTCTCACTTCTGCTTCACC
ATCCCCCTGCCAGCGTTGGTCTAG

TGGAAAGAACTAACGCTGACGGGGGCGATTTCTTGCAGCTGTGCTAGCGGGAGGCTCTGGTGCTGGGGATAAGATGTGGCCCTTAGCACC

ACAGTCTCTGCGCCTTTTGGGTTCGAATCTTCCCCACGCAGCTTCCGCTTTTTACGCCCTGTTGCACACCCTTTCTAGAGATACCTGAAA

ATCTCAGCTCGCACCCTGAGGAAGGTTGTGGCTCAGAGGTTAAAATAGCTCGAGCCGCAACCTCCCTTTCTTTTTATTCCACCCTCGCAA

GGCCCCGGGTTCTGAGCCCCCTAACGGAGGTTCAAAATTTCCTCTACTAGGGGATGCTCGGGTCCAAGTGTGCACAATATCTCTTCCAAA

AGGTCCTGATGAACGTCTTCCCATGTAACAAGCCCCAGCAGAGACATTCCAGCCACATCCAGCAGCATTTGGGCCGCCTTTTCTAACAGT

GCCCATAAAGTCCCTTCCGTTTCCACAACGGCTGCCTCTGCTTCTTCTATTTCCACCTCGGCACCGACTCCCCCGCCGAGCCCTTCGAGC

TCTTCGGGATCCATTACCTGATAACGACAAAATTATTTCTTGTCTTTTAAGCAAGTGTTGTTGGTTGGGGGCCCCACTCTCTACATGCCT

GCCCGGCCCTGGTTTTGTCCAATGATGTCACCATCGATGCCTGGTGCCCCCTCTGCGGGCCCCATAAGCGACTCCAATTCGAAAGGATCG

ACACCACGCTCACCTGCGAGACCCACCGTATCAACTGGACCGCCGATGGACGACCTTGCGGCCTCAATGGAACGTTGTTCCCTCGACTAC

ATGTCTCCGAGACCCGCCCCCAAGGGCCCCGACGACTCTGGATCAACTGCCCCCTTCCGGCCGTTCGCGCTCAGCCCGGCCCGGTTTCAC

TTTCCCCCTTCGAGCGGTCCCCCTTCCAGCCCTACCAATGCCAATTGCCCTCGGCCTCTAGCGACGGTTGCCCCATTATCGGGCACGGCC

TTCTTCCCTGGAACAACTTAGTAACGCATCCTGTCCTCGGAAAAGTCCTTATATTAAATCAAATGGCCAATTTTTCCTTACTCCCCTCCT

TCGATACCCTCCTTGTGGACCCCCTCCGGCTGTCCGTCTTTGCCCCAGACACCAGGGGAGCCATACGTTATCTCTCCACCCTTTTGACGC

TATGCCCAGCTACTTGTATTCTACCCCTAGGCGAGCCCTTCTCTCCTAATGTCCCCATATGCCGCTTTCCCCGGGACTCCAATGAACCCC

CCCTTTCAGAATTCGAGCTGCCCCTTATCCAAACGCCCGGCCTGTCCTGGTCTGTCCCCGCGATCGACCTATTTCTAACCGGTCCCCCTT

CCCCATGCGACCGGTTACACGTATGGTCCAGTCCTCAGGCCTTACAGCGCTTCCTTCATGACCCTACGCTAACCTGGTCCGAATTGGTTG

CTAGCAGAAAGATAAGACTTGATTCCCCCTTAAAATTACAACTGCTAGAAAATGAATGGCTCTCCCGCCTTTTTTGAGGGGGAGTCATTT

GTATGAAAGATCATGCCGACCTAGGCGCCGCCACCGCCCCGTAAACCAGACAGAGACGTCAGCTGCCAGAAAAGCTGGTGACGGCAGCTG

GTGGCTAGAATCCCCGTACCTCCCCAACTTCCCCTTTCCCGAAAAATCCACACCCTGAGCTGCTGACCTCACCTGCTGATAAACTAATAA

AATGCCGGCCCTGTCGAGTTAGCGGCACCAGAAGCGTTCTTCTCCTGAGACCCTCGTGCTCAGCTCTCGGTCCTGAGCTCTCTTGCTCCC

GAGACCTTCTGGTCGGCTATCCGGCAGCGGTCAGGTAAGGCAAACCACGGTTTGGAGGGTGGTTCTCGGCTGAGACCACCGCGAGCTCTA

TCTCCGGTCCTCTGACCGTCTCCACGTGGACTCTCTCCTTTGCCTCCTGACCCCGCGCTCCAAGGGCGTCTGGCTTGCACCCGCGTTTGT

TTCCTGTCTTACTTTCTGTTTCTCGCGGCCCGCGCTCTCTCCTTCGGCGCCCTCTAGCGGCCAGGAGAGACCGGCAAACA

>Bos-1439_BLV_consensus 8720 bp

TGTATGAAAGATCATGCCGACCTAGGCGCCGCCACCGCCCCGTAAACCAGACAGAGACGTCAGCTGCCAGAAAAGCTGGTGACGGCAGCT

GGTGGCTAGAATCCCCGTACCTCCCCAACTTCCCCTTTCCCGAAAAATCCACACCCTGAGCTGCTGACCTCACCTGCTGATAAATTAATA

AAATGCCGGCCCTGTCGAGCTAGCGGCACCAGAAGCGTTCTTCTCCTGAGACCCTCGTGCTCAGCTCTCGGTCCTGAGCTCTCTTGCTCC

CGAGACCTTCTGGTCGGCTATCCGGCAGCGGTCAGGTAAGGCAAACCACGGTTTGGAGGGTGGTTCTCGGCTGAGACCACCGCGAGCTCT

ATCTCCGGTCCTCTGACCGTCTCCACGTGGACTCTCTCCTTTGCCTCCTGACCCCGCGCTCCAAGGGCGTCTGGCTTGCACCCGCGTTTG

TTTCCTGTCTTACTTTCTGTTTCTCGCGGCCCGCGCTCTCTCCTTCGGCGCCCTCTAGCGGCCAGGAGAGACCGGCAAACAATTGGGGGC

TCGTCCGGGATTGATCACCCCGGAACCCTAACAACTCTCTGGACCCACCCCCTCGGCGGCATTTTGGGTCTCTCCTTCAAATTATATCAT

GGGAAATTCCCCCTCCTATAACCCCCCCGCTGGTATCTCCCCCTCAGACTGGCTCAACCTTCTGCAAAGCGCGCAAAGGCTCAATCCGCG

ACCCTCTCCTAGCGATTTTACCGATTTAAAGAATTACATCCATTGGTTTCATAAGACCCAGAAAAAACCATGGACTTTCACTTCTGGTGG

CCCCACCTCATGTCCACCCGGGAGATTCGGCCGGGTTCCCCTTGTCTTGGCCACCCTAAACGAAGTGCTCTCAAACGAAGGGGGCGCCCC

GGGTGCATCGGCCCCCAGAAGAACAACCCCCCCCTTATGACCCCCCCGCCGTTTTGCCAATCATATCTGAAGGGAATCGCAACCGCCATCG

TGCTTGGGCACTCCGAGAATTACAAGATATCAAAAAGGAAATTGAAAATAAGGCACCGGGTTCGCAAGTATGGATACAAACACTACGACT

TGCAATCCTGCAGGCCGACCCTACTCCGGCTGACCTAGAACAACTTTGCCAATATATTGCTTCCCCGGTCGACCAAACGGCCCATATGAC

CAGCCTAACGGCAGCAATAGCCGCCGCTGAAGCGGCCAACACCCTCCAGGGTTTTAACCCCCAAAACGGGACCCTAACCCAACAATCCGC

TCAGCCCAACGCCGGGGATCTTAGAAGTCAATATCAAAACCTCTGGCTTCAGGCCTGGAAAAATCTCCCTACTCGTCCTTCAGTACAACC

TTGGTCCACCATCGTCCAAGGCCCCGCCGAAAGCTATGTAGAGTTTGTCAACCGGTTACAAATTTCATTAGCTGACAACCTTCCCGACGG

AGTCCCTAAGGAACCCATTATTGACTCCCTTAGTTATGCAAATGCTAACAAAGAGTGTCAGCAAATT
TTGCAGGGGCGAGGCCTAGTGGC

CGCCCCGGTGGGGCAAAAACTGCAGGCTTGCGCACATTGGGCCCCCAAGATGAAACAGCCTGCAA
TTCTCGTCCACACCCCAGGGCCCAA

GATGCCCGGGCCTCGGCAACCGGCCCCCAAAAGGCCTCCCCCAGGACCATGCTATCGATGCCTC
AAAGAAGGCCATTGGGCCCGGGATTG

TCCTACCAAGGCCACCGGCCCCCCTCCGGGACCTTGCCCCATATGTAAAGATCCTTCCCATTGGAA
ACGAGACTGTCCAACCCTCAAATC

AAAAAACTAATAGAGGGGGGACTTAGCGCCCCCCAAACCATAACACCTATAACAGATTCTCTTAGT
GAGGCCGAATTAGAATGCTTACTT

TCTATTCCTCTGGCTCGCAGCCGTCCCTCCGTGGCTGTATACCTGTCTGGCCCCTGGCTGCAGCC
CTCTCAGAATCAAGCCCTCATGCTT

GTGGACACCGGGGCTGAAAATACGGTTCTCCCACAAAATTGGCTGGTTCGAGATTACCCACGGATC
CCCGCCGCAGTGCTCGGAGCAGGG

GGAGTCTCCCGGAACAGATACAATTGGCTACAAGGCCCTCTGACCCTGGCTCTAAAACCAGAGGG
TCCCTTTATCACCATCCCAAAAATT

TTAGTTGACACTTTCGACAAATGGCAAATTTTAGGACGGGACGTTCTCTCCCGCCTACAGGCCTCTA
TCTCCATACCTGAGGAAGTACGC

CCCCCTGTGGTAGGCGTCTTGGATGCCCCCCCGAGCCACATTGGATTAGAACATCTGCCCCCCCC
ACCTGAGGTGCCTCAATTCCCTTTA

AACTAGAACGCCTCCAGGCCCTTCAAGACCTGGTCCATCGCTCTCTGGAGGCAGGTTATATCTCCC
CCTGGGACGGGCCAGGCAATAATC

CAGTCTTCCCGGTACGGAAACCAAATGGCGCCTGGAGGTTTGTGCATGACCTACGAGCTACAAATG
CTCTTACAAAGCCCATTCCGGCAC

TCTCCCCCGGACCGCCAGACCTTACCGCTATCCCTACACACCTTCCACATATCATTTGCCTAGATCT
CAAAGATGCCTTCTTCCAGGTTC

CAGTCGAAGACCGCTTCCGCTCCTACTTTGCTTTCACCCTCCCTACCCCCGGGGGACTCCAACCTC
ATAGACGCTTTGCCTGGCGGGTCC

TACCTCAAGGCTTCATTAACAGCCCAGCTCTTTTCGAACGAGCACTACAGGAACCTCTTCGCCAAG
TTTCCGCCGCCTTTTCCCAGTCTC

TTCTGGTGTCCTATATGGACGATATCCTTATCGCTTCGCCTACAGAAGAACAGCGGTCACAATGTTA
TCAAGCCCTGGCTGCCCGCCTCC

GGGACCTAGGGTTTCAGGTGGCATCCGAAAAGACTCGCCAGACGCCTTCGCCCGTCCCCTTCTTG
GGACAAATGGTCCATGAGCAGATTG

TCACCTACCAGTCCCTACCTACCTTGCAGATCTCATCCCCAATTTCTCTTCACCAATTACAGGCGGT
CTTAGGAGACCTCCAATGGGTCT

CTAGGGGCACACCCACTACCCGCCGGCCCCTGCAACTTCTCTACTCTTCCCTTAAAGGCATCGATG
ACCCTAGGGCCATCATCCAGCTTT

CCCCGGAACAGCTGCAAGGCATTGCAGAGCTTCGACAAGCCCTGTCCCATAACGCAAGATCTAGA
TATAACGAGCAAGAACCCCTGCTAG

CCTACGTACACCTAACCCGGGCGGGGTCCACCCTGGTACTCTTCCAAAAGGGCGCTCAATTTCCCCTGGCCTACTTTCAGACCCCCTTGA

CTGACAACCAAGCCTCACCTTGGGGCCTCCTTCTCCTGCTGGGATGCCAATACCTGCAGACTCAGGCCTTAAGCTCGTATGCCAAGCCCA

TACTTAAATATTATCACAATCTTCCTAAAACCTCTCTAGACAATTGGATTCAATCATCTGAGGACCCTCGAGTCCAGGAGTTGCTGCAAT

TGTGGCCCCAGATTTCCTCTCAGGGAATACAGCCCCCGGGCCCTTGGAAGACCTTAATCACCAGGGCAGAGGTTTTTTTGACGCCCCAGT

TCTCCCCTGATCCGATTCCTGCGGCCCTTTGCCTCTTTAGTGACGGGGCTACAGGACGAGGAGCATATTGCTTGTGGAAAGACCACCTTT

TAGACTTTCAGGCCGTTCCGGCTCCAGAATCCGCTCAAAAGGGAGAACTAGCAGGTCTCTTGGCGGGCTTAGCAGCCGCCCCGCCTGAAC

CTGTAAATATATGGGTAGATTCCAAATACCTGTACTCTTTGCTCAGAACCCTAGTTCTGGGAGCTTGGCTTCAACCTGACCCCGTACCCT

CCTACGCCCTCCTATATAAAAGCCTCCTCCGACATCCAGCAATCTTTGTTGGTCATGTCCGGAGCCACTCTTCAGCATCCCACCCTATTG

CTTCCCTGAACAATTATGTAGATCAACTGCTTCCCTTAGAAACTCCAGAGCAATGGCATAAGCTCACCCACTGCAACTCGCGGGCCTTAT

CTCGATGGCCGAACCCACGTATCTCTGCCTGGGACCCCCGTTCCCCCGCTACGCTGTGTGAAACCTGTCAAAAGCTTAATCCAACTGGAG

GAGGAAAGATACGAACTATTCAGAGAGGGTGGGCCCCGAATCATATTTGGCAGGCCGATATAACCCATTATAAATACAAACAGTTCACCT

ACGCTCTGCATGTGTTTGTAGATACTTACTCTGGAGCTACTCATGCCTCGGCGAAGCGTGGGCTCACCACTCAAATGACCATTGAGGGCC

TTCTTGAGGCCATAGTGCATCTGGGTCGTCCAAAAAAGCTAAACACTGACCAAGGTGCAAACTACACCTCCAAAACCTTTGTCAGGTTTT

GCCAGCAGTTCGGAGTTTCCCTTTCTCATCATGTTCCCTACAACCCCACAAGTTCGGGGTTAGTAGAACGGACAAATGGACTGCTCAAAC

TTCTTCTATCTAAATATCACCTAGACGAACCCCACCTTCCCATGACTCAGGCCCTTTCTCGAGCCCTCTGGACTCACAATCAGATTAACC

TCCTACCAATTCTAAAGACCAGATGGGAGCTACACCATTCACCCCCACTTGCTGTCATTTCAGAGGGCGGAGAAACACCCAAGGGCTCTG

ATAAACTCTTTTTGTACAAGCTCCCCGGGCAAAACAATCGTCGGTGGCTAGGACCACTCCCGGCCCTAGTTGAAGCCTCGGGAGGCGCTC

TCCTGGCTACTGACCCCCCCGTGTGGGTTCCCTGGCGTTTGCTGAAAGCCTTCAAATGCCTAAAGAACGACGGTCCCGAAGACGCCCACA

ACCGATCGTCAGATGGGTAAGTCTCACTCTCACTCTCCTCGCTCTCTGTCGGCCCATCCAGACTTGGAGATGCTCCCTGTCCCTAGGAAA

CCAACAATGGATGACAGCATATAACCAAGAGGCAAAATTTTCCATCTCCATTGACCAAATACTAGAGGCTCATAATCAGTCACCTTTCTG

TGCCAAGTCTCCCAGATACACCTTGGACTCTGTAAATGGCTATCCTAAGATCTACTGGCCCCCCCC
ACAAGGGCGGCGCCGGTTTGGAGC

CAGGGCCATGGTCACATATGATTGCGAGCCCCGATGCCCTTATGTGGGGGCAGATCGCTTCGACT
GCCCCCACTGGGACAATGCCTCCCA

GGCCGATCAAGGGTCCTTTTATGTCAATCATCAGATTTTATTCCTGCATCTCAAACAATGTCATGGA
ATTTTCACTCTAACCTGGGAGAT

ATGGGGATATGATCCCCTGATCACCTTTTCTTTACATAGGATCCCTGATCCCCCTCAACCCGACTTT
CCCCGGTTGAGCAGTGACTGGGT

TCCCTCTGTCAGATCATGGGCCCTGCTTTTAAATCAAACAGCACGGGCCTTCCCAGACTGTGCTAT
ATGTTGGGAACCTTCCCCTCCCTG

GGCTCCCGAAATATTAGTATATAACAAAACCATCTCCAGCTCTGGACCCGGCCTCGCCCTCCCGGA
CGCCCAAATCTTCTGGGTCAACAC

GTCCTCGTTTAACACCACCCAAGGATGGCACCACCCTTCCCAGAGGTTGTTGTTCAATGTTTCTCAA
GGCAACGCCTTGTTATTACCTCC

TATCTCCCTGGTTAATCTCTCTACGGCTTCCTCCGCCCCTCCTACCCGGGTCAGACGTAGTCCCGT
CGCAGCCCTGACCTTAGGCCTAGC

CCTGTCAGTGGGGCTCACTGGAATTAATGTGGCCGTGTCTGCCCTTAGCCATCAGAGACTCACCTC
CCTGATCCACGTTCTGGAGCAAGA

