

MX-HIV-0401 PROTEASE FRAGMENT (Fwd & Rev)

>401_PR_F
CAACAACTcCCTCTCAGAAACAGGAGGGGTCAGACAAGGAACTGTATCCTTTTTCTTCCCTCAAATCACTCTTTGGCAACGACCCTCGTCACAATAAAGATAGGGGGCAACTA
AAGGAAGCTCTATTAGATACAGGGGCAGATGATACAGTATTAGAAGAAATGAATTTGTTAGGAAGATGGAACCAAAAATGATAGGGGGAATGGAGGTTTTATCAAGGTAAGAC
AGTATGATCAGGTACCCATAGAAATCTGTGGACATAAAGCTGTAGGTACAGTATTAGTAGGACCTACACCTGTCAACACAATGGGAAGAAATCTGTTGACTCAGATTGGATGCAC
TTTAAATTTTCCATTAGTCTATTGAAACTGTACCAGTAAAATTTAAACCAGGAATGGATGGCCAAAAGTTAAACAATGGCCATTGACAGAAG

>401_PR_R
GAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTTTGGGGAAGAGACAACAACAACCTCCCTCTCAGAAACAGGAGGGGTCAGACAAGGAACTGTATCCTTTTTCTTCCCTCAAAT
CACTCTTTGGCAACGACCCTCGTCACAATAAAGATAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGGGCAGATGATACAGTATTAGAAGAAATGAATTTGTTAGGAAGA
TGGAACCAAAAATGATAGGGGGAATGGAGGTTTTATCAAGGTAAGACAGTATGATCAGGTACCCATAGAAATCTGTGGACATAAAGCTGTAGGTACAGTATTAGTAGGACCTA
CACCTGTCAACACAATGGGAAGAAATCTGTTGACTCAGATTGGATGCACCTTTAAATTTTCCATTAGTCTATTGAAACTGTACCAGTAAAATTTAAACCa

Nucleotide Alignments

```

      1
      |
401_PR_F *****CAACAACTCCCTCTCAGAAACAGGAG
401_PR_R GAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTTTGGGGAAGAGACAA-----

     101
     |
401_PR_F GGGTCAGACAAGGAACTGTATCCTTTTTCTTCCCTCAAATCACTCTTTGGCAACGACCCTCGTCACAATAAAGA
401_PR_R -----

     201
     |
401_PR_F TAGGGGCAACTAAAGGAAGCTCTATTAGATACAGGGGCAGATGATACAGTATTAGAAGAAATGAATTTGTTAG
401_PR_R -----

     301
     |
401_PR_F GAAGATGGAAACCAAAAATGATAGGGGGAATGGAGGTTTTATCAAGGTAAGACAGTATGATCAGGTACCCATAG
401_PR_R -----

     401
     |
401_PR_F AAATCTGTGGACATAAAGCTGTAGGTACAGTATTAGTAGGACCTACACCTGTCAACACAATGGGAAGAAATCTGT
401_PR_R -----

     501
     |
401_PR_F TGACTCAGATTGGATGCACCTTTAAATTTTCCATTAGTCTATTGAAACTGTACCAGTAAAATTTAAACCAGGAA
401_PR_R -----*****

     601
     |
401_PR_F TGGATGGCCAAAAGTTAAACAATGGCCATTGACAGAAG
401_PR_R *****
```

Written by James Robinson, Last modified 11/02/03.

>401_PR
GAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTTTGGGGAAGAGACAACAACAACCTcCCTCTCAGAAACAGGAGGGGTCAGACAAGGAACTGTATCCTTTTTCTTCCCTCAAAT
CACTCTTTGGCAACGACCCTCGTCACAATAAAGATAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGGGCAGATGATACAGTATTAGAAGAAATGAATTTGTTAGGAAGA
TGGAACCAAAAATGATAGGGGGAATGGAGGTTTTATCAAGGTAAGACAGTATGATCAGGTACCCATAGAAATCTGTGGACATAAAGCTGTAGGTACAGTATTAGTAGGACCTA
CACCTGTCAACACAATGGGAAGAAATCTGTTGACTCAGATTGGATGCACCTTTAAATTTTCCATTAGTCTATTGAAACTGTACCAGTAAAATTTAAACCAGGAATGGATGGCC
AAAAGTTAAACAATGGCCATTGACAGAAG

MX-HIV-0401 REVERSE TRANSCRIPTASE FRAGMENT A (Fwd & Rev)

>401_RTa_F

CTATTGACTCAGATTGGTTGTACTTTAAATTTCCCATAGTCTATTGAACTGTACCAGTAAAATTAAGCCAGGAATGGATGGCCAAAAGTGAACAATGGCCATTGACAG
AAGAAAAAATAAAGCATTAGTAGAAATTCGCACAGAAATGGAAAAAGAAGAAAAATTTCCAGGAATGGGCCTGAAAATCCATATAAATACTCCAATATTTGCCATAAAGAAAA
AGACAGTACTAAATGGAGAAAATTAGTAGATTTCCAGAGAACCTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCTGCAGGGCTAAAAAGAAAA
TCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTCCCCTTAGATGAAGACTTCAGAAAGTATACTGCTTTTACCATACCTAGTACAAAACATGAGACACCAGGGATCA
GATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCACCAGCAATTTTCCAAAGCAGCATGACAAAAATCTTAGAACCTTTTGGGAAACGATATCCAGATATAGTTATCTA
TCAATACATGGATGATTTATATGTAGGATCTGACTTAGAAAATAGGGCAGCATAGAGAAAAATAGAGGAACCTGAGACAACATCTGTTAAAGTGGGGGTTTACCACACCAGACAAA
AAACATCAGAAaGAACCTCCATTCCTTTGGATGGGTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGCAGCTGGACAGTCAATGACaTAC
AGAAGTTAGTGGGAAAATTAATTTGGGCAAGTCAgATTTATGCaGGGATTAa

>401_RTa_R

CTcCAatattTTGCCaTAAaGAAAAAGaCAGTACTAAATGGagAAAATTaGtaGATTTCCAGAGAACCTAaTaaGAGAACTCaGACTTCTGGGAAGTTCAATTAGGAATACCACA
TCCTGCAGGGCTAAAAAGAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTCCcCTTAGATGAAGACTTCAGAAAGTATACTGCTTTTACCATACCTAGT
ACAAACaATGAGACACCAGGGATCAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCACCAGCAATTTTCCAAAGCAGCATGACAAAAATCTTAGAACCTTTTGGGA
AACGATATCCAGATATAGTTATCTATCAATACATGGATGATTTATATGTAGGATCTGACTTAGAAAATAGGGCAGCATAGAGAAAAATAGAGGAACCTGAGACAACATCTGTTAA
GTGGGGGTTTACCACACCAGACAAAAAACATCAGAAAGAACTCCATTCCTTtGGATGGGTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAA
GACAGCTGGACAGTCAATGACATACAGAAGTTAGTGGGAAAATTAATTTGGGCAAGTCAgATTTATGCaGGGATTAa

