



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGCAGAGCGACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGTGCCTA
MfeI (82)
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACAGCGCCGCCGCCCTACCTGAGGCC
PvuII (239)
301 GCCATCCACGCGGTTGAGTGCAGTCTGCCGCTCCCGCCTGTGGTGCTCTGAACTGCGTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMI (441)
NaeI (441) 401 GGGCCTTTGTCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGTCTTTGTTTCGTTT

KasI (535) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCCTA
AgeI (552) CCTGAGATCACCGGTCATCATGAAGGAGTCATCTTTGCAAAATAGCTCCTGCAGCCT
BspHI (560) 1▶Me tLysGI uSer Ser LeuGI nAsnSer Ser CysSer Le

601 GGGAAAGGAGACTAAAAAGGAAAACATGTTCCCTCAGTCGTCATGCTGACCTGCATTTTCCCTGCTAAATCTGGTTCCTGTAGTTATGCGCCGAAGAA
13▶uGI yLysGI uThr LysLysGI uAsnMetPheLeuGI nSer Ser Me tLeuThr CysI l ePheLeuLeuI l eSer GI ySer CysGI uLeuCysAl aGI uGI u

701 AATTTTTCTAGAAGCTATCCTTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGCAGAGTGCAGCAATCGTGCAGTACAGGAAGTCCCCAAACGGTGG
47▶AsnPheSer ArgSer TyrP roCysAspGI uLysLysGI nAsnAspSer Val I l eAl aGI uCysSerAsnArgArgLeuGI nGI uVal P roGI nThr Val I G
801 GCAAAATATGTGACAGAAGCTAGACCTGTCTGATAATTTTCATCACACACATAACGAATGAATCATTCAAGGGCTGCAAAATCTCACTAAAAATAAATCTAAA
80▶l yLysTyrVal l Thr GI uLeuAspLeuSer AspAsnPheI l eThr Hi s I l eThrAsnGI uSer PheGI nGI yLeuGI nAsnLeuThr LysI l eAsnLeuAs

BsrGI (912) 901 CCACAACCCCAATGTACAGCACCAGAACGGAAATCCCGGTATACAATCAAAATGGCTTGAATATCACAGACGGGCATTCTCAACCTAAAAAACCTAAGG
113▶nHi sAsnP roAsnVal l GI nHi sGI nAsnGI yAsnP roGI y l l eGI nSerAsnGI yLeuAsnI l eThrAspGI yAl aPheLeuAsnLeuLysAsnLeuArg
1001 GAGTTACTGCTTGAAGACAACCCAGTTACCCCAAATACCCCTCTGGTTTGCAGAGTCTTTGACAGAAGCTTAGTCTAATTCAAAACAATATATACAACATAA
147▶GI uLeuLeuLeuGI uAspAsnGI nLeuP roGI nI l eP roSer GI yLeuP roGI uSer LeuThr GI uLeuSer LeuI l eGI nAsnAsnI l eTyrAsnI l eT
1101 CTAAGAGGGCATTTCAGACTTATAAACTTGA AAAATCTCTATTTGGCTGGAAGTCTATTTTAAACAAAGTTTGCAGAAAACCTAACATAGAGATGG
180▶hr LysGI uGI y l l eSer ArgLeuI l eAsnLeuLysAsnLeuTyrLeuAl aTrpAsnCysTyrPheAsnLysVal CysGI uLysThrAsnI l eGI uAspGI
1201 AGTATTTGAAACGCTGACAAATTTGGAGTTGCTATCATTCTTTCAATCTCTTTACATGTCCACCCAACTGCCAAGCTCCCTACGCAAACTTTT
213▶yVal l PheGI uThr LeuThrAsnLeuGI uLeuLeuSer LeuSer PheAsnSer LeuSer Hi sVal P roP roLysLeuP roSer Ser LeuArgLysLeuPhe
1301 CTGAGCAACCCAGATCAAATACATTAGTGAAGAAGATTTCAAGGGATTGATAAATTAACATTACTAGATTTAAGCGGGAAGTGTCCGAGGTGCTTCA
247▶LeuSerAsnThr GI nI l eLysTyrI l eSer GI uGI uAspPheLysGI yLeuI l eAsnLeuThr LeuLeuAspLeuSer GI yAsnCysP roArgCysPheA
1401 ATGCCCATTTCCATGCTGCTTGTGATGGTGGTCTCAATTAATATAGATCGTTTGTCTTTCAAACTTGACCCAACTTCGATACCTAAACCTCTC
280▶snAl aP roPheP roCysVal P roCysAspGI yGI yAl aSer I l eAsnI l eAspArgPheAl aPheGI nAsnLeuThrI nLeuArgTyrLeuAsnLeuSe