TCAGCAACGCTTGATCACAGCAATTAACCAGACCCACTATAATTTGCTTAATGTGGCCTCTGTGGTT
GCCCAGAACCGACGGGGGCTTGA

TTGGTTGTACATCCGGCTGGGTTTTCAAAGCCTATGTCCCACAATTAATGAGCCTTGCTGTTTCCTG
CGCATTCAAAATGACTCCATTAT

CCGCCTTGGTGATCTCCAGCCTCTCTCGCAAAGAGTCTCTACAGACTGGCAGTGGCCCTGGAATTG
GGATCTGGGGCTCACTGCCTGGGT

GCGAGAAACCATTCATTCTGTTCTAAGCCTGTTCCTATTAGCCCTTTTTTTGCTCTTCCTGGCCCCC
TGCCTGATAAAATGCTTGACCTC

TCGCCTTTTAAAGCTCCTCCGGCAGGCTCCCCACTTCCCTGAAATCTCCTTAACCCCTAAACCCGAT
TCTGATTATCAGGCCTTGCTACC

ATCTGCACCAGAGATCTACTCTCACCTCTCCCCCGTCAAACCCGATTACATCAACCTCCGACCCTG
CCCTTGATACCCCCGCGTTTCACG

CACCCCCAGGCTGTGGTGGTGCACTGGCTTAGTGGAATAGTCAGTGTACCATCACAAGCCTCTTCT
TGCTGCCAGCACCGAGTTCGAACA

CAGCTCTACCCTGAGCCTCTCTGAGTGCATGACTGAGTGTAGCGCAGAGAGATTGTCGCTTCTGCG
TGTCGCTCAGTCATTTTTTATAGC

CGATTGGGGTTCGCGCCCTTCCGTTGCCTGTGACACAGATAAGACCTCTCTCACTTCTGCTTCACC
ATCCCCCTGCCAGCGTTGGTCTAG

TGGAAAGAACTAACGCTGACGGGGGCGATTTCTTGCAGCTGTGCTAGCGGGAGGCTCTGGTGCTG
GGGATAAGATGTGGCCCTTAGCACC

ACAGTCTCTGCGCCTTTTGGGTTCGAATCTTCCCCACGCAGCTTCCGCTTTTTACGCCCTGTTGCA
CACCCTTTCTAGAGATACCTGAAA

ATCTCAGCTCGCACCCTGAGGAAGGTTGTGGCTCAGAGGTTAAAATAGCTCGAGCCGCAACCTCCCTTTCTTTTTATTCCACCCTCGCAA

GGCCCCGGGTTCTGAGCCCCCTAACGGAGGTTCAAAATTTCCTCTACTAGGGGATGCTCGGGTCCAAGTGTGCACAATATCTCTTCCAAA

AGGTCCTGATGAACGTCTTCCCATGTAACAAGCCCCAGCAGAGACATTCCAGCCACATCCAGCAGCATTTGGGCCGCCTTTTCTAACAGT

GCCCATAAAGTCCCTTCCGTTTCCACAACGGCTGCCTCTGCATCTTCTATTTCCACCTCGGCACCGACTCCCCCGCCGAGCCCTTCGAGC

TCTTCGGGATCCATTACCTGATAACGACAAAATTATTTCTTGTCTTTTAAGCAAGTGTTGTTGGTTGGGGGCCCCACTCTCTACATGCCT

GCCCGGCCCTGGTTTTGTCCAATGATGTCACCATCGATGCCTGGTGCCCCCTCTGCGGGCCCCATGAGCGACTCCAATTCGAAAGGATCG

ACACCACGCTCACCTGCGAGACCCACCGTATCAACTGGACCGCCGATGGACGACCTTGCGGCCTCAATGGAACGTTGTTCCCTCGACTGC

ATGTCTCCGAGACCCGCCCCCAAGGGCCCCGACGACTCTGGATCAACTGCCCCCTTCCGGCCGTTCGCGCTCAGCCCGGCCCGGTTTCAC

TTTCCCCCTTCGAGCGGTCCCCCTTCCAGCCCTACCAATGCCAATTGCCCTCGGCCTCTAGTGACGGTTGCCCCATTATCGGGCACGGCC

TTCTTCCCTGGAACAACTTAGTAACGCATCCTGTCCTCGGAAAAGTCCTTATATTAAATCAAATGGCCAATTTTTCCTTACTCCCCTCCT

TCGATACCCTCCTTGTGGACCCCCTCCGGCTGTCCGTCTTTGCCCCAGACACCAGGGGAGCCATACGTTATCTCTCCACCCTTTTGACGC

TATGCCCAGCTACTTGTATTCTACCCCTAGGCGAGCCCTTCTCTCCTAAGGTCCCCATATGCCGCTTTCCCCGGGACTCCAATGAACCCC

CCCTTTCAGAATTCGAGCTGCCCCTTATCCAAACGCCCGGCCTGTCTTGGTCTGTCCCCGCGATCGACCTATTCCTAACCGGTCCCCCTT

CCCCATGCGACCGGTTACACGTATGGTCCAGTCCTCAGGCCTTACAGCGCTTCCTTCATGACCCTACGCTAACCTGGTCCGAATTGGTTG

CTAGCAGAAAAATAAGACTTGATTCCCCCTTAAAATTACAACTGCTAGAAAATGAATGGCTCTCCCGCCTTTTTTGAGGGGGAGTCATTT

GTATGAAAGATCATGCCGACCTAGGCGCCGCCACCGCCCCGTAAACCAGACAGAGACGTCAGCTGCCAGAAAAGCTGGTGACGGCAGCTG

GTGGCTAGAATCCCCGTACCTCCCCAACTTCCCCTTTCCCGAAAAATCCACACCCTGAGCTGCTGACCTCACCTGCTGATAAATTAATAA

AATGCCGGCCCTGTCGAGCTAGCGGCACCAGAAGCGTTCTTCTCCTGAGACCCTCGTGCTCAGCTCTCGGTCCTGAGCTCTCTTGCTCCC

GAGACCTTCTGGTCGGCTATCCGGCAGCGGTCAGGTAAGGCAAACCACGGTTTGGAGGGTGGTTCTCGGCTGAGACCACCGCGAGCTCTA

TCTCCGGTCCTCTGACCGTCTCCACGTGGACTCTCTCCTTTGCCTCCTGACCCCGCGCTCCAAGGGCGTCTGGCTTGCACCCGCGTTTGT

TTCCTGTCTTACTTTCTGTTTCTCGCGGCCCGCGCTCTCTCCTTCGGCGCCCTCTAGCGGCCAGGA
GAGACCGGCAAACA

>HIV_U1_CHRX 9656 bp

TGGAAGGGCTAATTTGGTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAG
GCTACTTCCCTGATTGGCAGAACT

ACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCTTCAAGTTAGTACCAGTTG
AGCCAGGGCAGGTAGAAGAGGCCA

ATGAAGGAGAGAACAACACCTTGTTACACCCTATGAGCCTGCATGGGATGGAGGACCCGGAGGGA
GAAGTATTAGTGTGGAAGTTTGACA

GCCTCCTAGCATTTCGTCACATGGCCCGAGAGCTGCATCCGGAGTACTACAAAGACTGCTGACATC
GAGTTTTCTACAAGGGACTTTCCG

CTGGGGACTTTCCAGGGAGGTGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCTG
CATATAAGCAGCTGCTTTTTGCCTGT

ACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAGCTAGGGAACCCACT
GCTTAAGCCTCAATAAAGCTTGCCT

TGAGTGCTACAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCC
TTTTAGTCAGTGTGGAAAATCTCT

AGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGTAAAGCCAGAGGAGATCTCTCGACGCAG
GACTCGGCTTGCTGAAGCGCGCACGG

CAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGA
GAGAGATGGGTGCGAGAGCGTCGGTA

TTAAGCGGGGGAGAATTAGATAAATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAACAATAT
AGACTAAAACATATAGTATGGGCA

AGCAGGGAACTAGAACGATTCGCAGTTAATCCTGGCCTTTTAGAGACATCAGAAGGCTGTAGACAA
ATACTGAGACAGCTACAACCATCC

CTTCAGACAGGATCAGAAGAACGTAGATCATTATTTAATACAGTAGCAGTCCTCTATTGTGTGCATC
AAAGGATAGATGTAAAAGACACC

AAGGAAGCCTTAGATAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAGGCACAGCAAGCAGCA
GCTGACACAGGAAACAGCAGCCAG

GTCAGCCAAAATTACCCTATAGTGCAGAACCTCCAGGGGCAAATGGTACATCAGGCCATATCACCT
AGAACTTTAAATGCATGGGTAAAA

GTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTAATACCCATGTTTTCAGCATTATCAGAAGGAGCC
ACCCCACAAGATTTAAATACCATG

CTAAACACAGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCT
GCAGAATGGGATAGATTGCATCCA

GTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAAC
TACTAGTACCCTTCAGGAACAAATA

GGATGGATGACACATAATCCACCTATCCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTAT

AGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGATTCTATAAAACTCTAAGAGCCGAG

CAAGCTTCACAAGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCGAACCCAGATTGTAAGACTATTTTAAAAGCATTG

GGACCAGGAGCGACACTAGAAGAAATGATGACAGCATGTCAGGGAGTGGGGGGACCCGGCCATAAAGCAAGAGTTTTGGCTGAAGCAATG

AGCCAAGTAACAAATCCAGCTAACATAATGATACAGAAAGGCAATTTTAGGAACCAAAGAAAGACTGTTAAGTGTTTCAATTGTGGCAAA

GAAGGGCACATAGCCAAAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTACT

GAGAGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCCCACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCC

CCACCAGAAGAGAGCTTCAGGTTTGGGGAAGAGACAACAACTCCCTCTCAGAAGCAGGAGACGATAGACAAGGAACTGTATCCTTTAGCT

TCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAGGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATG

ATACAGTATTAGAAGAAATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATG

ATCAGATACTCATAGAAATCTGCGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGT

TGACTCAGATTGGCTGCACTTTAAATTTTCCCATTAGTCCTATTGAGACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAG

TTAAACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAGTAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTG

GGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTCAGAGAACTTA

ATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCACATCCTGCAGGGTTAAAACAGAAAAAATCAGTAACAGTACTGGATG

TGGGCGATGCATATTTTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAACAATGAGACACCAG

GGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAGTGTAGCATGACAAAAATCTTAGAGCCTT

TTAGAAAACAAAATCCAGACATAGTCATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAA

AAATAGAGGAACTGAGACAACATCTGTTGAGGTGGGGATTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGG

GTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAGGACAGCTGGACTGTCAATGACATACAGAAATTAG

TGGGAAAATTGAATTGGGCAAGTCAGATTTATGCAGGGATTAAAGTAAGGCAATTATGTAAACTTCTTAGGGGAACCAAAGCACTAACAG

AAGTAGTACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGGGAGATTCTAAAAGAACCGGTACATGGAGTGTATTATGACC

CATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAA

CAGGAAAGTATGCAAGAACGAAGGGTGCCCACACTAATGATGTGAAACAATTAACAGAGGCAGTGCAAAAAATAGCCACAGAAAGCATAG

TAATATGGGGAAAGACTCCTAAATTTAAATTACCCATACAAAAGGAAACATGGGAAGCATGGTGGACAGAGTATTGGCAAGCCACCTGGA

TTCCTGAGTGGGAGTTTGTCAATACCCCTCCCTTAGTGAAGTTATGGTACCAGTTAGAAAAAGAACCCATAGTAGGAGCAGAAACTTTCT

ATGTAGATGGGGCAGCCAATAGGGAAACTAAATTAGGAAAAGCAGGATATGTAACTGACAGAGGAAGACAAAAAGTTGTCCCCCTAACGG

ACACAACAAATCAGAAGACTGAGTTACAAGCAATTCATCTAGCTTTGCAGGATTCAGGATTAGAAGTAAACATAGTAACAGACTCACAAT

ATGCATTGGGAATCATTCAAGCACAACCAGATAAGAGTGAATCAGAGTTAGTCAGTCAAATAATAGAGCAGTTAATAAAAAAGGAAAAAG

TCTACCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTGGTCAGTGCTGGAATCAGGAAAGTACTAT

TTTTAGATGGAATAGATAAGGCCCAAGAAGAACATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACCTGCCACCTG

TAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGCCCCGGAATAT

GGCAGCTAGATTGTACACATTTAGAAGGAAAAGTTATCTTGGTAGCAGTTCATGTGGCCAGTGGATATATAGAAGCAGAAGTAATTCCAG

CAGAGACAGGGCAAGAAACAGCATACTTCCTCTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAGTACATACAGACAATGGCAGCAATT

TCACCAGTACTACAGTTAAAGCCGCCTGTTGGTGGGCGGGGATCAAGCAGGAATTTGGCATTCCCTACAATCCCCAAAGTCAAGGAGTAA

TAGAATCTATGAATAAAGAATTAAAGAAAATTATAGGACAGGTAAGAGATCAGGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTAT

TCATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTA

AAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGAGACAGCAGAGATCCAGTTTGGAAAGGACCAGCAAAGC

TCCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAAGAAGAAAAGCAAAGATCATCAGGGATT

ATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAACACATGGAAAAGATTAGTAAAACACCATATGTAT

ATTTCAAGGAAAGCTAAGGACTGGTTTTATAGACATCACTATGAAAGTACTAATCCAAAAATAAGTTCAGAAGTACACATCCCACTAGGG

GATGCTAAATTAGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCATTTGGGTCAGGGAGTCTCCATAGAATGGAGG

AAAAAGAGATATAGCACACAAGTAGACCCTGACCTAGCAGACCAACTAATTCATCTGCACTATTTTGATTGTTTTTCAGAATCTGCTATA

AGAAATACCATATTAGGACGTTTAGTTAGTCCTAGGTGTGAATATCAAGCAGGACATAACAAGGTAGGATCTCTACAGTACTTGGCACTA

GCAGCATTAATAAAACCAAAACAGATAAGGCCACCTTTGCCTAGTGTTAGGAAACTGACAGAGGACAGATGGAACAAGCCCCAGAAGACC

AAGGGCCACAGAGGGAGCCATACAATGAATGGACACTAGAGCTTTTAGAGGAACTTAAGAGTGAAGCTGTTAGACATTTTCCTAGGATAT

GGCTCCATAACTTAGGACAACATATCTATGAAACTTATGGGGATACTTGGGCAGGAGTGGAAGCCATAATAAGAATTCTGCAACAACTGC

TGTTTATTCATTTCAGTATTGGATGTCAACATAGCAGAATAGGCATTATTCGACAGAGGAGAGCAAGGAATGGAGCCAGTAGATCCTAGA

CTAGAGCCCTGGAAGCTTCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAAGTTTGT

TTCACAACAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCCTCAAGACAGTGAGACTCATCAAGTT

TCTCTATCAAAGCAGTGAGTAGTACATGTAATGCTTTCTTTCAATAGTTGTGTGGACCATAGTGTTCATAGAATACAGGAAAGTATTAAG

ACAAAGAAAAGTAGACAGGATAATTGATAGAATAAGAGAAAGAGCAGAAGACAGTGGCAATGAGAGCGAAGGGGACCAGGAAGAATTATT

AGCACTTGTGGAGATGGGGCACGATGCTCCTTGGGATGTTAATGATCTGTAGTGCTGCAGAACAATTGTGGGTCACAGTCTATTATGGGG

TACCTGTGTGGAAAGAAGCAACCACCACTCTATTTTGTGCCTCGGATGCTAAAGCATATGACACAGAGGTACATAATGTTTGGGCCACAC

ATGCCTGTGTACCCACAGACCCTAACCCACAAGAAGTAGTATTGCAAAATGTGACAGAAAATTTTAACATGTGGAAAAATAACATGGTAG

AACAAATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAACCATGTGTAAAATTAACCCCACTCTGTGTTACCTTAAATTGCA

CTGATTTGACGAATGCTACTTATGCCAATGGTAGTAGTGAGGAGAGAGGAGAAATAAGAAATTGCTCTTTCAATGTCACCACAATCATAA

GAAATAAGATACAAAAAGAATATGCACTTTTTTATAGACTTGATATAGAACCAATAGATAAGGATAATACAAGCTATACATTGATAAATT

GTAACACCTCAGTCATTACACAGGCTTGTCCAAAGGTATCCTTTGAACCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTC

TAAAGTGTAACGATAAGAAATTCAATGGAACAGGACCATGTACAAATGTCAGTACAGTACAATGTACACATGGAATTAGGCCAGTAGTGT

CAACTCATCTGCTATTAAATGGCAGTCTAGCAGAAGGAGAGGTAGTAATTAGGTCTGAAAATTTCACGAACAATGTTAAAACCATAATAG

TACAGCTGAATGAATCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAGAAAAGGTATAGCTATAGGACCGGGGAGAACACTCT

ATGCAAGAGAAAAAATAATAGGAGATATAAGACAAGCACATTGTAACCTTAGTAGAGCAAAATGGAATGACACTTTAAAACAGATAGTTA

CAAAATTAAAAGAACAATTTAGGAATAAAACAATAGTCTTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATT

GTGGAGGGGAATTTTTCTACTGTAAGACAACACAGCTGTTTAATAGTACTTGGCTGTTTAATAGTACTTGGAATGATACTGAAAGGTCAG

ATAACAATGAAACTATCATACTCCCATGCAGAATAAAACAAATTATAAACAGGTGGCAGGAAGTAGGAAAAGCAATATATGCCCCTCCCA

TCAGTGGACAAATTAGATGCTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTGATAAAGAGAACAGTACCACCGAGATCT

TCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAGCCATTAGGAGTAGCAC

CCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAGCGCTAGGAGCTCTGTTTCTTGGGTTCTTGGGAGCAGCAG

GAAGCACTATGGGCGCAGCGTCAATGGCGCTGACGGTACAGACCAGACAATTAATGTCTGGTATAGTGCAACAGCAAAACAATTTGCTGA

AGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAGTCCTGGCTGTGGAAAGATACC

TAAAGGATCAACAGCTCCTAGGGATGTGGGGTTGCTCTGGAAAACTCATCTGCACCACTACTGTGCCTTGGAATGCTAGTTGGAGTAATA

AATCTCTGGATAAGATTTGGGATAACATGACCTGGATGGAGTGGGAGAGAGAAATTGACAATTACACAGGCTTAATATACACCTTAATTG

AAGAATCGCAGTACCAACAAGAAAAGAATGAACAAGAATTATTGGAATTGGATAAGTGGGCAAGTTTGTGGAATTGGTTTGACATAACAA

AGTGGCTGTGGTATATAAAAATATTCATAATGATAGTAGGAGGTTTAATAGGTTTAAGAATAGTTTTTACTGTACTTTCTATAGTGAATA

GAGTTAGGCAGGGATACTCACCATTATCGTTCCAGACCCGCCTCCCAACCCAGAGGGGACCCGACAGGCCCGAAGGAATCGAAGAAGAAG

GTGGAGAGAGAGACAGAGACAGATCCGGACCATTAGTGAATGGATTCTTAGCACTCATCTGGGTCGATCTGCGGAGCCTGTGCCTCTTCA

GCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACTTCTGGGACGCAGGGGGTGGGAAGTCCTCAAATATTGTT

GGAATCTCCTGCAGTATTGGGGTCAGGAACTAAAGAATAGTGCTATTAGTTTGCTTAATGCCACAGCTGTAGCAGTAGCTGAGGGGACAG

ATAGGGTTATAGAAGTAGCACAAAGAATTGGTAGGGGTATCCTCCATATCCCTAGAAGAATAAGACAAGGCTTAGAAATTGCTTTGCTAT

AAGATGGGTGGCAAGTGGTCAAAACGTAGTATGGCTGGATGGTCTGCTGTAAGAGAAAGAATGAAACGAGCTGAACCAGCAGCAGATGGG

GTGGGAGCAGTATCTCGAGACTTGGAAAAACATGGAGCAATCACAAGTAGCAATACAGCAGCTACTAATGCTGATTGTGCCTGGCTAGAA

GCACAAGAGGATGAGGAGGTGGGTTTTCCAGTCAGACCTCAGGTACCTTTAAGACCAATAACTCGCAGGGCAGCTATGGATCTTAGCCAC

TTTTTAAAAGAAGAGGGGGGACTGGAAGGGCTAATTTGGTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGC

TACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCTTCAAGTTAGTACCAGTTGAG

CCAGGGCAGGTAGAAGAGGCCAATGAAGGAGAGAACAACACCTTGTTACACCCTATGAGCCTGCATGGGATGGAGGACCCGGAGGGAGAA

GTATTAGTGTGGAAGTTTGACAGCCTCCTAGCATTTCGTCACATGGCCCGAGAGCTGCATCCGGAGTACTACAAAGACTGCTGACATCGA

GTTTTCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGTGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCTGCATA

TAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAGCTAGGGAACCCACTGCT

TAAGCCTCAATAAAGCTTGCCTTGAGTGCTACAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTT

TTAGTCAGTGTGGAAAATCTCTAGCA

>HIV_U1_CHR2 9656 bp

TGGAAGGGCTAATTTGGTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACT

ACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCTTCAAGTTAGTACCAGTTGAGCCAGGGCAGGTAGAAGAGGCCA

ATGAAGGAGAGAACAACACCTTGTTACACCCTATGAGCCTGCATGGGATGGAGGACCCGGAGGGAGAAGTATTAGTGTGGAAGTTTGACA

GCCTCCTAGCATTTCGTCACATGGCCCGAGAGCTGCATCCGGAGTACTACAAAGACTGCTGACATCGAGTTTTCTACAAGGGACTTTCCG

CTGGGGACTTTCCAGGGAGGTGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCTGCATATAAGCAGCTGCTTTTTGCCTGT

ACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAGCTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCT

TGAGTGCTACAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCT

AGCAGTGGCGCCCGAACAGGGACCTGAAAGCGAAAGTAAAGCCAGAGGAGATCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGG

CAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCGGTA

TTAAGCGGGGGAGAATTAGATAAATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAACAATATAGACTAAAACATATAGTATGGGCA

AGCAGGGAACTAGAACGATTCGCAGTTAATCCTGGCCTTTTAGAGACATCAGAAGGCTGTAGACAAATACTGAGACAGCTACAACCATCC

CTTCAGACAGGATCAGAAGAACGTAGATCATTATTTAATACAGTAGCAGTCCTCTATTGTGTGCATCAAAGGATAGATGTAAAAGACACC

AAGGAAGCCTTAGATAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAGGCACAGCAAGCAGCAGCTGACACAGGAAACAGCAGCCAG

GTCAGCCAAAATTACCCTATAGTGCAGAACCTCCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCATGGGTAAAA

GTAGTAGAAGAGAAGACTTTCAGCCCAGAAGTAATACCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCACAAGATTTAAATACCATG

CTAAACACAGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAGGAGACCATCAATGAGGAAGCTGCAGAATGGGATAGATTGCATCCA

GTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAATTACTAGTACCCTTCAGGAACAAATA

GGATGGATGACACATAATCCACCTATCCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGGACTAAATAAAATAGTAAGAATGTAT

AGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAACTCTAAGAGCCGAG

CAAGCTTCACAGGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCGAACCCAGATTGTAAGACTATTTTAAAAGCATTG

GGACCAGCAGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTGGGGGGACCCGGCCATAAAGCAAGAGTTTTGGCTGAAGCAATG

AGCCAAGTAACAAATCCAGCTACCATAATGATACAGAAAGGCAATTTTAGGAACCAAAGAAAGACTATTAAGTGTTTTAATTGTGGCAAA

GAAGGGCACATAGCCAAAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAAGAAGGACACCAAATGAAAGATTGTACT

GAGAGACAGGCTAATTTTTTAGGGAAAATCTGGCCTTCCCACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCC

CCACCAGAAGAGAGCTTCAGGTTTGGGGAAGAGACAACAACTCCCTCTCAGAAGCAGGAGACGATAGACAAGGAACTGTATCCTTTAGCT

TCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATG

ATACAGTATTAGAAGAAATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATG

ATCAGATACCCATAGAAATCTGCGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGT

TGACTCAGATTGGCTGCACTTTAAATTTTCCCATTAGTCCTATTGAAACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAG

TTAAACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAGTAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAAAAATTG

GGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTCAGAGAACTTA

ATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCACATCCTGCAGGATTAAAACAGAAAAAATCAGTAACAGTACTGGATG

TGGGTGATGCATATTTTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAACAATGAGACACCAG

GGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAGTGTAGCATGACAAAAATCTTGGAGCCTT

TTAGAAAACAAAATCCAGACATAGTCATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAA

AAATAGAGGAACTGAGACAACATCTGTTGAGGTGGGGATTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGG

GTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAGGACAGCTGGACTGTCAATGACATACAGAAATTAG

TGGGAAAATTGAATTGGGCAAGTCAGATTTATGCAGGGATTAAAGTAAGGCAATTATGTAAACTTCTTAGGGGAACCAAAGCACTAACAG

AAGTAGTACCACTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGGGAGATTCTAAAAGAACCGGTACATGGAGTGTATTATGACC

CATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCATTTAAAAATCTGAAAA

CAGGAAAGTATGCAAGAATGAAGGGTGCCCACACTAATGATGTGAAACAATTAACAGAGGCAGTACAAAAAATAGCCACAGAAAGCATAG

TAATATGGGGAAAGACTCCTAAATTTAAATTACCCATACAAAAGGAAACATGGGAAGCGTGGTGGACGGAGTATTGGCAAGCCACCTGGA

TTCCTGAATGGGAGTTTGTCAATACCCCTCCCTTAGTGAAGTTATGGTACCAGTTAGAAAAAGAACCCATAGTAGGAGCAGAAACTTTCT

ATGTAGATGGGGCAGCCAATAGGGAAACTAAATTAGGAAAAGCAGGATATGTAACTGACAGAGGAAGACAAAAAGTTGTCCCCCTAACGG

ACACAACAAATCAGAAGACTGAGTTACAAGCAATTCATCTAGCTTTGCAGGATTCAGGATTAGAAGTAAACATAGTAACAGACTCACAAT

ATGCATTGGGAATCATTCAAGCACAACCAGATAAGAGTGAATCAGAGTTAGTCAGTCAAATAATAGAGCAGTTAATAAAAAAGGAAAAAG

TCTACCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTGGTCAGTGCTGGAATCAGGAAAGTACTAT

TTTTAGATGGAATAGATAAGGCCCAAGAAGAACATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACCTGCCACCTG

TAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGCCCCGGAATAT

GGCAGCTAGATTGTACACATTTAGAAGGAAAAGTTATCTTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGAAGTAATTCCAG

CAGAGACAGGGCAAGAAACAGCATACTTCCTCTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAGTACATACAGACAATGGCAGCAATT

TCACCAGTACTACAGTTAAAGCCGCCTGTTGGTGGGCGGGGATCAAGCAGGAATTTGGCATTCCCTACAATCCCCAAAGTCAAGGAGTAA

TAGAATCTATGAATAAAGAATTAAAGAAAATTATAGGACAGGTAAGAGATCAGGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTAT

TCATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTA

AAGAATTACAAAAACAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAGATCCAGTTTGGAAAGGACCAGCAAAGC

TCCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAAGAAGAAAAGCAAAGATCATCAGGGATT

ATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAACACATGGAAAAGATTAGTAAAACACCATATGTAT

ATTTCAAGGAAAGCTAAGGACTGGTTTTATAGACATCACTATGAAAGTACTAATCCAAAAATAAGTTCAGAAGTACACATCCCACTAGGG

GATGCTAAATTAGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCATTTGGGTCAGGGAGTCTCCATAGAATGGAGG

AAAAAGAGATATAGCACACAAGTAGACCCTGACCTAGCAGACCAACTAATTCATCTGCACTATTTTGATTGTTTTTCAGAATCTGCTATA

AGAAATACCATATTAGGACGTTTAGTTAGTCCTAGGTGTGAATATCAAGCAGGACATAACAAGGTAGGATCTCTACAGTACTTGGCACTA

GCAGCATTAATAAAACCAAAACAGATAAGGCCACCTTTGCCTAGTGTTAGGAAACTGACAGAGGACAGATGGAACAAGCCCCAGAAGACC

AAGGGCCACAGAGGGAGCCATACAATGAATGGACACTAGAGCTTTTAGAGGAACTTAAGAGTGAAGCTGTTAGACATTTTCCTAGGATAT

GGCTCCATAACTTAGGACAACATATCTATGAAACTTATGGGGATACTTGGGCAGGAGTGGAAGCCATAATAAGAATTCTGCAACAACTGC

TGTTTATTCATTTCAGAATTGGATGTCAACATAGCAGAATAGGCATTATTCGACAGAGGAGAGCAAGGAACGGAGCCAGTAGATCCTAGA

CTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAAGTTTGT

TTCACAACAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCCTCAAGACAGTGAGACTCATCAAGTT

TCTCTATCAAAGCAGTGAGTAGTACATGTAATGCTTTCTTTCAATAGTTGTGTGGACCATAGTGTTCATAGAATACAGGAAAGTATTAAG

ACAAAGAAAAGTAGACAGGATAATTGATAGAATAAGAGAAAGAGCAGAAGACAGTGGCAATGAGAG
CGAAGGGGACCAGGAAGAATTATC

AGCACTTGTGGAGATGGGGCACGATGCTCCTTGGGATGTTAATGATCTGTAGTGCTGCAGAACAAT
TGTGGGTCACAGTCTATTATGGGG

TACCTGTGTGGAAAGAAGCAACCACCACTCTATTTTGTGCCTCGGATGCTAAAGCATATGACACAG
AGGTACATAATGTTTGGGCCACAC

ATGCCTGTGTACCCACAGACCCTAACCCACAAGAAGTAGTATTGCAAAATGTGACAGAAAATTTTAA
CATGTGGAAAAATAACATGGTAG

AACAAATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAACCATGTGTAAAATTAACCCC
ACTCTGTGTTACCTTAAATTGCA

CTGATTTGACGAATGCTACTTATGCCAATGGTAGTAGTGAGGAGAGAGGAGAAATAAGAAATTGCT
CTTTCAATGTCACCACAATCATAA

GAAATAAGATACAAAAAGAATATGCACTTTTTTATAGACTTGATATAGAACCAATAGATAAGGATAAT
ACAAGCTATACATTGATAAATT

GTAACACCTCAGTCATTACACAGGCTTGTCCAAAGGTATCCTTTGAACCAATTCCCATACATTATTGT
GCCCCGGCTGGTTTTGCGATTC

TAAAGTGTAACGATAAGAAATTCAATGGAACAGGACCATGTACAAATGTCAGTACAGTACAATGTAC
ACATGGAATTAGGCCAGTAGTGT

CAACTCATCTGCTATTAAATGGCAGTCTAGCAGAAGGAGAGGTAGTAATTAGGTCTGAAAATTTCAC
GAACAATGTTAAAACCATAATAG

TACAGCTGAATGAATCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAGAAAAGGTATAGC
TATAGGACCGGGGAGAACACTCT

ATGCAAGAGAAAAAATAATAGGAGATATAAGACAAGCACATTGTAACCTTAGTAGAGCAAAATGGAA
TGACACTTTAAAACAGATAGTTA

CAAAATTAAAAGAACAATTTAGGAATAAAACAATAGTCTTTAATCAATCCTCAGGAGGGGACCCAGA
AATTGTAATGCACAGTTTTAATT

GTGGAGGGGAATTTTTCTACTGTAAGACAACACAGCTGTTTAATAGTACTTGGCTGTTTAATAGTAC
TTGGAATGATACTGAAAGGTCAG

ATAACAATGAAACTATCATACTCCCATGCAGAATAAAACAAATTATAAACAGGTGGCAGGAAGTAGG
AAAAGCAATATATGCCCCTCCCA

TCAGTGGACAAATTAGATGCTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGGTGATAA
AGAGAACAGTACCACCGAGATCT

TCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAA
AATTGAGCCATTAGGAGTAGCAC

CCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAAAAAAAAGAGCAGTGGGAGCGCTAGGAGCTCTG
TTCCTTGGGTTCTTGGGAGCAGCAG

GAAGCACTATGGGCGCAGCGTCAATGGCGCTGACGGTACAGACCAGACAATTAATGTCTGGTATA
GTGCAACAGCAGAACAATTTGCTGA

AGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCA
AGAGTCCTGGCTGTGGAAAGATACC

TAAAGGATCAACAGCTCCTAGGGATTTGGGGTTGCTCTGGAAAACTCATCTGCACCACTACTGTGC
CTTGGAATGCTAGTTGGAGTAATA

AATCTCTGGATAAGATTTGGGATAACATGACCTGGATGGAGTGGGAGAGAGAAATTGACAATTACA
CAGGCTTAATATACACCTTAATTG

AAGAATCGCAGATCCAACAAGAAAAGAATGAACAAGAATTATTGGAATTGGATAAGTGGGCAAGTTT
GTGGAATTGGTTTGACATAACAA

AGTGGCTGTGGTATATAAAAATATTCATAATGATAGTAGGAGGTTTAATAGGTTTAAGAATAGTTTTT
ACTGTACTGTCTATAGTGAATA

GAGTTAGGCAGGGATACTCACCATTATCGTTCCAGACCCGCCTCCCAGCCCAGAGGGGACCCGAC
AGGCCCGAAGGAATCGAAGAAGAAG

GTGGAGAGAGAGACAGAGACAGATCCGGACCATTAGTGAATGGATTCTTAGCACTCATCTGGGTC
GATCTGCGGAGCCTGTGCCTCTTCA

GCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACTTCTGGGACGCAGGG
GGTGGGAAGCCCTCAAATATTGTT

GGAATCTCCTGCAGTATTGGGGTCAGGAACTAAAGAATAGTGCTATTAGTTTGCTTAATGCCACAGC
TGTAGCAGTAGCTGAGGGGACAG

ATAGGGTTATAGAAGTAGCACAAAGAATTGGTAGGGGTATCCTCCATATCCCTAGAAGAATAAGACA
AGGCTTAGAAATTGCTTTGCTAT

AAGATGGGTGGCAAGTGGTCAAAACGTAGTATGGCTGGATGGTCTGCTGTAAGAGAAAGAATGAAA
CGAGCTGAACCAGCAGCAGATGGG

GTGGGAGCAGTATCTCGAGACTTGGAAAAACATGGAGCAATCACAAGTAGCAATACAGCAGCTACT
AATGCTAATTGTGCCTGGCTAGAA

GCACACGAGGAGGAGGAGGTGGGTTTTCCAGTCAGACCTCAGGTACCTTTAAGACCAATAACTCG
CAAGGCAGCTATGGATCTTAGCCAC

TTTTTAAAAGAAGAGGGGGGACTGGAAGGGCTAATTTGGTCCCAAAGAAGACAAGATATCCTTGAT
CTGTGGATCTACCACACACAAGGC

TACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGG
TGCTTCAAGTTAGTACCAGTTGAG

CCAGGGCAGGTAGAAGAGGCCAATGAAGGAGAGAACAACACCTTGTTACACCCTATGAGCCTGCA
TGGGATGGAGGACCCGGAGGGAGAA

GTATTAGTGTGGAAGTTTGACAGCCTCCTAGCATTTCGTCACATGGCCCGAGAGCTGCATCCGGAG
TACTACAAAGACTGCTGACATCGA

GTTTTCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGTGTGGCCTGGGCGGGACTGGGGA
GTGGCGAGCCCTCAGATGCTGCATA

TAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTC
TGGCTAGCTAGGGAACCCACTGCT

TAAGCCTCAATAAAGCTTGCCTTGAGTGCTACAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGG
TAACTAGAGATCCCTCAGACCCTT

TTAGTCAGTGTGGAAAATCTCTAGCA

>02006_consensus (9737 bp) Consensus sequence of the proviruses from patient 02006

TGGAAGGGATGATTTACTCCCAGAAAAGACAAGAAATCCTTGACCTGTGGGTCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACT

ACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGGTGGTGCTTCAAGTTAGTACCAGTGGATCCAGATAAGGTAGAAGAGGCCA

CTGAAGGAGAAAACAACTGCTTGTTACACCCTATGAACCAGCACGGGATGGATGATCCGGAGAAGGAAGTGTTACAGTGGAAGTTTGACA

GTCGCCTAGCATTTCACCATGTGGCCCGAGAGATACATCCGGAGTACTACAAGAACTGATGACACTGAGCTTTCTACAAGGGACTTTCCA

CTGGGGACTTTCCAGGGGAGGTGTGGCCGGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCTGCATATAAGCAGCTGCTTTTTGCCTG

TACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTGACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCC

TTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCATTTAGTCAGTGTGGAAAATCTC

TAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGTAAAGCCAGAGGAGATCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACA

GCAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAAT

ATTAAGCGGGGGAAAATTAGATGCATGGGAAAAAATTCGGTTAAGGCCAGGAGGAAAGAAACAATATAAGCTAAAACATCTAGTATGGGC

AAGCAGGGAGTTAGAACGATTCGCGGTTAACCCTGGCCTGTTAGAAACATCAGAAGGCTGTAAACAAATAATAGGACAGCTACAACCATC

CTTACAGACAGGATCAGAAGAACTTAAATCATTATTTAATACAGTAGCAGTACTCTATTGTGTACACCAAAGGATAGAGGTAAAAGACAC

CAAGGAAGCTCTAGATAAGATAGAGGAAGAGCAAAATAAAAGTAAGAAAAAGGAACAGCAAGCAGCAGCTGACGCAGGAAGCAGCAGCAA

GGTCAGCCAAAATTACCCTATAGTGCAGAACCTACAGGGACAAATGGTACATCAGCCTATATCACCTAGAACTTTAAATGCATGGGTAAA

AGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTAATACCCATGTTTACAGCATTATCAGAAGGAGCCACCCCACAAGATTTAAACACCAT

GCTAAACACGGTGGGGGGACACCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCTGCAGAATGGGATAGATTGCATCC

AGTGCAGGCAGGGCCTGTTGCACCAGGCCAAATAAGAGAACCAAGGGGAAGTGACATAGCAGGAACTACTAGTACCCTTCAGGAACAAAT

AGCATGGATGACACATAATCCACCTGTCCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGGATTAAATAAGATAGTAAGAATGTA

TAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAGACTCTAAGAGCTGA

GCAAGCTTCACAGGATGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATTCAAACCCAGAT
TGTAAGACTATTTTAAAAGCATT

GGGACCAGGAGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTGGGGGGACCCGGCCAC
AAAGCAAGAGTTTTGGCGGAAGCAAT

GAGCCAAGCAACAGGTGCAGCTACCATAATGATGCAGAAAGGCAATTTTAGGAACCAAAGAAAAGC
TGTTAAGTGTTTCAATTGTGGCAA

AGAAGGGCACATAGCCAGAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGG
AAGGACACCAAATGAAAGATTGTAT

TGAGAGACAGGCTAATTTTTTAGGGAAAATCTGGCCTTCCCACAAGGGAAGGCCAGGGAATTTTCT
TCAGAGCAGACCAGAGCCAACAGC

CCCACCAGCCCCACCAGAAGAGAGCTTCAGGTTTGGGGAAGAGACAACAACTCCCCATCAGAAGC
CGGAGGCGATAGACAAGGAAGTGTA

TCCTTTGGCTTCCCTCAAATCACTCTTTGGCAGCGACCCCTCGTCACAATAAGGATAGGGGGGCAG
CTAAAGGAAGCTCTATTAGATACA

GGAGCAGATGACACAGTATTAGAAGAAATGAATTTGCCAGGAAAATGGAAACCAAAAATGATAGGG
GGAATTGGAGGTTTTATCAAAGTA

AGACAGTATGATCAGATACCCATAGAAATATGTGGACATAAAGCTATAGGTACAGTATTAGTAGGAC
CTACACCTGTCAACATAATTGGA

AGAAATCTGTTGACTCAGCTTGGCTGTACTTTAAATTTTCCCATTAGTCCTATTGAAACTGTACCAGT
AAAATTAAAGCCAGGAATGGAT

GGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAGTAGAAATTTGTACAG
AAATGGAAAAGGAAGGGAAAATT

TCAAAAATTGGGCCTGAAAATCCATACAACACTCCAGTATTTGCCATAAAGAAAAAAGATAGTACTA
AATGGAGAAAATTAGTAGATTTC

AGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCTGCAGGGT
TAAAAAAGAAAAAATCAGTAACA

GTACTAGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATACTGCAT
TTACCATACCTAGTACAAACAAT

GAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATA
TTCCAATGTAGCATGACAAAAATT

TTAGAGCCTTTTAGAAAACAACATCCAGATATAGTTATCTATCAATACATGGATGATTTGTATGTAGG
ATCTGACTTAGAAATAGGACAG

CATAGAACAAAAATAGAAGAACTGAGACAACACTTGTTGAAGTGGGGATTTACTACACCAGACAAAA
AACATCAGAAAGAACCTCCATTC

CTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAA
GACAGCTGGACTGTCAATGACATA

CAGAAGTTAGTGGGAAAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTAT
GTAAACTCCTTAGAGGAACCAAA

GCACTAACAGAAGTAATAACTCTAACAGCAGAAGCAGAGTTAGAACTGGCAGAAAACAGGGAAATT
CTAAAAGAACCAGTACATGGAGTG

TATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGAGCAAGGCCAATGGACATATCAAATTTATCAAGAACCATTTAAA

AACCTGAAAACAGGAAAATATGCAAGAGTGAGGGGTGCCCACACTAATGATGTAAGACAATTAACAGAGGCAGTGCAAAAAATAGCCACA

GAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAAACTACCCATACAAAAAGAAACATGGGAAGCATGGTGGACAGAGTATTGGCAA

GCCACCTGGATTCCTGAATGGGAGTTTGTCAATACCCCTCCCCTAGTGAAGTTATGGTACCAGTTAGAAAAAGAACCCATAGTAGGAGCA

GAAACTTTCTATGTAGATGGGGCAGCTAACAGAGAAACTAAATTAGGAAAAGCAGGATATGTTACTGACAGAGGAAGACAAAAGGTTGTC

CCCCTAACTGACACAACAAATCAGAAGACTGAGTTACAAGCAATTCAGCTAGCTTTGCAAGATTCGGGATTAGAAGTAAACATAGTAACA

GACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGATAAGAGCGAATCAGAGGTAGTCAGTCAAATAATAGAGCAGTTAATAAAA

AAGGAAAAGGTCTACCTGGCCTGGGTGCCAGCACACAAAGGGATAGGAGGAAATGAACAGGTAGATAAATTAGTCAGTTCTGGAATCAGG

AAAATACTATTTTTAGATGGAATAGATAAGGCCCAAGAAGACCATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAAC

CTGCCACCTGTAATAGCAAAAGAAATAATAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGT

CCAGGAATATGGCAACTAGATTGTACACACTTAGAAGGAAAAATTATCCTGGTAGCAGTTCATGTAGCCAGTGGATACATAGAAGCAGAA

GTTATTCCAGCAGAGACAGGGCAAGAAACAGCATACTTTATCTTAAAATTAGCAGGAAGGTGGCCAGTAAAAACAATACATACAGACAAT

GGCAGCAATTTCACCAGTACTACAGTTAAGGCCGCCTGCTGGTGGGCGGGGATCAAGCAGGAATTTGGCATTCCCTACAATCCCCAAAGT

CAAGGAGTAGTAGAATCTATGAATAAAGAATTAAAGAAAATTATAGGGCAAGTAAGAGATCAGGCTGAACATCTTAAGACAGCAGTACAA

ATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGAC

ATACAAACTAAAGAATTACAAAACCAAATTACAAAAATTCAAAAATTTCGGGTTTATTACAGGGACAGCAGAGATCCACTTTGGAAAGGA

CCAGCAAAACTTCTCTGGAAAGGTGAAGGGGCAGTGGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAAGGAGAAAAGCAAAGATC

ATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAGAACATGGAACAGTTTAGTAAAACA

CCATATGTATATTTCAAAGAAAGCTAGGGGATGGTCTTATAAACATCACTATGAAAACCCCCATCCAAGAGTAGGTTCAGAAGTACATAT

CCCACTAGGGGATGCTAAATTGGTAGTAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCATATGGGTCAGGGAGTCTCCAT

AGAATGGAGGAAAAAGGAATATAGCACACAAGTAGACCCTAACACAGCAGACCGACTAATCCATCTGTATTACTTTGATTGTTTTTCAGA

ATCTGCTATAAGAAAGGCCATATTAGGACATAGAGTTAGTCCTAGTTGTGAATATCAAGCAGGACATAACAAGGTAGGATCTCTACAGTA

CCTGGCACTAGTAGCATTAGTAAAACAAAAAAGGACAAAGCCACCTTTGCCTAGTATTACAAAGCTGACAGAGGATAGATGGAACAAGCC

CCAGAAGACCAAGGGCCACAGAGGGAGCCATACCATGAATGGACACTAGAGCTTTTAGAGGAGCTTAAGCATGAAGCTGTCAGACATTTT

CCTAGGCCATGGCTCCATGGCTTAGGGCAACACATCTATGAAACTTATGGGGATACTTGGGCAGGAGTGGAGGCCATAATAAGAATATTG

CAACAACTACTGTTTATTCATTTCAGAATTGGGTGTCGACATAGCAGAATAGGCATTACTCGACAGAGGAGAGCAAGAAATGGAGCCAGT

AGATCCTAGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTG

CCAAGTCTGTTTCACAAAAAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCCTCAAGGCAATCAGAA

TAATCAAGTTCCTCTACCAGAACAGTAAGTAATATATGTAATGCAATCTGAAGTAATATTAGCAATAGTAGGATTAGTAGTAGCAGCAAT

AATAGCAATAGTTGTGTGGACCATAGTATTCATAGAATATAGGAAAATATTAAGACAAAGAAAAATAGACAGGTTAATTGATAGAATAAG

AGATAGAGCAGAAGACAGTGGCAATGAGAGTGAAGGGGACCAGGAAGAATTGTCAGCACTTGTGGAGATGGGGCATCATGCTCCTTGGGA

TGTTGATGATCTGTAGTGCTGCAGACAATTTGTGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAAGAAGCAACCACCACTCTATTTT

GTGCATCAGATGCTAAAGCATATGATGCAGAAGTACATAATGTTTGGGCTACACATGCCTGTGTACCCACAGACCCCAGCCCACAAGAAA

TAAAATTGGAAAATGTGACAGAAAATTTTAATATGTGGAAAAATGGCATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATC

AAAGCCTAAAGCCATGTGTAAAATTAACCCCACTCTGTGTTACTTTACATTGCCAAAAAGTAACGATAGGCAATACCACTTATGTTAATA

ATAGTACAGAAGAAGAAATGAGAAACTGCTCTTTCAATACCACCACAAACATAAGAGATAAGATGCAAAAAGAATATGCACTTTTTTATA

GTCTTGATTTAGTATCACTAGATAATAATAAGCATAATAATAGTAGTAATAGTACCAGATATAGATTGACAAGTTGTAATACCTCAGTTA

TCACACAGGCCTGCCCAAAGGTATCCTTTGAACCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGCTTTTAAAGTGTAATGATA

AAAAGTTTAATGGAACAGGGCCATGCTCAAATGTCAGCACAGTACAATGTACACATGGAATTAAGCCAGTAGTATCAACTCAATTGCTGT

TAAATGGCAGTTTAGCAGAAGAAGAGGTAGTAATTAGATCTAAAAATTTCACTGACAATGCTAAGACCATAATAGTACAGCTGAATAAAC

CTGTAGAAATTAACTGTACAAGACCCAACAACAATACAAGTAAAAGTATAACTATAGGACCAGGGAG
AGCATTTTATGCAACAGAAAGAA

TAATAGGAAATATAAGACAAGCACATTGTAACATTAATGGATCAAAATGGAATAACACTTTAAGACAG
ATAGTTACAAAATTAAGAGAAC

AATTTGGGAATAAAACAATAGTCTTTAATCAATCCTCAGGAGGGGACCCAGAAATAGTAATGCACAT
CTTTAATTGTAGAGGGGAATTTT

TCTACTGTAATACAGCAAAGCTGTTTAATAGTACTTGGCTGTTTAATAGTACTGAAAAGCCAAATAAC
ACCATCAATGGCACCATCACAC

TCCCATGCAGAATAAAACAAATTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCC
CATCCAAGGAAACATTAGTTGTT

CATCAAATATTACAGGGCTGTTATTAACAAGAGATGGTGGTAATAACAATAATGAGAGCACCCCAGA
GATCTTCAGACCTGGAGGAGGAA

ACATGAAGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGATTA
GCACCCACCAGGGCAAAGAGAA

GAGTGGTGCAGAGAGAAAAAAGAGCAGTAACACTAGGAGCTGTGTTCCTTGGGTTCTTGGGAGCA
GCAGGAAGCACTATGGGCGCAGCGT

CACTGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAACAGCAGAACAACCTGC
TGAGGGCTATTGATGCGCAACAGC

ATCTGTTGCAGCTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAGTCCTGGCTGTAGAAAGA
TACCTACGGGATCAACAGCTCCTAG

GGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTACTGTGCCTTGGAATACTAGTTGGAGTAA
TAAATCTCTGGATCACATTTGGG

GTAACATGACCTGGATGGAGTGGGAAAGAGAAATTGACAATTACACAAACTTAATATACAACTTAAT
TGAAGAATCACAGAACCAACAAG

AAAAGAATGAACAAGATTTATTAGAATTAGATAAATGGGCAAGTTTGTGGAGTTGGTTTGACATAAC
AAACTGGCTGTGGTATATAAAAA

TATTCATAATGATAGTAGGAGGCTTGATAGGTTTAAGAATAGTTTTTACTGTAATTTCTATAGTGAAT
AGAGTTAGGCAGGGATACTCAC

CATTATCGTTGCAGACCCGCTTCCCAGCCCCGGGGGGACCCGACAGGCCCGAAGGAATCGAAGG
AGAAGGTGGAGAGAGAGACAGAGAAC

GATCAACTCGATTAGTGCATGGATTCTTAGCACTCATCTGGGACGATCTACGGAGCCTGTGCCTCT
TCAGCTACCACCGCTTGAGAGACT

TACTCTTGATTGCAGCGAGGGTTGTGGAACTTCTGGGACGCAGGGGGTGGGAAGCCCTCAAATAT
TGGGGGAGTCTTCTGCAGTATTGGA

GTCAGGAGCTAAAGAATAGTGCTGTTAGTTTGCTTAATACCACAGCTATAGTAGTAGCAGAAGGAAC
AGATAGGGTTATAGAAGTAGTAC

AAAGAATTTGTAGAGCTATTCGCAACATACCTAGAAGAATAAGACAGGGCTTTGAAAGAGTTTTGCT
ATAAGATGGGTGGCAAGTGGTCA

AAACGTAGTGGGGGTGGATGGTCAGTTGTAAGGGAAAGAATGAGACGAGCTGAGCCAGCATCAGA
GCCAGCAGCAGAGGGGGTGGGAGCA

GCATCTCGAGACTTGGAAAAATATGGAGCAATCACAAGTAGCAATACAGCAGCTACCAATCCTGATT
GTGCCTGGCTAGAAGCACAAGAG

GAGGAGGAGGTGGGTTTTCCAGTCAGACCTCAGGTACCTTTAAGACCAATGACTTACAAGGGAGCT
CTAGATCTTAGCCACTTTTTAAAA

GAAAAGGGGGGACTGGAAGGGATGATTCACTCCCAGAAAAGACAAGAAATCCTTGACCTGTGGGT
CTACCACACACAAGGCTACTTCCCT

GATTGGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGGTGGTGCTTCAA
GTTAGTACCAGTGGATCCAGATAAG

GTAGAAGAGGCCACTGAAGGAGAAAACAACTGCTTGTTACACCCTATGAACCAGCATGGGATGGAT
GATCCGGAGAAAGAAGTGTTACAG

TGGAAGTTTGACAGTCGCCTAGCATTTCACCATGTGGCCCGAGAGCTACATCCGGAGTACTACAAG
AACTGATGACACTGAGCTTTCTAC

AAGGGACTTTCCACTGGGGACTTTCCAGGGGAGGTGTGGCCAGGGCGGGACTGGGGAGTGGCGA
GCCCTCAGATGCTGCATATAAGCAGC

TGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTGAC
TAGGGAACCCACTGCTTAAGCCTC

AATAAAGCTTGCCTTGAGTGCTTTAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAG
ATCCCTCAGACCCTTTAGTCAGT

GTGGAAAATCTCTAGCA

>02006_chr10_sc (9748 bp) Sequence of the provirus with ID chr10:41845251-42110306

TGGAAGGGATGATTTACTCCCAGAAAAGACAAGAAATCCTTGACTTGTGGGTCTACCACACACAAG
GCTACTTCCCTGATTGGCAGAACT

ACACACCAGGGCCAGGGACCAGATATCCACTGACCTTTGGGTGGTGCTTCAAGTTAGTACCAGTG
GATCCAGATAAAGTAGAAGAGGCCA

CTGAAGGAGAAAACAACTGCTTGTTACACCCTCTGAACCAGCACGGGATGGATGATCCAGAGAAG
GAAGTGTTAATATGGCAGTTTGACA

GTCACCTAGCATTTCACCATGTAGCTCGAGAGAAACATCCGGAGTACTACAAGAACTGATGACACT
GAGCTTTCTACAAGGGACTTTCCA

CTGGGGACTTTCCAGGGGAGGTGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCT
GCATATAAGCAGCTGCTTTTCGCCTG

TACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTGACTAGGGAACCCACT
GCTTAAGCCTCAATAAAGCTTGCC

TTGAGTGTTTTAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCA
TTTAGTCAGTGTGGAAAATCTCT

AGCAGTGGCGCCCGAACAGGGACTTGAAAACGAAAGTAAGGCCAGAGGAGATCTCTCGACGCAG
GACTCGGCTTGCTGAAGCGCGCACAG

CAAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCATTTTTGACTAGCGGAGGCTAGAAGGAGAG
AGATGGGTGCGAGAGCGTCAATATTA

AGCGGGGGAAAATTAGATGCATGGGAAAAAATTCGGTTAAGGCCAGGAGGAAAGAAACAATATAAG
CTAAAACATTTAGTATGGGCAAGC

AGGGAGTTAGAACGATTCGCGGTTAACCCTGGCCTGTTAGAAACATCTGAAGGCTGTAAACAAATA
ATAGCACAGCTACAACCATCCTTA

AAGACAGGATCAGAAGAACTTAAATCATTATTTAATACAGTAGCAGTACTCTATTGTGTACACCAGA
GGATAGAGGTAAGAGACACCAAG

GAAGCTTTAGATAGGATAGAGGAAGAGCAAAATAAAAGTAAGAAAAAGGAACAGCAAGCAGCAGCT
GACGCAGGAAGCAGCAGCAAGGTC

AGCCAAAATTACCCTATAGTGCAAAATCTACAGGGACAAATGGTACATCAGCCTATATCACCTAGAA
CTTTAAATGCATGGGTGAAGGTA

GTAGAAGAGAAGGCTTTCAGCCCAGAAGTAATACCCATGTTTACAGCATTATCAGAAGGAGCCACC
CCACAAGATTTAAACACCATGCTA

AACACGGTGGGGGGACACCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCTGC
AGAATGGGATAGATTGCATCCAGTA

CAGGCAGGGCCTGTTGCACCAGGCCAGATAAGAGAACCAAGGGGAAGTGACATAGCAGGAACTAC
TAGTACCCTTCAGGAACAAATAGCA

TGGATGACACATAATCCACCTGTCCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGGATTAA
ATAAGATAGTAAGAATGTATAGC

CCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTC
TATAAGACTCTAAGAGCTGAGCAA

GCTTCACAGGATGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATTCAAACCCAGATTGTA
AGACTATTTTAAAAGCATTGGGA

CCAGGAGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGGGGACCCGGCCACAAAGC
AAGAATTTTGGCGGAAGCAATGAGC

CAAGCAACAGGTGCAGCTACCATAATGATGCAGAAAGGCAATTTTAGGAACCAAAGAAAAACTGTT
AAGTGTTTCAATTGTGGCAAAGAG

GGGCACATAGCCAGAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGGAAGG
ACATCAAATGAAAGATTGTATTGAG

AGACAGGCTAATTTTTTAGGGAAAATCTGGCCTTCCCACAAGGGAAGGCCAGGGAATTTTCTTCAG
AGCAGACCAGAGCCAACAGCCCCA

CCAGCCCCACCAGAAGAAAGCTTCAGGTTTGGGGAAGACACAACAACTCCCCATCAGAAGCCGGA
GGTGATAGACAAGGAAATGTATCCT

TTGACTTCCCTCAAATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAGGGGGGCAGCTAA
AGGAAGCTCTATTAGATACAGGAG

CAGATGACACAGTATTAGAAGAAATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAA
TTGGAGGTTTTATCAAAGTAAGAC

AGTATGATCAGGTGCCACTAGAAATATGTGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTAC
ACCTGTCAACATAATTGGAAGAA

ATCTGTTGACTCAGCTTGGCTGTACTTTAAATTTTCCCATTAGTCCTATTGAAACTGTACCAGTAAAA
TTAAAGCCAGGAATGGATGGCC

CAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAGTAGAAATTTGTACAGAAATGGAAAAGGAAGGAAAAATTTCAA

AAATTGGGCCTGAAAATCCATACAACACTCCAGTATTTGCTATAAAGAAAAAAGATAGTACTAAATGGAGAAAATTAGTAGATTTCAGAG

AACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAGTTAGGAATACCACATCCTGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTAC

TAGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATACTGCATTTACCATACCCAGTACAAACAATGAGA

CACCAGGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAATGTAGCATGACAAAAATTTTAG

AGCCTTTTAGAAAACAACACCCAGATATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGACAGCATA

GAACAAAAATAGAAGAACTGAGACAACACTTGTTAAAATGGGGATTTACTACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTT

GGATGGGTTATGAACTCCATCCTGATAAGTGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGA

AGTTAGTAGGAAAATTAAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGACAATTATGTAAACTCCTTAGAGGAACCAAAGCAC

TAACAGAAGTAATAACTCTAACAGAAGAAGCAGAGTTAGAACTGGCAGAAAACAGGGAAATTCTAAAAGAACCAGTACATGGAGTGTATT

ATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGAGCAAGGCCAGTGGACATATCAAATTTATCAAGAGCCATTTAAAAACC

TGAAAACAGGAAAATATGCAAGAGTGAGGGGTGCCCACACTAATGATGTAAGACAATTAACAGAGGCAGTGCAAAAAATAGCCACAGAAA

GCATAGTGATATGGGGAAAGACTCCTAAATTTAAACTGCCCATACAAAAAGAAACATGGGAAGCATGGTGGACAGAGTATTGGCAAGCCA

CCTGGATTCCTGAATGGGAGTTTGTCAATACCCCTCCCCTAGTGAAGTTATGGTACCAGTTAGAAAAAGAACCCATAGTAGGAGCAGAAA

CTTTCTATGTAGATGGGGCAGCTAACAGAGAAACTAAACTAGGAAAAGCAGGATATGTAACTGACAGAGGAAGACAAAAGGTTGTCCCCC

TAACTGACACAACAAATCAGAAGACTGAGTTACAAGCAATTCAGCTAGCTTTGCAAGATTCGGGATCAGAAGTAAACATAGTAACAGACT

CACAATATGCATTAGGAATCATTCAAGCACAACCAGATAAGAGCGAATCAGAGGTAGTCAGTCAAATAATAGAGCAGCTAATAAACAAGG

AAAAGGTCTACCTGGCATGGGTGCCAGCACACAAAGGAATAGGAGGAAATGGACAGGTAGATAAGTTAGTCAGTTCTGGAATCAGGAAGG

TACTATTCTTAGATGGAATAGATAAGGCCCAAGAAGACCATGAGAAATATCACAGTAATTGGAGAGCAATGGCTAGTGATTTTAACCTGC

CACCTGTAATAGCAAAAGAAATAATAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAG

GAATATGGCAGTTAGATTGTACACACTTAGAAGGAAAAATTATCCTGGTAGCAGTTCATGTAGCCAGTGGATACATAGAAGCAGAAGTTA

TTCCAGCAGAGACAGGACAAGAAACAGCATACTTTATCTTAAAATTAGCAGGAAGGTGGCCAGTAAAAACAATACATACAGACAATGGTA

GCAATTTCACCAGTACTACAGTTAAGGCCGCCTGCTGGTGGGCAGGGATCAAGCAGGAATTTGGCATTCCCTACAATCCCCAAAGTCAAG

GAGTAGTGGAATCTATGAATAAAGAATTAAAGAAAATTATAGGGCAAGTAAGAGATCAGGCTGAACATCTTAAGACAGCAGTACAAATGG

CAGTATTCATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATAC

AAACTAAAGAACTACAAAACCAAATTACAAAAATTCAAAAATTTCGGGTTTATTACAGGGACAGCAGAGATCCACTTTGGAAAGGACCAG

CAAAACTTCTGTGGAAAGGTGAAGGGGCAGTGGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAAGGAGAAAAGCAAAGATCATTA

GGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAAAACATGGCACAGTTTAGTAAAACACCAT

ATGTATATTTCAAAGAAAGCTAAGGGATGGTCTTATAAACATCACTATGAAAACCCCCATCCAAGAGTAGGTTCAGAAGTACATATCCCA

CTAGGGGATGCTAAATTGGTAGTAATAACATATTGGGGTCTGCATACAGGAGAGAGAGACTGGCATATGGGTCAGGGAGTCTCCATAGAA

TGGAGGAAAAAGGAATATAGCACACAAGTAGACCCTAATACAGCAGACCAACTAATCCATCTGTATTACTTTGATTGTTTTTCAGAATCT

GCTATAAGAAGGGCCATATTAGGACATAGAGTTAGTCCTAGGTGTGAATATCAAGCAGGACATAACAAGGTAGGATCTCTACAGTACCTG

GCACTAGTAGCATTAGTAAAACAAAAAAGGACAAAGCCACCTTTGCCTAGTATTACAAAGCTGACAGAGGATAAATGGAACAAGCCCCAG

AAGACCAAGGGCCACAGAGGGAGCCATACAATGAATGGACACTAGAGCTTTTAGAGGAGCTTAAACAGGAAGCTGTCAGACATTTTCCTA

GGCCATGGCTTCATGGCTTAGGGCAACACATCTATGAAACTTATGGAGATACTTGGGCAGGAGTGGAGGCCATAACAAGAATATTGCAAC

AACTACTGTTTATTCATTTCAGAATTGGGTGTCGACATAGCAGAATAGGCATTATTCGACAGAGGAGAGCAAGAAATGGAGCCAGTAGAT

CCTAGACTAGAGCCCTGGAAACATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAA

GTCTGTTTCACAAAAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCCTCAAAGCAATCAGAATAAT

CAAGTTCCTCTATCAGAACAGTAAGTAATATATGTAATGCATTCTGAAGTAATATTAACAATAGTAGCATTAGTAGTAGCAGCAATAATA

GCAATAGTTGTGTGGACCATAGTACTCATAGAATATAGGAAAATATTAAGACAAAGAAAAATAGACAGGTTAATTGATAGAATAATAGTT

AGAGCAGAAGACAGTGGCAATGAGAGTGAAGGGGACCAGGAAGAATTGTCAGCACTTGTGGACATGGGGCATGATGCTCCTTGGGTTGTT

AATGATCTGTAGTGCTGCAGACAATTTGTGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAAGAAGCAACCACCACTCTATTTTGTGC

ATCAGATGCTAAAGCATATGATGCAGAGGTACATAATGTTTGGGCTACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAATAAA

ATTGGAAAATGTGACAGAAAATTTTAATATGTGGAGAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAG

CCTAAAGCCATGTGTAAAATTAACCCCACTCTGTGTTACTATGGATTGCCATCATGAAGTAAAGGTAGAAAATGCTACTGGATTACAGAA

TACCACTTTTATTAATAATAAAACAGAAGAAATGAGAAACTGCTCTTTCAATACCACCACAAGCATAAGAGATAAGATGCAGAAAGAATA

TGCAATTTTTTATAAACTTGATTTAGTACCACTAAATAATAATAAGGATAATAGTAGTAGTAATAGTAGCTATAGATTGATAAATTGTAA

TACCTCAGTTATCACACAGGCCTGTCCAAAGATAACCTTTGAACCAATTCCCATACATTATTGTACCCCGGCTGGTTTTGCACTTTTAAA

GTGTAATGATAAACAGTTTAATGGAACAGGGCCATGCTCAAATGTCAGCACAGTACAATGTACACATGGAATTAAGCCAGTAGTATCAAC

TCAATTGCTGTTAAATGGCAGTTTAGCAGAAGAAGAGGTAGTAATTAGATCTGAAAATTTCACTGACAATGCTAAGACCATAATAGTACA

GCTAAAAGAACCTGTAGAAATTAACTGTACAAGACCTGGCAACAATACAAGTAAAAGTATATCTATAGGACCAGGGAGAGCATTTTATGC

AACAGGAAGAATAATAGGAGATATAAGACAAGCACATTGTAACCTTAGTAGATCACAATGGAATAACACTTTAAAAGATATAGTTGCAAA

ATTAAGAGAACAATTTGGGAATAAAACAATAGTCTTTAATCGATCCTCAGGAGGGGACCCAGAAATAGTAATGCACATCTTTAATTGTAG

AGGGGAATTTTTCTACTGTAATACAACAAAGCTGTTTAATAGTACTTGGCTGTCTAATAGTACTTTGAATTACACTGACCAGTCAAATGA

CACCATCACACTCCCATGCAAAATAAAACAAATTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATCCAAGGAAA

CATTAGTTGTTCATCAAAAATTACAGGGCTGTTATTAACAAGAGATGGTGGTAATAACAATAATGAGAGCGCCCCAGAGATCTTCAGACC

TGGAGGAGGAAACATGAAGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAG

GGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTAACACTAGGAGCTGTGTTCCTTGGGTTCTTGGGCGCAGCAGGAAGCACTAT

GGGCGCAGCGTCACTGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAACAGCAGAACAATCTGCTGAGGGCTATTGA

TGCGCAACAGCATCTGTTGCAGCTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAGTCCTGGCTGTAGAAAGATACCTACGGGATCA

ACAGCTCCTAGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTACTGTGCCTTGGAATACTAGTTGGAGTAATAAATCTCTGGA

TCACATTTGGGGTAACATGACCTGGATGGAGTGGGAAAGAGAAATTGACAATTACACAAGCTTAATATATAACTTAATTGAAGAATCACA

GAACCAACAAGAAAAGAATGAAAAAGATTTATTAGAATTAGATAAATGGGACAGTTTGTGGAGTTGGTTTGACATAACAAACTGGCTGTG

GTATATAAGAATATTCATAATGATAGTAGGAGGCTTGATAGGTTTAAAAATAGTTTTTACTGTAATTTCTATAATAAATAGAGTTAGGCA

GGGATACTCACCATTATCGTTGCAGACCCGCTTCCCAGCCCCGGGGGGACCCGACAGGCCCGAAGGAATCGAAGGAGAAGGTGGAGAGAG

AGACAGAGAACGATCAACTCGATTAGTGCATGGATTCTTCGCACTCATCTGGGACGATCTACGGAGCCTGTGCCTCTTCAGCTACCACCG

CTTGAGAGACTTACTCTTGATTGCAGCGAGGGTTGTGGAACTTCTGGGACGCAGGGGGTGGGAAGCTCTCAAATATTGGGGGAGTCTTCT

GCAGTATTGGAGTCAGGAGCTAAAGAATAGTGCTGTTAGTTTGCTTAATACCACAGCTATAGTAGTAGCAGAAGGAACAGATAGGATTAT

AGAAGTAGCACAAAGAATTTGTAGAGCTATTCGCCATATACCTAGAAGAATAAGACAGGGCTTTGAGAGAGCTTTGCTATAAGATGGGTG

GCAAGTGGTCAAAACGTAGTAGGGAGGGATGGTCAGTTATAAGGGAAAGAATGAGACGAGCTGAGCCAGCAGCAGAGCCAGCAGCAGAGG

GGGTGGGAGCAGCATCTCGAGACTTGGAAAAATATGGAGCAATCACAAGTAGCAATACAGCAGCTACCAATGCTGATTGTGCCTGGCTGG

AAGCACAAGAGGAGGAGGAGGTGGGTTTTCCAGTCAGACCTCAGGTACCTTTAAGACCAATGAATTACAAGGGAGCTGTAGATCTTAGCC

ACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGATGATTTACTCCCAGAAAAGACAAGAAATCCTTGACTTGTGGGTCTACCACACACAAG

GCTACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGACCAGATATCCACTGACCTTTGGGTGGTGCTTCAAGTTAGTACCAGTGG

ATCCAGATAAAGTAGAAGAGGCCACTGAAGGAGAAAACAACTGCTTGTTACACCCTCTGAACCAGCACGGGATGGATGATCCAGAGAAGG

AAGTGTTAATATGGCAGTTTGACAGTCACCTAGCATTTCACCATGTAGCTCGAGAGAAACATCCGGAGTACTACAAGAACTGATGACACT

GAGCTTTCTACAAGGGACTTTCCACTGGGGACTTTCCAGGGGAGGTGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCTGC

ATATAAGCAGCTGCTTTTCGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTGACTAGGGAACCCACT

GCTTAAGCCTCAATAAAGCTTGCCTTGAGTGTTTTAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACC

ATTTAGTCAGTGTGGAAAATCTCTAGCA

>02006_cen 9760 bp Sequence of the provirus with ID chr13:16283554-17732159

TGGAAGGGTTGATTTATTCCCAGCAAAGACAAGAAATCCTTGACTTGTGGGTCTACCACACACAAG
GCTACTTCCCTGATTGGCAGAACT

ACACACCAGGGCCAGGGACCAGATATCCACTGACCTTTGGGTGGTGCTTCAAGTTAGTACCAGTG
GATCCAGATAAAGTAGAAGAGGCCA

CTGAAGGAGAAAACAACTGCTTGTTACACCCTCTGAACCAGCACGGGATGGATGATCCAGAGAAG
GAAGTGTTAATGTGGAAGTTTGACA

GTCGCCTAGCATTTCACCATGTGGCCCGAGAGCTACATCCGAGTACTACAAGAACTGATGACACTG
TGCTTTCTACAAGGGACTTTCCAC

TGGGGACTTTCCAGGGGAGGTGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCTG
CATATAAGCAGCTGCTTTTTGCCTGT

ACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTGACTAGGGAACCCACT
GCTTAAGCCTCAATAAAGCTTGCCT

TGAGTGCTTTAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCA
TTTAGTCAGTGTGGAAAATCTCTA

GCAGTGGCGCCTGAACAGGGACTTGAAAGCGAAAGTAAGGCCAGAGGAGATCTCTCGACGCAGG
ACTCGGCTTGCTGAAGCGCGCACGGC

AAGAGGCGAGGGGCGGCGACTGGTGAGTACGCCATTTTTTGACTAGCGGAGGCTAGAAGGAGAG
AGATGGGTGCGAGAGCGTCAATATTA

AGCGGGGGAAAATTAGATGCATGGGAAAAAATTCGGTTAAGGCCAGGAGGAAAGAAACAATATAAG
CTAAAACATCTAGTATGGGCAAGC

AGGGAGTTAGAACGATTCGCGGTTAACCCTGGCCTGTTAGAAACATCAGAAGGCTGTAAACAAATA
ATAGGACAGCTACAACCATCCTTA

AAGACAGGATCAGAAGAACTTAAATCATTATTTAATACAGTAGCAGTACTCTATTGTGTACACCAGA
GGATAGAGGTAAGAGACACCAAG

GAAGCTCTGGATAAGATAGAGGAAGAGCAAAATAAAAGTAAGAAAAAGGAACAGCAAGCAGCAGCT
GACGCAGGAAGCAGCAGCAAGGTC

AGCCAAAATTACCCTATAGTGCAGAACTTACAGGGACAAATGGTACATCAGCCTATATCACCTAGAA
CTTTAAATGCATGGGTGAAAGTA

GTAGAAGAGAAGGCTTTCAGCCCAGAAGTAATACCCATGTTTACAGCATTATCAGAAGGAGCCACC
CCACAAGATTTAAACACCATGCTA

AACACGGTGGGGGGACACCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCTGC
AGAATGGGATAGATTGCATCCAGTG

CAGGCAGGGCCTGTCGCACCAGGCCAAATAAGAGAACCAAGGGGAAGTGACATAGCAGGAACTAC
TAGTACCCTTCAGGAACAAATAGCA

TGGATGACACATAATCCACCTGTCCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGGATTAA
ATAAAATAGTAAGAATGTATAGC

CCTACCAGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTC
TATAAGACTCTAAGAGCTGAGCAA

GCTTCACAGGATGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATTCAAACCCAGATTGTA
AGACTATTTTAAAAGCATTGGGA

CCAGGAGCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTGGGGGGACCCGGCCACAAAG
CAAGAATTTTGGCGGAAGCAATGAGC

CAAGCAACAGGTGCAGCTACCATAATGATGCAGAAAGGCAATTTTAGGAACCAAAGGAAAACTGTT
AAGTGTTTCAATTGTGGCAAAGAA

GGGCATATAGCCAAAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGGAAGG
ACACCAAATGAAAGATTGTATTGAG

AGACAGGCTAATTTTTTAGGGAAAATCTGGCCTTCCCACAAGGGAAGGCCAGGGAATTTTCTTCAG
AGCAGACCAGAGCCAACAGCCCCA

CCAGCCCCACCAGAAGAGAGCTTCAGGTTTGGGGAAGAGACAACAACTCCCCATCAGAAGCCGGA
GGTGATAGACAAGGAAATGTATCCT

TTGGCTTCCCTCAAATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAGGGGGGCAGCTAA
AGGAAGCTCTATTAGATACAGGAG

CAGATGACACAGTATTAGAAGAAATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAA
TTGGAGGTTTTATCAAAGTAAGAC

AGTATGATCAGGTACACATAGAAATATGTGGACATAAAGCTATAGGTACAGTATTAATAGGACCTAC
ACCTGTCAACATAATTGGAAGAA

ATCTGTTGACTCAGCTTGGCTGTACTTTAAATTTTCCCATTAGTCCTATTGAAACTGTACCAGTAAAA
TTAAAACCAGGAATGGATGGCC

CAAAAGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAATAGAAATTTGTACAGAAATG
GAAAAGGAAGGGAAAATTTCAA

AAATTGGGCCTGAAAATCCATACAACACTCCAGTATTTGCCATAAAGAAAAAGGATAGTACTAAATG
GAGAAAATTAGTAGATTTCAGAG

AACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCACACCCTGCAGGGTTAAA
AAAGAAAAAATCAGTAACAGTAC

TAGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATACTGCATTTACC
ATACCTAGTACAAACAATGAGA

CACCAGGGATTAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCACCAGCAATATTCC
AATGTAGTATGACAAAAATTTTAG

AGCCTTTTAGAAAACAACATCCAGATATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCT
GACTTAGAAATAGGACAGCATA

GAACAAAAATAGAAGAACTGAGACAACACTTGTTGAAGTGGGGATTTACTACACCAGACAAAAAACA
TCAGAAAGAACCTCCATTCCTTT

GGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACA
GCTGGACTGTCAATGACATACAGA

AATTAGTGGGAAAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAA
ACTCCTTAGGGGAACCAAAGCAC

TAACAGAAGTAATAACTCTAACAGCAGAAGCAGAGTTAGAACTGGCAGAAAACAGGGAAATTCTAAA
AGAACCAGTACATGGAGTGTATT

ATGACCCCTCAAAAGACTTAATAGCAGAAATACAGAAGCAGGATCAAGGCCAATGGACATATCAAAT
TTATCAAGAGCCATTTAAAAACC

TGAAAACAGGAAAATATGCAAAAATGAGGGGTGCCCACACTAACGATGTAAGACAATTAACAGAGG
CAGTGCAAAAAATAGCCACAGAAA

GCATAGTAATATGGGGAAAGACTCCTAAATTTAAACTACCCATACAAAAAGAAACATGGGAAGCATG
GTGGACAGAGTATTGGCAAGCCA

CCTGGATTCCTGAATGGGAGTTTGTCAATACCCCTCCCCTAGTGAAGTTATGGTACCAGTTAGAAAA
AGAACCCATAGTAGGAGCAGAAA

CTTTCTATGTAGATGGGGCAGCTAACAGGGAAACTAAACTAGGAAAAGCAGGATATGTTACTGACA
GAGGAAGACAAAAGGTTGTCCCCC

TAACTGACACAACAAATCAGAAGACTGAGTTACAAGCAATTCAGCTAGCTTTGCAAGATTCGGGATC
AGAAGTAAACATAGTAACAGACT

CACAATATGCATTGGGAATCATTCAAGCACAACCAGATAAGAGCGAATCAGAGGTAGTCAGTCAAA
TAATAGAGCAGCTAATAAACAAGG

AAAAGGTCTACCTGGCCTGGGTGCCAGCACATAAAGGGATAGGAGGAAATGAACAGGTAGATAAAT
TAGTCAGTTCTGGAATCAGGAAAA

TACTATTTTTAGATGGAATAGATAAAGCCCAAGAAGATCATGAGAAATATCACAGTAATTGGAGAGC
AATGGCTAGTGATTTTAACCTGC

CACCGGTAATAGCAAAAGAAATAATAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGC
ATGGACAAGTAGACTGTAGTCCAG

GAATATGGCAACTAGATTGTACACATTTAGAAGGAAAAATTATCCTGGTAGCAGTTCATGTAGCCAG
TGGATACATAGAAGCAGAAGTTA

TTCCAGCAGAGACAGGGCAAGAAACAGCATACTTTATCTTAAAATTAGCAGGAAGGTGGCCAGTAA
AAACAATACATACAGACAATGGCA

GCAATTTCACCAGTACTACAGTTAAGGCCGCCTGCTGGTGGGCGGGGATCAAGCAGGAATTTGGC
ATTCCCTACAATCCCCAAAGTCAAG

GAGTAGTGGAATCCATGAATAAAGAATTAAAGAAAATTATAGGGCAAGTAAGAGATCAGGCTGAACA
TCTTAAGACAGCAGTACAAATGG

CAGTATTCATCCACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAG
TAGACATAATAGCAACAGACATAC

AAACTAAAGAACTACAAAATCAAATTACAAAAATTCAAAAATTTCGGGTTTATTACAGGGACAGCAGA
GATCCACTTTGGAAAGGACCAG

CAAAACTTCTGTGGAAAGGTGAAGGGGCAGTGGTAATACAAGATAATAGTGACATAAAAGTAGTGC
CAAGGAGAAAAGCAAAGATCATTA

GGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAAGATTAGAACA
TGGCACAGTTTAGTAAAACATCAT

ATATATATTTCAAAGAAAGCTAGGGGGTGGTCTTATAAACATCACTATGAAAACCCCCATCCAAGAG
TAGGTTCAGAAGTACATATCCCA

CTAGGGGATGCTAAATTGGTAGTAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCAT
ATGGGTCAGGGAGTCTCCATAGAA

TGGAGGAAAAAGGAATATAGCACACAAGTAGACCCTAACACAGCAGACCGACTAATCCATCTGTATTACTTTGATTGTTTTTCAGAATCT

GCTATAAGAAAGGCCATATTAGGACATAGAGTTAGTCCTAGGTGTGAATATCAAGCAGGACATAACAAGGTAGGATCTCTACAGTACCTG

GCACTAGTAGCATTAGTAAAACAAAAAAGAACAAAGCCACCTTTGCCTAGTATTTCAAAGCTGACAGAGGATAGATGGAACAAGCCCCAG

AAGATCAAGGGCCACAGAGGGAGCCATACAATGAATGGACACTAGAGCTTTTAGAGGAGCTTAAGCAAGAAGCTGTCAGACATTTTCCTA

GGCCATGGCTTCATGGCTTAGGGCAACACATCTATGAAACCTATGGAGATACTTGGGCAGGAGTGGAGGCCATAACAAGAATATTGCAAC

AACTACTGTTTATTCATTTCAGAATTGGGTGTCGACATAGCAGAATAGGCATTACTCGACAGAGGAGAGCAAGAAATGGAGCCAGTAGAT

CCTAGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGTACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAA

GTCTGTTTCACAAAAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCCTCAAGGCAATCAGAATAAT

CAAGTTCCTCTATCAGAACAGTAAGTAATATATGTAATGCATTCTGAAGTAATATTAACAATAGTAGCATTAGTAGTAGCAGCAATAATA

GCAATAGTTGTGTGGACCATAGTATTAATAGAATATAGGAGAATATTAAGACAAAGAAAAATAGACAGGTTAATTGATAGAATAATAGAT

AGAGCAGAAGACAGTGGCAATGAGAGTGAAGGGGACCAGGAAGAATTGTCAGCACTTGTGGAGATGGGGCATGATGCTCCTTGGGTTGTT

AATGATCTGTAGTGCTGCAGACAATTTGTGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAAGAAGCAACCACCACTCTATTTTGTGC

ATCAGATGCTAAAGCATATGATGCAGAGGTACATAATGTTTGGGCTACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAATAAA

ATTGAAAAATGTGACAGAAAATTTTAATATGTGGAGAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAG

CCTAAAGCCATGTGTGAAATTAACCCCACTCTGTGTTACTTTAGATTGCCATTATGAAGTAAAGATAGAAAATGCCACTGGATTTGGGAA

TACCACTTTTATTAATAATAAAACAGAAGAAATGAGAAACTGCTCTTTCAATACCACCACAAGCATAAGAGATAAGATGCAGAAAGAATA

TGCAATTTTTTATAAACTTGATTTAGTACAACTAGATAGTAATAAGCATAATAGTAGTAATAATACCAGATATAGATTGATAAATTGTAA

TACCTCAGTTATCACACAAGCCTGTCCAAAGATAACCTTTGAACCAATTCCCATACATTATTGTACCCCGGCTGGTTTTGCGCTTCTAAA

GTGTAATGATAAAAAGTTTAATGGAACAGGGCCTTGCTCAAATGTCAGCACAGTACAATGCACACATGGAATTAAGCCAGTAGTATCAAC

TCAATTGCTGTTAAATGGCAGTTTAGCAGAAGAAGAGGTAGTAATTAGATCTGAAAATTTCACTGACAATGCTAAGACCATAATAGTACA

GCTAAAAGAACCTGTAGAAATTAACTGTACAAGACCCGGCAACAATACAAGTAAAAGTATAACTATAGGACCAGGGAGAGCATTTTATGC

AACAGGGAGAATAATAGGAGATATAAGACAAGCACATTGTAACCTTAGTAGATCACAATGGAATAACACTTTAAAAGATATAGTTAAAAA

ATTAAGAGAACAATTTGGGAATAAAACAATAGCCTTTAAAAATTCCTCAGGAGGGGACCCAGAAATAGTAATGCACATCTTTAATTGTAG

AGGGGAATTTTTCTACTGTAATACAACAAAGCTGTTTAGTAGTATTTGGCTGTCTAATAGTACTTTGAATTACACTAAAAAGTCAAATAG

CACCATCAATGACACCATCACACTCCCATGCAGAATAAAACAAATCATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCC

CATCCAAGGAAACATTAGTTGTTCATCAAAAATTACAGGGCTGTTATTAACAAGAGATGGTGGTAATAACAATAATGAGAGCGCCCCAGA

AATCTTCAGACCTGGAGGAGGAAACATGAAGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGT

AGCACCCACCAGGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTAACACTAGGAGCTGTGTTCCTTGGGTTCTTGGGAGCAGC

AGGAAGCACTATGGGCGCAGCGTCATTGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAACAGCAGAACAACCTGCT

GAGGGCTATTGATGCGCAACAGCATCTGTTGCAGCTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAGTCCTGGCTGTAGAAAGATA

CCTAAGAGATCAACAGCTCCTAGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTACTGTGCCCTGGAATACTAGTTGGAGTAA

TAAATCTCTGGATCACATTTGGGGTAACATGACCTGGATGGAGTGGGAAAGAGAAATTGACAATTACACAAGCTTAATATACACCTTAAT

TGAAGAATCACAGAACCAACAAGAAAAGAATGAACAAGATTTATTAGAATTAGATAAATGGGACAGTTTGTGGAGTTGGTTTGACATAAC

AAACTGGCTGTGGTATATAAGAATATTCATAATGATAGTAGGAGGCTTGATAGGTTTAAGAATAGTGTTTACTGTAATTTCTATAGTAAA

TAGAGTTAGGCAGGGATACTCACCATTATCATTGCAGACCCGCTTCCCAGCCCCGGGGGGACTCGACAGGCCCGAAGGAATCGAAGGAGA

AGGTGGAGAGAGAGACAGAGAACGATCAACTCGATTAGTGCATGGATTCTTCGCACTCATCTGGGACGATCTACGGAGCCTGTGCCTCTT

CAGCTACCACCGCTTGAGAGACTTACTCTTGATTGCAGCGAGGGTTGTGGAACTTCTGGGACGCAGGGGGTGGGAAGCCCTCAAGTATTG

GGGGAGTCTTCTGCAGTATTGGAGTCAGGAGCTAAAGAATAGTGCTGTTAGTTTGCTTAATACCACAGCTATAGTAGTAGCAGAAGGAAC

AGATAGGATTATAGAAGCAATACAAAGAATTTGCAGAGCTATTCGCAACGTACCTAGAAGAATAAGACAGGGCTTTGAAAGAGCTTTGCT

ATAAGATGGGTGGCAAGTGGTCAAAACGTGTTGGGGATGGATGGTCAGTTGTAAGGGAAAGAATGAGACGAGCTGAGCCAGTAGCAGAAC

CAGCAGCAGAGGGGGTGGGAGCAGCATCTCGAGACTTGCAAAAATATGGAGCAATCACAAGTAGC
AATACAGCAGCTACCAATGCTGATT

GTGCCTGGCTGGAAGCACAAGAGGATGAGGAGGTGGGTTTTCCAGTCAAACCTCAGGTACCTTTA
AGACCAATGACTTACAAGGGAGCTG

TAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGTTGATTTATTCCCAGCAAAGACA
AGAAATCCTTGACTTGTGGGTCT

ACCACACACAAGGCTACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGACCAGATATCCA
CTGACCTTTGGGTGGTGCTTCAAGT

TAGTACCAGTGGATCCAGATAAAGTAGAAGAGGCCACTGAAGGAGAAAACAACTGCTTGTTACACC
CTCTGAACCAGCACGGGATGGATG

ATCCAGAGAAGGAAGTGTTAATGTGGAAGTTTGACAGTCGCCTAGCATTTCACCATGTGGCCCGAG
AGCTACATCCGGAGTACTACAAGA

ACTGATGACACTGTGCTTTCTACAAGGGACTTTCCACTGGGGACTTTCCAGGGGAGGTGTGGCCTG
GGCGGGACTGGGGAGTGGCGAGCC

CTCAGATGCTGCATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGA
GCCTGGGAGCTCTCTGGCTGACT

AGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTTAAGTAGTGTGTGCCCGTCT
GTTGTGTGACTCTGGTAACTAGAG

ATCCCTCAGACCATTTAGTCAGTGTGGAAAATCTCTAGCA

>BosT_ERV_chr11_APOB 6777 bp

TGCGGGGAGCCGGTGAGGCATTCCACTCGTGACAAAGGTCATGAGGAAGGAGGCTCGGCATACG
CAAAGGCGGGATCGAGCCTCAGGAGT

CCCCCCGGATATTCTCGAGCATTTTCCCCCAAAAAACCAGAGTCTGCCTACTTTATTGCTTTGTGCT
CTCACCTCTGACTTTACTGGGGG

CTGTCCCCTACCACCGTCTCTCTCTCTCTGTGTCAAAGAGTTAACTTACAGCTCCAATTAATAAAGTT
CCTGGGCAATTAGGAGTGTTTA

AATCCAAACCCCTCTGATGGCTCTCTAACTCGCCTGACAAGTTTACCCGGACTCCTGCAGCTATGC
ATACGATTGTTTACAGTCTCCCAG

CCTCGAGAGGCATGGGAAGCTTAAGATATTCAAATAGCTTAGAGCCTCTCAGAGAGTTAAAAACTG
TCAGAATAAACTAGTAAAGGATTT

CATTGATGAGTCAATGCTTGTTGCCAAGTTTTCACATCCCCTGAATTGTATCCTTGAATATGTATCAA
TTAATAGTGGGTATGTAGAAAA

AATAAGTAGTGGCCTTGGTGTTAGTAACTTTAGACCCTTAAGGTAATAAATTCTTTCTTTGTTGTAAA
CCCATTACACATCCGCCCTATA

GGAATGCAATTTTATCTTTGGAAGATGGTGCCAAACCTTGAAATAATTACTCTTAGAGAAAGTAAGT
CTTTGTTGATAAGTCCTTGTCAA

GAGTCATAAAATGTTAGTAGGCCTTCTGGCCAGAAGATGATGTAAATCACCTAAACCATTTGTATAC
GATACATTTGCAGGAAAGAAACC

TTGGTTTTTGATAAGAATCAAAGACTGCTGACTTTGCATCCCCTATTATCCTCTATGTGTAACTTAGG
GTATAAAAGCCCCTGTTAAAAA

TAAAGCTACGGGCCTTGCTCACCAACGCTTGGTCTCCCCATGTCATTCTTTTAACTTCCAGCTGAGT
CTCCATCTGGAGCGCGGAACCCA

CCACGCTTACTAATCATGCCTGGGCTTCTAAGACCCACTCGAGAAGGTGTCTAGGGTGAGACACCT
TCCGCTATTCGAGAGGGCGCCTGC

GGCCTACGTAAGTGGTGCAAACTTCTTGTCTTGAAGTTTTATTGGTCTCCCGCGTAAACCAAGCTAC
TCAGCTTCTTTTCTCCACTGAAA

TTTCCTACTGAGCTATCCTCATTCTATTGTTCTCTATATCCCTAATTAGCATATAAATAGTCGCCGAC
GCCGTCTCCCCTTCGAATACCC

TGGATCAGCCGGGGCTGGTCCTCGGCAGGTGGCGCCCGATACAGGGATTTCGAAGGTAAGTCCC
CAACCCCATTCCCCCAGTGTTGAGTT

TTCGGGACGGATAGGACCCCACTTAGGGTGCTGCAGACCCCCCTGTAGAATAGACAGGGAGAGAG
GAGTGGGAAGTGTTGAAAGTGTTGA

AGAGTTAGAGAGAAAAAACGATTTCTAAAAAGGCTGACAAAAAAGGCCTAATCTTTTGATAGCTAAA
AGCAAAATCTTTTACTATACTTA

ATTTTCTGACATGGGTAATACTGAATCAAATGAAAGACAGCTCTTTATAGGAGTAATTTTACAGTTAT
TAGGTAAAAGAGGAATTAAAGT

TAAAAAATCTGCCATTCAATCATTTCTTTCATTTGTACAAGAGCACTGTCCCTGGTTTCCAGACGAAG
GCTCTGTTAACTTAGATGTCTG

GGAAAAAGTAGGAAAACAGTTAAAAACTTACCATGCAGAACATGGCTCAGAAAAGGTGCCTAATGA
CGCCTTTTCCTTATGGAATATTAT

TAGAGATGTCTTAGACCCTGCCCCAGATTCAGAAAAAGTACATCTTAAAAGGGATAGTGAAGAAAAT
GCTGTAGTTAAACCTACCCCTGA

ACCTAAAAAAGTAACCTTTAAAGAGGAAAATGAAGTCGAAATCGTAGTTAAACCTGAAGAAAATGAG
AAAAACGAAGACCCGCCTGACTA

TCGGCAACTAAGAAAAATGTTAATAGCCGTGACTACTCAGGAACAACCTAAAGATAGGGATGAGGA
ACAAGATCAGCCTTCCCCAAAACA

AAAAGAGAATTTAGGCGAGATAACAACTAAATATCACTCTGATGGAGATTGGCATCTCTTAAACAAA
GATGCCCCAAAAGGCCTAAGAGA

AACTTCCCCTGTCAGACCATCCGCTCCTAGGAGCTTGAGGGAAACATCCCCGATCAGACAGTTTGT
AAGATTTAGTCACCAAACAGTTAA

AGGATCTCCGGAGCAGGAGAAGAGAAATAGGGGACGCTCCCATCCCTCCATGCCTCCTCCTCTGT
GTGGGGGTAAAGGGCCTCCACTAGG

AGTTTCCCCGAGAAGGGTTTGCTCTAGTCCCAAGGATAAATTTGATCCTCACCTTGAGGCTTCTTTT
CTAGTCCCCTCTGCAGTTTAGAT

AGGGTAAAGAAGGTACAAAATAAATAAAAATATTCTCAAAAAAAAAAAAATAAATAAAAAAAAAAATG
GGTTTCTCTGCATCTAAGAATA

TACAAATGTTTATTGGCTCTGTGAAATATGATTAGAAATGTCCTGGACCCCTCTCATGAGGCTGTTA
AATAGCCTTTGAGGGAGGCTATA

GCGGTAGATGGTGGTGAGTCTCCAGTTTCTGCATAGATCATAATTTCTGCTATTGGAAAAAAAAAAA
AAAAGGAGGGAGAAAGTGCTTTA

CCTCCCGAGCAAGGAGAGGGTTTACAGGACGCTGCGGCCGAGCGCCCTGGCGAGCTCCCTCCCC
CTCTGCGCAAACCTTTAAAAAACTTA

TCCACCACTTTCTGATTTAAGGTGATCGCCACCGCCTCCGTTTGTGCCTCCCAGGACCCAGGAGTT
CCCCACTTTGCCCCCAAAGCCTCT

CAAAGCGTTGCCTAAGCCTGAGGAAGATAAAAAGTTTTTTGAACAGTGGAGCTTTTAAAACAGACTG
CGTGCACAATATGCAGAGCATGC

TTCTTAGAGAAAACCCCCTCAGGGGGGGTCTCACCCAGGCTAGGAGAAAAGCGAAACCGCAAGGG
GATTTAGCAACAATATTAACCCCTG

ATGTGCCGCTTACTCCAATTCCAACGGGAGTGGCTGGCCCCCTACCTTTGGACATTGTAAGATCTG
TCCTGGGGCATAGCTCGCTTTCTT

TTCAAAAAAAAAAAAAAAAAAAATTCGGTGGTGCATGGTGTAGTAGATTCTGATTATATTGAGAAATT
AAAGTCTTGATATCACCGCCTA

CCAAAACTGTACAAATTAATAAAGGTCAAAGAGTAACACAGCCTTTGCTTTTACCTTATTATCAGACA
AGAAAAAACTTGACTTCTCAAG

TTAAGAGCCACGAAACATTTGGATCTAGTGATCTAGCCTTTTGGATGCAGGAAATTACAGCTCCTGG
CCAACACATACTACTGAAAATGA

GTTGGTGGGATTAGAGAAAGTGGTATGTGAGGTGGGGATGCTATACAGGTGGTACATTTAACAATA
CTTTATATATTGTTATAAATACAT

AATATAAATATATTTGCCTAATTACAGTTTGCCTAATTAGGTATGGGTATAAATAAAATATAATAAAGG
TATACCTAATTCTACTTATAG

ATCTATACTTAATTTGTATATAAATTAATATGTATACTATATATACATATATATATACTGTATAATTAAAT
GTATATTGCAGTAATATAA

TGTGTGTGGCTGATATTAATTGGTATGGATTGGTGTGCTGTGATGATATATGTTCTATATACTGTATA
ATTGTATATGTGTGACATGCAT

TATTGCAGTACATTGGTTTGTGTTGGTTGGTATTAGCTGTGTATGTGTTGTATTGCTGTAATATATAT
TAATTAATATATTAATTATAAA

GATTAATTCAAATATATTGATTTATATATATTATATGTCAATAAGTTTATACTTGTTGCCATATATAAAT
ACATATATAAATACATTTTA

CATTTAGGTATGTCTGTATACCAAAATGAGATAAGGAGGTTATACATTTATTACTTAAATTTATACAG
AGACCTAAACTAAACATTAGAC

CTAAATTAACATGTCCCTAGGAAAACAGCTGGACAAATAGTAAAGTTATGTCCAAATTGTAATTAATC
ACTTAAAGGTAAACACACAGGG

GCAGACTACAATGATGAGACACTAGATGCCTGAATTACAGGCTACGACTCGGCCCCTGGTCAAGTA
AAAAGACCTCCTTACTGGAGAGTG

TAAAAGGCCAGATGTGTTGCTAACTTGTGGGAGAGGGTATGCTTATATATTTCTACAGAATGCAGAT
TCTCCGATTTGGATCTCAGAAAA

ATTCGTCATGTCACTTTCCCCAAAGACCAGGTTCGCGGCGCGCGCACTCGCTCGCTCCCCCTCCC
CCAGCCAGCTCTCTCCGCCGCCAGC

GGCGCCCTGCACCCTCCTTGCTGCACCCCGAGACCTAGAGCAAAGAAGTCTGTGCGGCGAGTGAGGGCCAGAGAGGAAAGCGCGCCCGCG

CAGTGCAGTCCAGACCAGCGCGCTCCCGCCGCGGGCGACAGGAAAACGCCCCGAGCCCAAGCGGCGCGGCTAGCCCGAGTCCGCGAACCC

CGCCCCTCCGCCCGCCGTAGACGCCTCGGCCCACAACCTGCCGTTCGCGTTCTGCCTCCTGATGCTGCGGACACTGCTCCGCCGAGCCGA

TGCCTACAGCTGCTCCTTGTCTCATCTTGTTTAAATCCTCAATATAATGTTTGCTCCTTTGTAATGTTACAATGTCCACCTTATCTTATG

ATGTAACCACTTATTGAATTATGCTCTTGCTGTATTACAACAACTGCAGGATTTAATGCGAACACGACGGTTTGTGGGCTTACTTATCTT

GGGAATAGCAGCTTTGATAAGCGCAATTACTTCTGTTACTGTGGCAGCAATATCATTGACTCAACAAGTACATACTGCTCAATATGTTGA

TTCTATGTCCAAAAATGTTTCTTTAGCATTGGCAACACAGGAAGCTATAGACAGGAAATTAGAAATGAGGGTAGATGCCCTAGAGGAAGC

AGTAATACATATTGGGACTGAATTGCAGGCTTTAAAGGTGAAAATGGCATTGTCCTGTCATGCTGACTACCGGTGGATATGTGTAACACC

CCTGAAAGTAAATGAGACAGATTTTGAATGGGAAAAGATTAAAAACCATATTTCAGGTATTTGGAACAGCTCTGACATTAGCTTAGACTT

AGGGAAACTTCACAATCAAATAGCAACCCTGGAACACTCCCGATTAGATTTTACTGCCGCTGGAACAGCAAATGATTTCTTCCATACTTT

CTCTAACTACATTTCAGGAAAAAATATTCTGTCTACCTTCCTCGGCTACGCTACCTTGGCTGTTTTAATTTTATTTCTAATAATCATTCT

TCCTTGTATTGTCAGGATTCTTCGGCAGAGCATTCAGAGGCTCGCGACTGAGCTACATCTGGCTGTTTTAAGAAATAAAAAAGGGGGAGA

TGCGGGGAGCCGGTGAGGCATTCCACTCGTGACAAAGGTCATGAGGAAGGAGGCTCGGCATACGCAAAGGCGGGATCGAGCCTCAGGAGT

CCCCCCGGATATTCTCGAGCATTTTCCCCCAAAAAACCAGAGTCTGCCTACTTTATTGCTTTGTGCTCTCACCTCTGACTTTACTGGGGG

CTGTCCCCTACCACCGTCTCTCTCTCTCTGTGTCAAAGAGTTAACTTACAGCTCCAATTAATAAAGTTCCTGGGCAATTAGGAGTGTTTA

AATCCAAACCCCTCTGATGGCTCTCTAACTCGCCTGACAAGTTTACCCGGACTCCTGCAGCTATGCATACGATTGTTTACAGTCTCCCAG

CCTCGAGAGGCATGGGAAGCTTAAGATATTCAAATAGCTTAGAGCCTCTCAGAGAGTTAAAAACTGTCAGAATAAACTAGTAAAGGATTT

CATTGATGAGTCAATGCTTGTTGCCAAGTTTTCACATCCCCTGAATTGTATCCTTGAATATGTATCAATTAATAGTGGGTATGTAGAAAA

AATAAGTAGTGGCCTTGGTGTTAGTAACTTTAGACCCTTAAGGTAATAAATTCTTTCTTTGTTGTAAACCCATTACACATCCGCCCTATA

GGAATGCAATTTTATCTTTGGAAGATGGTGCCAAACCTTGAAATAATTACTCTTAGAGAAAGTAAGTCTTTGTTGATAAGTCCTTGTCAA

GAGTCATAAAATGTTAGTAGGCCTTCTGGCCAGAAGATGATGTAAATCACCTAAACCATTTGTATAC
GATACATTTGCAGGAAAGAAACC

TTGGTTTTTGATAAGAATCAAAGACTGCTGACTTTGCATCCCCTATTATCCTCTATGTGTAACTTAGG
GTATAAAAGCCCCTGTTAAAAA

TAAAGCTACGGGCCTTGCTCACCAACGCTTGGTCTCCCCATGTCATTCTTTTAACTTCCAGCTGAGT
CTCCATCTGGAGCGCGGAACCCA

CCACGCTTACTAATCATGCCTGGGCTTCTAAGACCCACTCGAGAAGGTGTCTAGGGTGAGACACCT
TCCGCTATTCGAGAGGGCGCCTGC

GGCCTACGTAAGTGGTGCAAACTTCTTGTCTTGAAGTTTTATTGGTCTCCCGCGTAAACCAAGCTAC
TCAGCTTCTTTTCTCCACTGAAA

TTTCCTACTGAGCTATCCTCATTCTATTGTTCTCTATATCCCTAATTAGCATATAAATAGTCGCCGAC
GCCGTCTCCCCTTCGAATACCC

TGGATCAGCCGGGGCTGGTCCTCGGCA