Nucleotide Alignments

```

1
|
401_RTa_F  CTATTGACTCAGATTGGTTGTACTTTAAATTTCCCATAGTCTATTGAACTGTACCAGTAAAATTAAGCCA
401_RTa_R  *****

101
|
401_RTa_F  GGAATGGATGGCCAAAAGTGAACAATGGCCATTGACAGAAGAAAAATAAAAGCAATTAGTAGAAATTCGCACA
401_RTa_R  *****

201
|
401_RTa_F  GAAATGGAAAAAGAAGGAAAAATTTCCAGGAATGGGCCTGAAAATCCATATAAATACTCCAATATTTGCCATAAAG
401_RTa_R  *****-----

301
|
401_RTa_F  AAAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTCCAGAGAACCTAATAAGAGAACTCAAGACTTCTGGGAA
401_RTa_R  -----

401
|
401_RTa_F  GTTCAATTAGGAATACCACATCCTGCAGGGCTAAAAAGAAAAATCAGTAACAGTACTGGATGTGGGTGATGCA
401_RTa_R  -----

501
|
401_RTa_F  TATTTTTCAGTCCCCTTAGATGAAGACTTCAGAAAGTATACTGCTTTTACCATACCTAGTACAAACAATGAGACA
401_RTa_R  -----

601
|
401_RTa_F  CCAGGGATCAGATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCACCAGCAATTTTCCAAAGCAGCATG
401_RTa_R  -----

701
|
401_RTa_F  ACAAAAATCTTAGAACCTTTTGGGAAACGATATCCAGATATAGTTATCTATCAATACATGGATGATTTATATGTA
401_RTa_R  -----

801
|
401_RTa_F  GGATCTGACTTAGAAAATAGGGCAGCATAGAGAAAAATAGAGGAACCTGAGACAACATCTGTTAAAGTGGGGTTT
401_RTa_R  -----

901
|
401_RTa_F  ACCACACCAGACAAAAACATCAGAAAGAACTCCATTCCTTTGGATGGGTATGAACTCCATCCTGATAAATGG
401_RTa_R  -----
```

```

1001
|
401_RTa_F ACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGCTGGACAGTCAATGACATACAGAAGTTAGTGGGAAAATTA
401_RTa_R -----

1101
|
401_RTa_F AATTGGGCAAGTCAGATTTATGCAGGGATTAAA*****
401_RTa_R -----GTAAGGCAATTATGTAAACTCCTTAGAGGAGCTAAAGCACTA

1201
|
401_RTa_F *****
401_RTa_R ACAGAAGTAGTACCACTAACAAAAGAAGCAGAACTAGAAGTGGCAGAAAACAGGGAGAT

```

Written by James Robinson, Last modified 11/02/03.

```

>401_RTa
CTATTGACTCAGATTGGTTGTTACTTTAAATTTTCCCATAGTCTTATTGAACTGTACCAGTAAAATTTAAAGCCAGGAATGGATGGCCAAAAGTGAAACAATGGCCATTGACAG
AAGAAAAAATAAAGCATTAGTAGAAATTCGCACAGAAATGGAAAAAGAGAAAAATTTTCAGGAATTGGGCTGAAAATCCATAAATACTCCAATATTTGCCATAAAGAAAA
AGACAGTACTAAATGGAGAAAATTAGTAGATTTTCAGAGAAGTAAATAAGAGAAGTCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCTGCAGGGCTAAAAAGAAAA
TCAGTAAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTCCCTTAGATGAAGACTTCAGAAAAGTATACTGCTTTTACCATACCTAGTACAAACAATGAGACACCAGGGATCA
GATATCAGTACAATGTGCTTCCACAAGGATGGAAAGGATCACCAGCAATTTTCCAAAGCAGCATGCAAAAATCTTAGAACCTTTTGGGAAACGATATCCAGATATAGTTATCTA
TCAATACATGGATGATTTATATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAGAAAAAATAGAGGAAGTGGAGCAACATCTGTAAAGTGGGGTTTACCACACCAGACAAA
AAACATCAGAAaGAACCTCCATTCCTTTGGATGGGTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAGACAGCTGGACAGTCAATGACaTAC
AGAAGTTAGTGGGAAAATTAATTTGGCAAGTCaGATTTATGCaGGGATTAAGTAAGGCAATTATGTAAACTCCTTAGAGGAGCTAAAGCACTAACAGAAGTAGTACCACTAAC
AAAAGAAGCAGAACTAGAAGTGGCAGAAAACAGGGAGAT

```

MX-HIV-0401 REVERSE TRANSCRIPTASE FRAGMENT B (Fwd & Rev)

>401_RTb_F

AgAaGcAgAacTAGAacTGGCAGAAAAcAgGgAAATTCtaaAACAgCCaGTACaTGGgGTGTATtATGACCcatCaaAagATtAatAgCgAAATaCAGaAgCaGGGacaGGGC
CAatGgacaTATCAaATTTATcAgGaaCCaTtTaAaaaCTGAaaaCaggaAaatATGCaAAAAATGAGGAgTCCCaCACtaatGatGTAAACaaTTAACagAggcaGTGCaaa
aAACaGCCACagaaaGCatAgtaaatAtgAGGAAAGACTCCTAAgTTTAgATTACCCATACaaagGAAGCatGGgaagcATGgTggtCagAGTATTGGCAaGcCACCTGgATcCC
TGAgTGGGAGTTGTcAATACTCCTCCTTAGtaaAAttaTGGtATCAaTtagagACaGaaCCcaTAgTAggAgCAGAAaCtTTCTATgtGgAtGgggCagCTaaTaGGgaaACT
aaatCaggtAAAGCAGATATGTTACTGcAgAgGACAAaAGGtAgTcCCcTtagcggACaCaACaAatCaGAA