EcoRI (1570) 1501 TAGCACTTCCCTCAGGAAGATTAATGCTGCCTGGTTAAAAATATGCCTCATCTGAAGGTGCTGGATCTTGAATCAACTATTTAGTGGGAGAAATAGCC
313▶r Ser Thr Ser LeuArgLysI l eAsnAl aAl aTrpPheLysAsnMe tP roHi sLeuLysVal l LeuAspLeuGI uPheAsnTyrLeuVal l GI yGI u l l eAl a
1601 TCTGGGCAATTTTAAACGATGCTGCCCCGTTAGAAATACTTGACTTGTCTTTAACTATATAAAGGGGAGTTATCCACAGCATATTAATATTTCCAGAA
347▶Ser GI yAl aPheLeuThr Me tLeuP roArgLeuGI u l l eLeuAspLeuSer PheAsnTyrI l eLysGI ySer TyrP roGI nHi sI l eAsnI l eSer ArgA
1701 ACTTCTCTAAACTTTTGTCTCTACGGGCAATGCAATTAAGAGTTATGTGTTCCAGAACTCAGAGAAGATGATTTCCAGCCCCTGATGCAGCTTCCAAA
380▶snPheSer LysLeuLeuSer LeuArgAl aLeuHi sLeuArgGI yTyrVal l PheGI nGI uLeuArgGI uAspAspPheGI nP roLeuMe tGI nLeuP roAs

ClaI (1840) 1801 CTTATCGACTATCAACTTGGGTATTAATTTTATTAAGCAAATCGATTTCAAACTTTTCCAAAATTTCTCCAATCTGGAATATTTACTTGTGAGAAAAC
413▶nLeuSer Thr I l eAsnLeuGI y l l eAsnPheI l eLysGI nI l eAspPheLysLeuPheGI nAsnPheSerAsnLeuGI u l l e l l eTyrLeuSer GI uAsn

BspEI (1967) 1901 AGAATATCACGTTGGTAAAAGATACCCGGCAGAGTTATGCAAATAGTTCTCTTTTCAACGTATATCCGGAACGACGCTCAACAGATTTTGGAGTTG
447▶ArgI l eSer P roLeuVal l LysAspThr ArgGI nSer TyrAl aAsnSer Ser Ser PheGI nArgHi sI l eArgLysArgArgSer Thr AspPheGI uPheA

BstBI (2007) 2001 ACCCACATTCGAACCTTTATCATTTACCCGCTCTTAATAAAGCCACAAATGTGCTGCTATGGAAAAGCCTTAGATTTAAGCCTCAACAGTATTTTCTT
480▶spP roHi sSerAsnPheTyrHi sPheThr ArgP roLeuI l eLysP roGI nCysAl aAl aTyrGI yLysAl aLeuAspLeuSer LeuAsnSer I l ePhePh
2101 CATTGGGCCAAACCAATTTGAAAATCTTCTGACATTGCTGTTAAATCTGTCTGCAAAATAGCAATGCTCAAGTGTTAAGTGGAACTGAATTTTCAGCC
513▶e l l eGI yP roAsnGI nPheGI uAsnLeuP roAspI l eAl aCysLeuAsnLeuSer Al aAsnSerAsnAl aGI nVal l LeuSer GI yThr GI uPheSer Al a

DraIII (2045) 2201 ATTCCTCATGTCAAATATTTGGATTTGACAAAACATAGACTAGACTTTGATAATGCTAGTGTCTTACTGAATGTCCGACTTGAAGTTCTAGATCTCA
547▶I l eP roHi sVal l LysTyrLeuAspLeuThrAsnAsnArgLeuAspPheAspAsnAl aSer Al aLeuThr GI uLeuSerAspLeuGI uVal l LeuAspLeuS
2301 GCTATAATTCACACTATTTGAGAATAGCAGGCGTAACACATCATCTAGAAATTTATCAAAATTCACAAATCTAAAAGTTTTAAACTGAGCCACAACAA
580▶er TyrAsnSer Hi sTyrPheArgI l eAl aGI yVal l Thr Hi sHi sLeuArgI uPheI l eGI nAsnPheThrAsnLeuLysVal l LeuAsnLeuSer Hi sAsnAs
2401 CATTATACTTTAACAGATAAGTATAACCTGAAAAGCAAGTCCCTGGTGAAGATTAAGTTTTCAGTGGCAATCGCCTTGACATTTTGGAAATGATGATGAC
613▶nI l eTyrThr LeuThrAspLysTyrAsnLeuGI uSer LysSer LeuVal l GI uLeuVal l PheSer GI yAsnArgLeuAspI l eLeuTrpAsnAspAsp
2501 AACAGGTATATCTCCATTTCAAAGGTCTCAAGAACTGACACGTCTGGATTTATCCCTTAATAGGCTCAAGCACATCCCAAATGAAGCATTCTTAAT
647▶AsnArgTyrI l eSer I l ePheLysGI yLeuLysAsnLeuThr ArgLeuAspLeuSer LeuAsnArgLeuLysHi sI l eP roAsnGI uAl aPheLeuAsnL