>401_RTb_R

CTAGAACTGGCAGAAAAcAGGGAAATTCtAAAAcAgCCAGTACATGGGGtGTATTATGACCcATCAAAAGATTTAATAGCAGAAATACAGAAGCAGGGACAGGGCCAATGGACAT
ATCAAAATTTATCAGGAACCAATTTAAAAATCTGAAAAcAGGAAAAATATGCAAAAAATGAGGAGTGCCcACACTAATGATGTAAAAcAATTAACAGAGGCAGTGCAAAAAACAGCCAC
AGAAAGCATAGTAATATAGGAAAGACTCCTAAGTTTAGATTACCCATACAAAAGGAAGCATGGGAAGCATGGTGGTcAGAGTATTGGCAAGCCACCTGGATCCCTGAGTGGGAG
TTTGTCAATACTCCTCCTTAGTAAATTTAGGATCAATTAGAGACAGAAcCCATAGTAGGAGCAGAAACTTTCTATGTGGATGGGGCAGCTAATAGGGAAACTAAATCAGGTA
AAGCAGGATATGTTACTGACAGAGGAAGACAAAAGGTAGTCCcCTAGCGGACACAACAATCAGAAGACTGAGTTACAAGCAATTCATCTAGCTTTGCAGGATTCAGGATCAGA
AGTAAATATAGTAACAGACTCCCAATATGCATTAGGAATTAATCAAGCGCAACCAGATAAGAGTGAATCAGAAATAGTCAGTCAAATAATAGAACAGTTAATAAAAAAGGAAAA
GTCTACCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGTAATGAACAGGTAGACAAATAGTCAGTAAAGGAATTAGAAAAGTACTATTtCTTAGATGGAATAGATAAG

Nucleotide Alignments

```

1
|
401_RTb_F  AGAAGCAGAACTAGAACTGGCAGAAAAcAGGGAAATTCtAAAAcAGCCAGTACATGGGGTGTATtATGACCcATC
401_RTb_R  *****-----

101
|
401_RTb_F  AAAAGATTTAATAGCAGAAATACAGAAGCAGGGACAGGGCCAATGGACATATCAAAATTTATCAGGAACCATTTAA
401_RTb_R  -----

201
|
401_RTb_F  AAATCTGAAAAcAGGAAAAATATGCAAAAAATGAGGAGTGCCcACACTAATGATGTAAAAcAATTAACAGAGGCAGT
401_RTb_R  -----

301
|
401_RTb_F  GCAAAAAACAGCCACAGAAAGCATAGTAATATAGGAAAGACTCCTAAGTTTAGATTACCCATACAAAAGGAAGC
401_RTb_R  -----

401
|
401_RTb_F  ATGGAAGCATGGTGGTcAGAGTATTGGCAAGCCACCTGGATCCCTGAGTGGGAGTTTGTCAATACTCCTCCTT
401_RTb_R  -----

501
|
401_RTb_F  AGTAAATTTATGGTATCAATTAGAGACAGAACCCATAGTAGGAGCAGAACTTTCTATGTGGATGGGGCAGCTAA
401_RTb_R  -----

601
|
401_RTb_F  TAGGAAACTAAATCAGGTAAAGCAGGATATGTTACTGACAGAGGAAGACAAAAGGTAGTCCcCTAGCGGACAC
401_RTb_R  -----

701
|
401_RTb_F  AACAAATCAGAA*****
401_RTb_R  -----GACTGAGTTACAAGCAATTCATCTAGCTTTGCAGGATTCAGGATCAGAAGTAAATATAGTAAC

801
|
401_RTb_F  *****
401_RTb_R  AGACTCCAATATGCATTAGGAATTAATCAAGCGCAACCAGATAAGAGTGAATCAGAAATAGTCAGTCAAATAAT

901
|
401_RTb_F  *****
401_RTb_R  AGAACAGTTAATAAAAAAGGAAAGGTCtACCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGTAATGAACA
```

1001

401_RTb_F

|

GGTAGACAAATTAGTCAGTAAAGGAATTAGAAAAGTACTATTTCTTAGATGGAATAGATAAG

401_RTb_R

Written by James Robinson, Last modified 11/02/03.

>401_RTb

AGAAGCAGAACTAGAACTGGCAGAAAaCAGGGAAATTCTAAAACAgCCAGTACATGGGGtgTATTATGACCCATCAAAAGATTTAATAGCAGAAATACAGAAGCAGGGACAGGGC
CAATGGACATATCAAATTTATCAGGAACCATTTAAAAATCTGAAAACAGGAAAATATGCAAAAATGAGGAGTGCCACACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAA
AAACAGCCACAGAAAGCATAGTAATATGAGGAAAGACTCCTAAGTTTAGATTACCCATACAAAAGGAAGCATGGGAAGCATGGTGGTCAGAGTATTGGCAAGCCACCTGGATCCC
TGAGTGGGAGTTGTCAATACTCCTCCTTAGTAAAATTATGGTATCAATTAGAGACAGAACCATAGTAGGAGCAGAAACTTTCTATGTGGATGGGCAGCTAATAGGAAACT
AAATCAGGTAAGCAGGATATGTTACTGCAGAGGAAGACAAAAGGTAGTCCCCCTAGCGGACACAACAAATCAGAAGACTGAGTTACAAGCAATTCATCTAGCTTTGCAGGATT
CAGGATCAGAAGTAAATATAGTAACAGACTCCCAATATGCATTAGGAATTATTCAAGCGCAACCAGATAAGAGTGAATCAGAAATAGTCAGTCAAATAATAGAACAGTTAATAAA
AAAGGAAAAGGTCTACCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGTAATGAACAGGTAGACAAATTAGTCAGTAAAGGAATTAGAAAAGTACTATTCTTAGATGGAATA
GATAAG

MX-HIV-0401 INTEGRASE FRAGMENT (Fwd & Rev)

>401_IN_F

TGCTATTTTGGATGGaatAGATAAGGCCAAGAAGAcCATGAAAAATATCACACTAATTGGAGGGCAATGGCTAGTGATTTAACCTGCCACCTGTAGTAGCAAAAGAAATAGT
AGCCAGCTGTGATAAATGTGCTAGCTAAAAGGGGAAGCCATACATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGAAAAATTTCTGGTA
GCAGTACATGTAGCCAGTGGATATATAGAAGCAGAAGTCATTCAGCAGAGACAGGACAGGAAACAGCATACTTCCCTCTTAAAATTAGCAGGAAGATGGCCAGTAAAATGATAC
ATACAGACAATGGCCCAATTTTCATCAGTAATGCGGTTAAGGCCCGCTGTTGGTGGGCAGGGATCAAGCAGGAATTTGGCATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGA
ATCTATGAATAAAGAATTAAGAAAAATATAGGACAGGTAAGAGATCAAGCTGAACATCTTAAAGACAGCAGTACAGATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGG
ATTGGGGGATACAGTGCAGGGGAAAGAATAATAGACATAATAGCAACAGACATACAAAGTAAAGAACTACAAAAACAATTAACAAAATTTTCGGGTTTATTACAGGG
ACAGCAGAGACCACCTTTGAAAGGACCAGCAAAGCTTCTCTGGAAGGTTAAGGGGCAGTAGTAATACAAGATAAATAGTGACATAAAAGTAGTGCCAAGAGAAAAGCAAAGAT
CATTAGGGATTATGGAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAGAACATGG

>401_IN_R

TGGAATCAGAAAAGTCTATTTTGGATgGAATAGATAAagCcCAAGAAGaCCATGAAAAATATCACACTAATTGGAGGGCAATGGCTAGTGATTTAACCTGCCACCTGTAGTA
GCAAAAGAAATAGTAGCCAGCTGTGATAAATGTGCTAGCTAAAAGGGGAAGCCATACATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAA
AAATTTCTGGTAGCAGTACATGTAGCCAGTGGATATATAGAAGCAGAAGTCATTCAGCAGAGACAGGACAGGAAACAGCATACTTCCCTCTTAAAATTAGCAGGAAGATGGCC
AGTAAAATGATACATACAGACAATGGCCCAATTTTCATCAGTAATGCGGTTAAGGCCCGCTGTTGGTGGGCAGGGATCAAGCAGGAATTTGGCATTCCCTACAATCCCCAAAGT
CAAGGAGTAGTAGAATCTATGAATAAAGAATTAAGAAAAATATAGGACAGGTAAGAGATCAAGCTGAACATCTTAAAGACAGCAGTACAGATGGCAGTATTCATCCACAATTTTA
AAAGAAAAGGGGGATTTGGGGATACAGTGCAGGGGAAAGAATAATAGACATAATAGCAACAGACATACAAAGTAAAGAACTACAAAAACAATTAACAAAATTTTCGG
GGTTTATTACAGGGACAGCAGAGACCACCTTTGGAAGGACCAGCAAAGCTTCTCTGGAAGGTTAAGGGGCAGTAGTAATACAAGATAAATAGTGACATAAAAGTAGTGCCAAGA
AGAAAAGCAAAGATCATTAGGGATTATGGAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGA

Nucleotide Alignments

```

      1
      |
401_IN_F *****TGCTATTTTGGATGGAAATAGATAAGGCCAAGAAGACCATGAAAAATATCACACTAATTG
401_IN_R TGGAATCAGAAAAG-----

      101
      |
401_IN_F GAGGGCAATGGCTAGTGATTTAACCTGCCACCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCA
401_IN_R -----

      201
      |
401_IN_F GCTAAAAGGGGAAGCCATACATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGA
401_IN_R -----

      301
      |
401_IN_F AGGAAAAATTTATCTGGTAGCAGTACATGTAGCCAGTGGATATATAGAAGCAGAAGTCATTCAGCAGAGACAGG
401_IN_R -----

      401
      |
401_IN_F ACAGGAAACAGCATACTTCCCTTAAAATTAGCAGGAAGATGGCCAGTAAAATGATACATACAGACAATGGCCC
401_IN_R -----

      501
      |
401_IN_F CAATTTTCATCAGTAATGCGGTTAAGGCCCGCTGTTGGTGGGCAGGGATCAAGCAGGAATTTGGCATTCCCTACAA
401_IN_R -----

      601
      |
401_IN_F TCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAGAAAAATTTAGGACAGGTAAGAGATCAAGC
401_IN_R -----

      701
      |
401_IN_F TGAACATCTTAAAGACAGCAGTACAGATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGGGATTTGGGGGATA
401_IN_R -----

      801
      |
401_IN_F CAGTGCAGGGGAAAGAATAATAGACATAATAGCAACAGACATACAAAGTAAAGAACTACAAAAACAATTAACAAA
401_IN_R -----
```

```

          901
          |
401_IN_F  ACTTCAAAAATTTTCGGGTTTATTACAGGGACAGCAGAGACCCACTTTGGAAAGGACCAGCAAAGCTTCTCTGGAA
401_IN_R  -----

          1001
          |
401_IN_F  AGGTGAAGGGGCAGTAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAAGAAGAAAAGCAAAGATCATTAG
401_IN_R  -----

          1101
          |
401_IN_F  GGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAGAACATGG
401_IN_R  -----*****

```

Written by James Robinson, Last modified 11/02/03.

```

>401_IN
TGGAATCAGAAAAGTCTATTTTTCGATGGaatAGATAAGGCCCAAGAAGAcCATGAAAAATATCACAATAATTGGAGGGCAATGGCTAGTGATTTTAACTGCCACCTGTAGTA
GCAAAAAGAAATAGTAGCCAGCTGTGATAAAATGTCAGCTAAAAGGGGAAAGCCATACATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAA
AAATTTATCTGGTAGCAGTACATGTAGCCAGTGGATATATAAGAAGCAGAAGTCATTCAGCAGAGACAGGACAGGAAACAGCATACTTCTCTTAAAATTAGCAGGAAGATGGCC
AGTAAAAATGATACATACAGACAATGGCCCAATTCATCAGTAATGCGGTTAAGGCCGCCCTGTTGGTGGGCAGGGATCAAGCAGGAATTTGGCATTCCCTACAATCCCCAAAGT
CAAGGAGTAGTAGAATCTATGAATAAAGAATTAAGAAAATATAGGACAGGTAAGAGATCAAGCTGAACATCTTAAGACAGCAGTACAGATGGCAGTATTCATCCACAATTTTA
AAAGAAAAGGGGGGATTGGGGGATACAGTGCAGGGGAAAGAATAATAGACATAATAGCAACAGACATACAAAGTAAAGAACTACAAAAACAAATTACAAAACCTTCAAAAATTTTCG
GGTTTATTACAGGGACAGCAGAGACCCACTTTGGAAAGGACCAGCAAAGCTTCTCTGGAAAGGTGAGGGGGCAGTAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAAGA
AGAAAAGCAAAGATCATTAGGATTATGAAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAGAACATGG

```

MX-HIV-0401 CONTIG ASSEMBLY (4 Fragments)

Nucleotide Alignments

```
2085 | 2159  
B.FR.83.HXB2 | TTT TTT AGG GAA GAT CTG GCC TTC CTA CAA GGG AAG GCC AGG GAA TTT TCT TCA GAG CAG ACC AGA GCC AAC AGC |  
401_PR | *** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** |  
401_RTa | *** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** **^ ^ ^ ^ ^ ^ |  
401_RTb | *** ** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_IN | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
      | F F R E D L A F L Q G K A R E F S S E Q T R A N S
```

```
2160 | 2231  
B.FR.83.HXB2 | CCC ACC AGA AGA GAG CTT CAG GTC TGG GGT AGA GAC AAC *** AAC TCC CCC TCA GAA GCA GGA GCC GAT AGA CAA |  
401_PR | --- --- --- --- --- --- --- --- --- --- AAC --- -T- --- -TC --- --- --- |  
401_RTa | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_RTb | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_IN | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
      | P T R R E L Q V W G R D N N N S P/L S E A/T G A/G D/V R Q
```

```
2232 | 2253 | 2306  
B.FR.83.HXB2 | GGA ACT GTA TCC TTT AAC TTC CCG CAG GTC ACT CTT TGG CAA CGA CCC CTC GTC ACA ATA AAG ATA GGG GGG CAA |  
401_PR | --- --- --- --- --- TT- --- -A A- --- --- --- AAC --- -A --- --- --- -A --- |  
401_RTa | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_RTb | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_IN | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
      | G T V S F N/F F C Q V/I T L W Q R P L V T I K I G G Q
```

```
2307 | 2381  
B.FR.83.HXB2 | CTA AAG GAA GCT CTA TTA GAT ACA GGA GCA GAT GAT ACA GTA TTA GAA GAA ATG AGT TTG CCA GGA AGA TGG AAA |  
401_PR | --- --- --- --- --- TT- --- --- -G --- --- --- -A --- --- --- -A --- |  
401_RTa | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_RTb | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_IN | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
      | L K E A L L D T G A D D T V L E E M S/N L P/L G R W K
```

```
2382 | 2456  
B.FR.83.HXB2 | CCA AAA ATG ATA GGG GGA ATT GGA GGT TTT ATC AAA GTA AGA CAG TAT GAT CAG ATA CTC ATA GAA ATC TGT GGA |  
401_PR | --- --- --- --- --- --- --- --- --- --- -G --- --- --- --- --- G- -C- TT- --- --- --- |  
401_RTa | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_RTb | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_IN | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
      | P K M I G G I G G F I K V R Q Y D Q I/V L/P I E I C G
```

```
2457 | 2531  
B.FR.83.HXB2 | CAT AAA GCT ATA GGT ACA GTA TTA GTA GGA CCT ACA CCT GTC AAC ATA ATT GGA AGA AAT CTG TTG ACT CAG ATT |  
401_PR | --- --- --- --- --- --- --- --- --- --- --- --- --- --- -C- --- --- --- --- -A --- |  
401_RTa | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_RTb | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_IN | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
      | H K A I/V G T V L V G P T P V N I/T I G R N L L T Q I
```

```
2532 | 2549 | 2550 | 2606  
B.FR.83.HXB2 | GGT TGC ACT TTA AAT TTT CCG ATT AGC CCT ATT GAG ACT GTA CCA GTA AAA TTA AAG CCA GGA ATG GAT GGC CCA |  
401_PR | --- --- --- --- --- --- --- --- --- --- --- -T --- --- --- -A --- --- --- -A --- |  
401_RTa | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_RTb | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_IN | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
      | G C T L N F C I S P I E T V P V K L K P G M D G P
```

```
2607 | 2681  
B.FR.83.HXB2 | AAA GTT AAA CAA TGG CCA TTG ACA GAA GAA AAA ATA AAA GCA TTA GTA GAA ATT TGT ACA GAG ATG GAA AAG GAA |  
401_PR | --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- C-C --- -A --- -A --- |  
401_RTa | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_RTb | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
401_IN | *** ** **^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ ^ |  
      | K V K Q W P L T E E K I K A L V E I C/R T E M E K E
```


2682 2756
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
GGG AAA ATT TCA AAA ATT GGG CCT GAA AAT CCA TAC AAT ACT CCA GTA TTT GCC ATA AAG AAA AAA GAC AGT ACT
*** **

2757 2831
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
AAA TGG AGA AAA TTA GTA GAT TTC AGA GAA CTT AAT AAG AGA ACT CAA GAC TTC TGG GAA GTT CAA TTA GGA ATA
*** **

2832 2906
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
CCA CAT CCC GCA GGG TTA AAA AAG AAA AAA TCA GTA ACA GTA CTG GAT GTG GGT GAT GCA TAT TTT TCA GTT CCC
*** **

2907 2981
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
TTA GAT GAA GAC TTC AGG AAG TAT ACT GCA TTT ACC ATA CCT AGT ATA AAC AAT GAG ACA CCA GGG ATT AGA TAT
*** **

2982 3056
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
CAG TAC AAT GTG CTT CCA CAG GGA TGG AAA GGA TCA CCA GCA ATA TTC CAA AGT AGC ATG ACA AAA ATC TTA GAG
*** **

3057 3131
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
CCT TTT AGA AAA CAA AAT CCA GAC ATA GTT ATC TAT CAA TAC ATG GAT GAT TTG TAT GTA GGA TCT GAC TTA GAA
*** **

3132 3206
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
ATA GGG CAG CAT AGA ACA AAA ATA GAG GAG CTG AGA CAA CAT CTG TTG AGG TGG GGA CTT ACC ACA CCA GAC AAA
*** **

3207 3281
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
AAA CAT CAG AAA GAA CCT CCA TTC CTT TGG ATG GGT TAT GAA CTC CAT CCT GAT AAA TGG ACA GTA CAG CCT ATA
*** **

3282 3356
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
GTG CTG CCA GAA AAA GAC AGC TGG ACT GTC AAT GAC ATA CAG AAG TTA GTG GGG AAA TTG AAT TGG GCA AGT CAG
*** **
--- ---
*** **
*** **
V L P E K D S W T V N D I Q K L V G K L N W A S Q

3357 3431
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
ATT TAC CCA GGG ATT AAA GTA AGG CAA TTA TGT AAA CTC CTT AGA GGA ACC AAA GCA CTA ACA GAA GTA ATA CCA
*** **
--- -T G---
*** **
*** **
I Y P/A G I K V R Q L C K L L R G T/A K A L T E V I/V P

3432 3506
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
CTA ACA GAA GAA GCA GAG CTA GAA CTG GCA GAA AAC AGA GAG ATT CTA AAA GAA CCA GTA CAT GGA GTG TAT TAT
*** **
--- --- A---
*** **
*** **
L T E/K E A E L E L A E N R E I L K E/Q P V H G V Y Y

3507 3581
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
GAC CCA TCA AAA GAC TTA ATA GCA GAA ATA CAG AAG CAG GGG CAA GGC CAA TGG ACA TAT CAA ATT TAT CAA GAG
*** **
--- ---
*** **
*** **
D P S K D L I A E I Q K Q G Q G Q W T Y Q I Y Q E