XhoI (2683) 2601 TGCCAGCGAGTCTCACTGAACACTACATATAAATGATAATATGTTAAAGTTTTTAACTGGACATTACTCCAGCAGTTTCTCGTCTCGAGTTGCTTGACTT
680▶euP roAl aSer LeuThr GI uLeuHi sI l eAsnAspAsnMe tLeuLysPhePheAsnTrpThr LeuLeuGI nGI nPheP roArgLeuGI uLeuLeuAspLe
2701 ACGTGGAAACAACTACTCTTTTAACTGATAGCTATCTGACTTTACATCTTCCCTTCGGACTGCTGCTGAGTCATAACAGGATTTCCACCTACCC
713▶uArgGI yAsnLysLeuLeuPheLeuThrAspSer LeuSerAspPheThr Ser Ser LeuArgThr LeuLeuLeuSer Hi sAsnArgI l eSer Hi sLeuP ro
2801 TCTGGCTTTCTTCTGAAGTCAGTAGTCTGAAGCACCTCGATTTAAGTTCCAATCTGCTAAAAACAATAAACAAATCCGCACTTGAACCTAAGACCACCA
747▶Ser GI yPheLeuSer GI uVal l Ser Ser LeuLysHi sLeuAspLeuSer SerAsnLeuLeuLysThr I l eAsnLysSer Al aLeuGI uThr LysThr Thr T

BsaBI (2977)

2901 CCAAATTATCTATGTTGGAACACACGGAAACCCCTTTGAATGCACCTGTGACATTGGAGATTCCGAAGATGGATGGATGAACATCTGAATGTCAAAT
780▶hr LysLeuSer MetLeuGI uLeuHi sGI yAsnProPheGI uCysThr CysAspI l eGI yAspPheArgArgTrpMetAspGI uHi sLeuAsnVal I Lys I l
3001 TCCAGACTGGTAGATGTCATTTGTGCCAGTCCCTGGGATCAAAGAGGGAAGAGTATTGTGAGTCTGGAGCTAACACTTGTGTTTCAGATGTCACCTGCA
813▶eP roArgLeuValI AspVal I l eCysAl aSer P roGI yAspGI nArgGI yLysSer I l eVal Ser LeuGI uLeuThr Thr CysVal SerAspVal Thr Al a

NcoI (3134)

3101 GTGATATTATTTTTCTTCACGTTCTTTATCACCACCATGGTTATGTTGGCTGCCTGGCTCACCATTTGTTTTACTGGGATGTTTGGTTTATATATAATG
847▶Val I l eLeuPhePhePheThr PhePhe I l eThr Thr MetValMetLeuAl aAl aLeuAl aHi sHi sLeuPheTyrTrpAspVal I TrpPhe I l eTyrAsnV
3201 TGTGTTTAGCTAAGGTAAAAGGCTACAGGTCTCTTTCCACATCCCAAACCTTCTATGATGCTTACATTTCTTATGACACCAAAGATGCCTCTGTTACTGA
880▶al CysLeuAl aLysVal I LysGI yTyrArgSer LeuSer Thr Ser GI nThr PheTyrAspAl aTyrI l eSer TyrAspThr LysAspAl aSer Val ThrAs

SmaI (3384)

SandI (3380)

Psp1406I (3349)

EcoO109I (3380) MscI (3392)