3583 3656
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
CCA TTT AAA AAT CTG AAA ACA GGA AAA TAT GCA AGA ATG AGG GGT GCC CAC ACT AAT GAT GTA AAA CAA TTA ACA
*** **
--- ---
*** **
*** **
P F K N L K T G K Y A R/K M R G/S A H T N D V K Q L T

3657 3731
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
GAG GCA GTG CAA AAA ATA ACC ACA GAA AGC ATA GTA ATA TGG GGA AAG ACT CCT AAA TTT AAA CTG CCC ATA CAA
*** **
--- ---
*** **
*** **
E A V Q K I/T T/A T E S I V I X G K T P K F K L P I Q

3732 3806
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
AAG GAA ACA TGG GAA ACA TGG TGG ACA GAG TAT TGG CAA GCC ACC TGG ATT CCT GAG TGG GAG TTT GTT AAT ACC
*** **
--- G---
*** **
*** **
K E T W E T W W T E Y W Q A T W I P E W E F V N T

3807 3869 3870 3881
B.FR.83.HXB2
401_PR
401_RTa
401_RTb
401_IN
CCT CCC TTA GTG AAA TTA TGG TAC CAG TTA GAG AAA GAA CCC ATA GTA GGA GCA GAA ACC TTC TAT GTA GAT GGG
*** **
--- ---
*** **
*** **
P P L V K L W Y Q L E K E P I V G A E T F Y V D G

3882 3956
 B.FR.83.HXB2 |-----
 401_PR GCA GCT AAC AGG GAG ACT AAA TTA GGA AAA GCA GGA TAT GTT ACT AAT AGA GGA AGA CAA AAA GTT GTC ACC CTA
 401_RTa ***
 401_RTb --- --- --T --- --A --- --- -C- --T --- --- --- --G-C --- --- --- --- --G --A --- C--- ---
 401_IN ***
 A A N R E T K L G K A G Y V T N R G R Q K V V T L

3957 4031
 B.FR.83.HXB2 |-----
 401_PR ACT GAC ACA ACA AAT CAG AAG ACT GAG TTA CAA GCA ATT TAT CTA GCT TTG CAG GAT TCG GGA TTA GAA GTA AAC
 401_RTa ***
 401_RTb G-G --- --- --- --- --- --- --- --- --- --- --- C--- --- --- --- --- --A --- -C- --- --- --T
 401_IN ***
 T D T T N Q K T E L Q A I Y L A L Q D S G L E V N

4032 4106
 B.FR.83.HXB2 |-----
 401_PR ATA GTA ACA GAC TCA CAA TAT GCA TTA GGA ATC ATT CAA GCA CAA CCA GAT CAA AGT GAA TCA GAG TTA GTC AAT
 401_RTa ***
 401_RTb --- --- --- --- --C --- --- --- --- --T --- --- --G --- --- --A-G --- --- --A A--- --- --G-
 401_IN ***
 I V T D S Q Y A L G I I Q A Q P D Q S E S E L V N

4107 4181
 B.FR.83.HXB2 |-----
 401_PR CAA ATA ATA GAG CAG TTA ATA AAA AAG GAA AAG GTC TAT CTG GCA TGG GTA CCA GCA CAC AAA GGA ATT GGA GGA
 401_RTa ***
 401_RTb --- --- --- --A ---
 401_IN ***
 Q I I E Q L I K K E K V Y L A W V P A H K G I G G

4182 4229 4230 4256
 B.FR.83.HXB2 |-----Pol p66 and RNase H end-----|-----Pol p31 Integrase start-----|
 401_PR AAT GAA CAA GTA GAT AAA TTA GTC AGT GCT GGA ATC AGG AAA GTA CTA TTT TTA GAT GGA ATA GAT AAG GCC CAA
 401_RTa ***
 401_RTb --- --- --G --- --C ---
 401_IN ***
 N E Q V D K L V S A G I R K V L L D G I D K A Q

4257 4331
 B.FR.83.HXB2 |-----
 401_PR GAT GAA CAT GAG AAA TAT CAC AGT AAT TGG AGA GCA ATG GCT AGT GAT TTT AAC CTG CCA CCT GTA GTA GCA AAA
 401_RTa ***
 401_RTb ***
 401_IN --A --C --- --A --- --- --C- --- --- --G --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
 D/E E/D H E K Y H S/T N W R A M A S D F N L P P V V A K

4332 4406
 B.FR.83.HXB2 |-----
 401_PR GAA ATA GTA GCC AGC TGT GAT AAA TGT CAG CTA AAA GGA GAA GCC ATG CAT GGA CAA GTA GAC TGT AGT CCA GGA
 401_RTa ***
 401_RTb ***
 401_IN ---
 E I V A S C D K C Q L K G E A M/I H G Q V D C S P G

4407 4481
 B.FR.83.HXB2 |-----
 401_PR ATA TGG CAA CTA GAT TGT ACA CAT TTA GAA GGA AAA GTT ATC CTG GTA GCA GTT CAT GTA GCC AGT GGA TAT ATA
 401_RTa ***
 401_RTb ***
 401_IN ---
 I W Q L D C T H L E G K V/I I L V A V H V A S G Y I

4482 4556
|-----|
B.FR.83.HXB2 GAA GCA GAA GTT ATT CCA GCA GAA ACA GGG CAG GAA ACA GCA TAT TTT CTT TTA AAA TTA GCA GGA AGA TGG CCA
401_PR *** ** *
401_RTa *** ** *
401_RTb *** ** *
401_IN --- --- --- C --- --- --- G --- A --- C --- C ---
E A E V I P A E T G Q E T A Y F L L K L A G R W P

4557 4631
|-----|
B.FR.83.HXB2 GTA AAA ACA ATA CAT ACT GAC AAT GGC AGC AAT TTC ACC GGT GCT ACG GTT AGG GCC GCC TGT TGG TGG GCG GGA
401_PR *** ** *
401_RTa *** ** *
401_RTb *** ** *
401_IN --- --- --- TG --- --- --- A --- --- CC --- --- --- T --- A --- AA --- G --- --- A --- --- A --- G ---
V K T/M I H T D N G S/P N F T/I G/S A/N T/A V R/K A A C W W A G

4632 4706
|-----|
B.FR.83.HXB2 ATC AAG CAG GAA TTT GGA ATT CCC TAC AAT CCC CAA AGT CAA GGA GTA GTA GAA TCT ATG AAT AAA GAA TTA AAG
401_PR *** ** *
401_RTa *** ** *
401_RTb *** ** *
401_IN --- --- --- --- --- --- --- C --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
I K Q E F G I P Y N P Q S Q G V V E S M N K E L K

4707 4781
|-----|
B.FR.83.HXB2 AAA ATT ATA GGA CAG GTA AGA GAT CAG GCT GAA CAT CTT AAG ACA GCA GTA CAA ATG GCA GTA TTC ATC CAC AAT
401_PR *** ** *
401_RTa *** ** *
401_RTb *** ** *
401_IN --- --- --- --- --- --- --- A --- --- --- --- --- --- --- --- --- --- --- G --- --- --- --- --- --- --- ---
K I I G Q V R D Q A E H L K T A V Q M A V F I H N

4782 4856
|-----|
B.FR.83.HXB2 TTT AAA AGA AAA GGG GGG ATT GGG GGG TAC AGT GCA GGG GAA AGA ATA GTA GAC ATA ATA GCA ACA GAC ATA CAA
401_PR *** ** *
401_RTa *** ** *
401_RTb *** ** *
401_IN --- --- --- --- --- --- --- A --- --- --- --- --- --- --- --- --- --- --- A --- --- --- --- --- --- --- ---
F K R K G G I G G Y S A G E R I V/I D I I A T D I Q

4856 4931
|-----|
B.FR.83.HXB2 ACT AAA GAA TTA CAA AAA CAA ATT ACA AAA ATT CAA AAT TTT CGG GTT TAT TAC AGG GAC AGC AGA AAT CCA CTT
401_PR *** ** *
401_RTa *** ** *
401_RTb *** ** *
401_IN -G- --- --- C- --- --- --- --- --- C- --- --- --- --- --- --- --- --- --- --- G-C --- --- --- ---
T/S K E L Q K Q I T K I/L Q N F R V Y Y R D S R N/D P L

4932 5006
|-----|
B.FR.83.HXB2 TGG AAA GGA CCA GCA AAG CTC CTC TGG AAA GGT GAA GGG GCA GTA GTA ATA CAA GAT AAT AGT GAC ATA AAA GTA
401_PR *** ** *
401_RTa *** ** *
401_RTb *** ** *
401_IN --- --- --- --- --- --- --- T --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
W K G P A K L L W K G E G A V V I Q D N S D I K V

5007 5081
|-----|
B.FR.83.HXB2 GTG CCA AGA AGA AAA GCA AAG ATC ATT AGG GAT TAT GGA AAA CAG ATG GCA GGT GAT GAT TGT GTG GCA AGT AGA
401_PR *** ** *
401_RTa *** ** *
401_RTb *** ** *
401_IN ---
V P R R K A K I I R D Y G K Q M A G D D C V A S R

```

5082          5096
|---Integrase end---|
B.FR.83.HXB2  CAG GAT GAG GAT TAG *** **
401_PR       *** **
401_RTa     *** **
401_RTb     *** **
401_IN      --- --- AAC ATG G**
              Q  D  E  D  X

```

Protein Alignments

```

1      10      20      30      40      50      60      70      80      90      100
B.FR.83.HXB2  FFREDLAFLQ GKAREFSSEQ TRANSPTRRE LQWGRDN.N SPSEAGADRO GTVSFNFQGV TLWQRPLVTI KIGGQLKEAL LDTGADDTVL EEMSLPGRWK
401_PR       *****
401_RTa     *****
401_RTb     *****
401_IN_F     *****

```

```

101     110     120     130     140     150     160     170     180     190     200
B.FR.83.HXB2  PKMIGGIGGF IKVROYDQIL IEICGHKAIK TVLVGPTPVN IIGRNLLTQI GCTLNFQISP IETVPVKLKP GMDGPKVKQW PLTEEKIKAL VEICTEMEKE
401_PR       -----
401_RTa     -----
401_RTb     -----
401_IN_F     -----

```

```

201     210     220     230     240     250     260     270     280     290     300
B.FR.83.HXB2  GKISKIGPEN PYNTPVFAIK KKDSTKWRKL VDFRELNKRT QDFWEVQLGI PHPAGLKKKK SVTVLDVGDG YFSVPLDEDF RKYTAFTIPS INNETPGIRY
401_PR       *****
401_RTa     *****
401_RTb     *****
401_IN_F     *****

```

```

301     310     320     330     340     350     360     370     380     390     400
B.FR.83.HXB2  QYNVLPQGWK GSPAIFQSSM TKILEPFRKQ NPDIVIQYQM DDLYVGSLE IGQHRTKIEE LRQHLLRWGL TTPDKKHQKE PPFLWMGYEL HPDKWTVQPI
401_PR       *****
401_RTa     *****
401_RTb     *****
401_IN_F     *****

```

```

401     410     420     430     440     450     460     470     480     490     500
B.FR.83.HXB2  VLPEKDSWTV NDIQKLVGKL NWSAQIYPIG KVRQLCKLLR GIKALTEVTP LEEAELELA ENREILKEPV HGVVYDPSKD LIAEIQKQGG QGWTYQIQE
401_PR       *****
401_RTa     *****
401_RTb     *****
401_IN_F     *****

```

```

501     510     520     530     540     550     560     570     580     590     600
B.FR.83.HXB2  PFKNLKTGKY ARMRGAHTND VKQLTEAVQK ITTESIVIWG KTPKFKLPIQ KETWETWTE YWQATWIPEW EFVNTPLPVK LWYQLEKEPI VEAETFYVDG
401_PR       *****
401_RTa     *****
401_RTb     *****
401_IN_F     *****

```

```

601     610     620     630     640     650     660     670     680     690     700
B.FR.83.HXB2  AANRETKLGG AGYVTNRGRQ KVVTLTDTTN QKTELQAIYL ALQDSGLEVN IVDTSQYALG IIQAQPDQSE SELVNQIEQ LIKKEKVYLA WVPAHKGIGG
401_PR       *****
401_RTa     *****
401_RTb     *****
401_IN_F     *****

```

```

701     710     720     730     740     750     760     770     780     790     800
B.FR.83.HXB2  NEQVDKLVSA GIRKVLFLDG IDKAQDEHEK YHSNWRAMAS DFNLPVPAK EIVASCDKQK LKGEAMHGQV DCSPGIWQLD CTHLEKGVIL VAVHVASGYI
401_PR       *****
401_RTa     *****
401_RTb     *****
401_IN_F     *****

```

```

      801      810      820      830      840      850      860      870      880      890      900
B.FR.83.HXB2  EAEVIPAETG QETAYFL LKL AGRWPVK T I H TDNGSNFTGA TVRAACWAG IKQEFGIPYN PQSQGVVESH NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN
401_PR       *****
401_RTa      *****
401_RTb      *****
401_IN_F     -----M-----P-----ISN-----A-----K-----

      901      910      920      930      940      950      960      970      980      990      1000
B.FR.83.HXB2  FKRRGGIGGY SAGERI V D I I ATDIQ T K E L Q KQITK I Q N F R VYRDRSRNPL WKGPAKLLWK GEGAVVIQDN SDIKVPRRK AKIIRDYGKQ MAGDDCVASR
401_PR       *****
401_RTa      *****
401_RTb      *****
401_IN_F     -----I-----S-----L-----D-----

      1001
B.FR.83.HXB2  QDEDX
401_PR       *****
401_RTa      *****
401_RTb
401_IN_F     -----

```

Written by James Robinson, Last modified 11/02/03.