3301 CTGGGTGATAAATGAGCTGCGCTACCACCTGAAGAGAGCCGAGACAAAAACGTTCTCCTTTGTCTAGAGGAGAGGGATTGGGACCCGGGATTGGCCATC
913▶pTrpVal I l eAsnGI uLeuArgTyrHi sLeuGI uGI uSer ArgAspLysAsnVal I LeuLeuCysLeuGI uGI uArgAspTrpAspP roGI yLeuAl a l l e
3401 ATCGACAACCTCATGAGAGCATCAACCAAAGCAAGAAAACAGTATTTGTTTTAACAAAAAATGCAAAAAGCTGGAACTTAAAAACGCTTTTTACT
947▶I l eAspAsnLeuMetGI nSer I l eAsnGI nSer LysLysThr Val PheVal I LeuThr LysLysTyrAl aLysSer TrpAsnPheLysThr Al aPheTyrL
3501 TGGCTTTGCAGAGGCTAATGGATGAGAACATGGATGTGATTATTTATCCTGCTGGAGCCAGTGTACAGCATTCTCAGTATTTGAGGCTACGGCAGCG
980▶euAl aLeuGI nArgLeuMetAspGI uAsnMetAspVal I l e l l ePhe I l eLeuLeuGI uP roVal I LeuGI nHi sSer GI nTyrLeuArgLeuArgGI nAr
3601 GATCTGTAAGAGCTCCATCCTCCAGTGGCTGACAACCCGAAGGCAGAAGGCTTGTGTTGGCAAACCTCTGAGAAATGTGGTCTTGACTGAAAATGATTCA
1013▶gl I eCysLysSer Ser I l eLeuGI nTrpP roAspAsnP roLysAl aGI uGI yLeuPheTrpGI nThr LeuArgAsnVal I Val I LeuThr GI uAsnAspSer

MscI (3763)

NheI (3757)

3701 CGGTATAACAATATGTATGTCGATTCCATTAAGCAATACTAACTGACGTTAAGTCATGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTGGAC
1047▶ArgTyrAsnAsnMetTyrValI AspSer I l eLysGI nTyr●●●

HpaI (3895)

3801 AAACCACAAC TAGAATGCAGTGA AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAA

MfeI (3906)

EcoRI (3991)

3901 CAACAACAATTGCATTCATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATCTAA

4001 AATACAGCATAGCAAACCTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGT

4101 GCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAAGTACTGCTCTTCATTTCTTTATGTTTTAAATG

SwaI (4244)

4201 CACTGACCTCCACATTCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGAT

EcoO109I (4305)

4301 GCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCT

4401 TTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCA
141▶●●●AsnArgThr TyrLysLeuP roI l eLeuGI uGI u l l eThr Thr LysVal I LeuLysGI yAsnMetGI u l l eLeuVal I PheCysAspP roAl aTyrAspS

BstXI (4534)

4501 GAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATGTGCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGT
107▶er I l eLeuGI uArgCysMetGI yCysP roSer Val Val ArgI l eSer ArgAspVal GI uAspSer TyrP roHi sArgVal Al aVal I l eThr AspPheAs

StuI (4669)

4601 CCTTCTGCCGTTGCTCAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAGATGATCTC
74▶pLysGI nGI yAsnSer Val Al aSer GI y l l eAl a l l eAl aGI uAl aCysVal Thr Val ArgGI y l l eTyrAl aGI u l l eHi sVal Al aSer I l e l l eGI u
4701 CCCAGTCTGGTCTGATGGCCGCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACAGCTCCAGATCTCCG
41▶GI yThr LysThr ArgI l eAl aAl aGI yVal Hi sHi sLysAsnAspGI uTyrLeuMetThr I l eLysGI uThr Al aVal GI uVal I LeuGI uLeuAspGI nG

BspHI (4819)

4801 TGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGCATGATTATACTATGCGGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGGA
7▶I nSer I l eAsnPheThr LysMet

4901 TGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGT

SpeI (5032)

5001 TGTTACGACATTTTGAAAGTCCCCTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACC

SnaBI (5160)

5100 GCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGGAAAGTCCCATAAGGT

NdeI (5265)

5200 CATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTAAGTGGCA

5300 GTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGAAACATACGTCAATATTGACGTCAATGGCGGGGGTCTG

SdaI (5443)PacI (5451)

5400 TGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCCTGCAGGTTAA TTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACC

5498 GTAAAAAGGCCGCTTGTCTGGCCTTTTCCATAGGCTCCGCCCTGACGAGCATCAGAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG

5598 ACTATAAAGATACAGGCGTTTTCCCTGGAAGCTCCCTGCTGCGCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCTTTCTCCCTTCG

ApaI (5775)

5698 GGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTTCAGC
5798 CCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG
5898 AGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA
5998 GTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTGTGGCAAGCAGCAGATTACGCGCAGAA

PacI (6191)

6098 AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATT

EagI (6211)

Swal (6200) **NotI (6210)**

6198 AACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAA
6298 ACAAACGAAACAAACAAACTAGCAAATAGGCTGTCCCAAGTGCAAGTGCCAGTGCCAGAACATTTCTCTATCGAA