>401_pol

```

GAGCCACAGCCCCACCAGAAGAGAGCTTCAGGTTTGGGGAAGAGACAACAACAACCTcCCTCTCAGAAACAGGAGGGGTGAGACAAGGAACGTGATCCTTTTTCTTCCCTCAAAT
CACTCTTTGGCAACGACCACCTCGTCACAATAAAGATAGGGGGACAACATAAGGAAGCTCTATTAGATACAGGGGCAGATGATACAGTATTAGAAGAAATGAATTTGTAGGAAGA
TGGAAACCAAAAATGATAGGGGGAATFGAGGTTTATCAAGGTAAGACAGTATGATCAGGTACCCATAGAAATCTGTGGACATAAAGCTGAGGTACAGTATTAGTAGGACCTA
CACCTGTCAACAACAATTGGAAGAAATCTRTTGACTCAGATGGWTGYACTTTAAATTTCCCATTAGTCCATTGAAACTGTACCAGTAAAAATTAARCCAGGAATGGATGGCCC
AAAAGTKAAAACAATGGCCATTGACAGAAAGAAAAATAAAGCATTAGTAGAAATTCGCACAGAAAATGGAAAAAGAAAGAAAAATTTTCAGGAATTTGGCCCTGAAAAATCCATATAAT
ACTCCAATATTTGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCAC
ATCCTGCAGGGCTAAAAAGAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTCCCTTAGATGAAGACTTCAGAAAGTATACTGCTTTTACCATACCTAG
TACAAAACAATGAGACACCAGGATCAGATATCAGTACAATGTCTTCCACAAGGATGGAAGGATCACCAGCAATTTTCCAAGCAGCATGACAAAAATCTTAGAACCTTTTGGG
AAACGATATCCAGATATAGTTATCTATCAATACATGGATGATTTATATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAGAAAAATAGAGAACTGAGACAACATCTGTAA
AGTGGGGGTTTACCACACCAGACAAAAACATCAGAAaGAACTCCATTCCTTTGGATGGGTATGAACTCCATCTGATAAATGGACAGTACAGCTATAGTGCTGCCAGAAAA
AGACAGCTGGACAGTCAATGACaTACAGAAGTTAGTGGGAAAAATTAATTTGGGCAAGTCAGATTTTATGCaGGGATTAAGTAAGGCAATTTATGTAAACTCCTTAGAGGAGCTAAA
GCACATAACAGAAGTAGTACCATAACAAAAGAAAGCAGAACTAGAACTGGCAGAAAACAGGGARATTCTAAAACAgCCAGTACATGGGGtgTATTATGACCCATCAAAGATTTAA
TAGCAGAAATACAGAAGCAGGGACAGGGCCAAATGGACATATCAAATTTATCAGGAACATTTAAAAATCTGAAAACAGGAAAAATAGCAAAAATGAGGAGTGCCCACTAATGA
TGTAACAATTAACAGAGGAGTGCAAAAACAGCCACAGAAAGCATAGTAATATGAGGAAAGACTCCTAAGTTTAGATTACCCATACAAAAGGAGCATGGGAAGCATGGTGG
TCAGAGTATTGGCAAGCCCTGGATCCCTGAGTGGGAGTTTGTCAATACCTCCTCTTAGTAAAAATTTAGGTATCAATTAGAGACAGAACCATAGTAGGAGCAGAAAATTTCT
ATGTGGATGGGGCAGCTAATAGGGAACTAAATCAGGTAAGCAGGATATGTACTGACAGAGGAAGACAAAAGGTAGTCCCTCCTAGCGGACACAAACAATCAGAAGACTGAGTT
ACAAGCAATTCATCTAGCTTTGCAAGGATTCAGGATCAGAAGTAAATATAGTAACAGACTCCCAATATGCATTAGGAATTATCAAGCGCAACCAGATAAGAGTGAATCAGAAATA
GTCAGTCAAAATAATAGAACAGTAAATAAAAAAGGAAAGGTTACCTGGCATGGGTACCAGCACACAAGGAATTTGGAGGTAATGAACAGGTAGACAAAATTAGTCAGTAAWGGAA
TYAGAAAAGTRCTATTTYTTRGATGGAAATAGATAAGGCCCAAGAAGAcCATGAAAAATATCACACTAATTGGAGGGCAATGGCTAGTGATTTTAACTGCCACCTGTAGTAGCAAA
AGAAATAGTAGCCAGCTGTGTATAATGTCACTAAAAGGGGAGCCATACATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAAAAAT
ATTCTGGTAGCAGTACATGTAGCCAGTGGATATATAGAAGCAGAAGTCATTCAGCAGAGACAGGACAGGAAACAGCATACTTCCCTTAAAATTAGCAGGAAGATGGCCAGTAA
AAATGATACATACAGACAATGGCCCAATTTTCATCAGTAATGCGGTTAAGGCCGCTGTTGGTGGGCAGGGATCAAGCAGGAATTTGGCATTCCTTACAATCCCCAAAGTCAAGG
AGTAGTAGAATCTATGAATAAAGAAATTAAGAAAAATATAGGACAGGTAAGAGATCAAGCTGAACATCTTAAAGACAGCAGTACAGATGGCAGTATTCATCCACAATTTTAAAAGA
AAAGGGGGGATTTGGGGGATACAGTGCAGGGGAAAGAAATAATAGACATAATAGCAACAGACATACAAAGTAAAGAACTACAAAAACAATTAACAAAATTTCAAATTTTCGGGTTT
ATTACAGGGACAGCAGAGCCACTTTGGAAAGGACCAGCAAGCTTCTCTGGAAAGGTGAAGGGCAGTAGTAATACAAGATAATAGTGACATAAAAGTAGTGCCAGAAGAAAGAA
AGCAAGATCATTAGGATTATGGAACAAGATGGCAGGTGATGATTTGTGTGGCAAGTAGACAGGATGAGGATTAGAACAATGG

```