



PvuI (7)
SgfI (6)
EcoNI (96)

1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203)
HindIII (245)
Bsu36I (291)

201 GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTACAGCGCCCGCCGCTACCTGAGGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCCTGTGGTGCCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMI (441)
NaeI (441)

401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGTCTTTGTTTCGTTT

KasI (535)
AgeI (552)
BspLU11I (560)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCCTACCTGAGATCACCGGTCAACATGTGCAATGGGTATTCCACAGACGAGAATTTCCGCTA
1▶MetSerAsnGlyTyrSerThrAspGluAsnPheArgTy

BsrGI (632)

601 TCTCATCTCGTGCTTTCAGGGCCAGGGTAAAAATGTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAGCAGATT
13▶rLeuIleSerCysPheArgAlaArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAlaGluValLysGluGlnIle
701 CAGAGGACAGTCCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGCTGAGCACCTTGGAGAAGGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCG
47▶GlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGluLeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheV

BspEI (809)
BsiBI (895)
BsaBI (895)

801 TGGAGGCCCTCCGAGAACCGGCAGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCACGGACTTGCCTCTCCATCGTTTGAACGCTCATGATGA
80▶alGluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThrAspLeuProSerProSerPheGluAsnAlaHisAspGlu

HindIII (943)

901 ATATCTCCAAGTCTGACCTCCTTCAGCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCATGGAGGAGGAAGTGTGACAATT
113▶uTyrLeuGlnLeuLeuAsnLeuLeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGluGluGluLeuLeuThrIle
1001 GAAGACAGAAACCGGATTGCTGCTGCAGAAAACATGAAATGAATCAGGTGTAAAGAGAGCTACTAAAAAGGATTGTGCAGAAAAGAACTGTTCTCTG
147▶GluAspArgAsnArgIleAlaAlaAlaGluAsnAsnGlyAsnGluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerA

HpaI (1149)

1101 CATTCTGAATGTTCTTCTCGTCAAACAGGAAACAATGAACCTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAAAACAATGCAGAGATTGAGAATTTATC
180▶IlePheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySerAspCysSerGluSerAsnAlaGluIleGluAsnLeuSe
1201 ACAAGTTGATGGTCTCAAGTGGAGAGCAACTCTTTCAACCACAGTTACAGCAAATCTGGAGAAGGAGGTCTGGGGCATGGAGAATAACTCATCAGAA
213▶rGlnValAspGlyProGlnValGluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMetGluAsnAsnSerSerGlu

PvuII (1362)

1301 TCATCTTTTGCAGATTCTTCTGTAGTTTCAGAATCAGACACAAGTTTGGCAGAAGGAAGTGTGAGCTGCTTAGATGAAAGTCTTGGACATAACAGCAACA
247▶SerSerPheAlaAspSerSerValValSerGluSerAspThrSerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnM

XmnI (1424)
StuI (1486)

1401 TGGGCAGTGATTCAGGCACCATGGGAAGTGATTCAGATGAAGAGAATGTGGCAGCAAGAGCATCCCCGGAGCCAGAATCCAGCTCAGGCCTTACCAAT
280▶etGlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAlaSerProGluProGluLeuGluLeuArgProTyrGlnMe
1501 GGAAGTTGCCAGCCAGCCTTGGAAAGGGAAGATATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACAGAGTGGCTGTTTACATTGCCAAGGATCAC
313▶tGluValAlaGlnProAlaLeuGluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIleAlaLysAspHis
1601 TTAGACAAGAAAGAAAAGCATCTGAGCCTGAAAAGTTATAGTTCTTGTCAATAAGGTAAGTACTGCTAGTTGAACAGCTCTCCGCAAGGAGTCCCAACAT
347▶LeuAspLysLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProP
1701 TTTTGAAGAAATGGTATCGTGTTATTGGATTAAGTGGTATACCAACTGAAAATATCATTTCCAGAAGTTGTCAAGTCTGTGATATTATTATCAGTAC
380▶heLeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPheProGluValValLysSerCysAspIleIleIleSerTh
1801 AGCTCAAATCCTTGAACCTCCCTTAAACTTGGAAAATGGAGAAGTGTGGTGTCAATTGTGACTTTTCCCTCATTATCATTGATGAATGTCAT
413▶rAlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIleIleAspGluCysHis
1901 CACACCAACAAAGAAGCAGTGTATAATAACATCATGAGGCATTATTTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAACAACAGTGATTC
447▶HisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIleP

DraIII (2030)

2001 CCCTTCTCAGATACTGGGACTAACAGCTTACCTGGTGTGGAGGGCCACGAAGCAAGCAAGCTGAAGAACACATTTTAAACTATGTGCCAATCT
480▶rLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLe

Ppu10I (2102)
NsiI (2102)

2101 TGATGCATTTACTATTAATAACTGTTAAAGAAAACCTTGATCAACTGAAAACCAAATACAGGAGCCATGCAAGAAGTTTGCATTGCAGATGCAACCAGA
513▶uAspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThrArg

XbaI (2223)

2201 GAAGATCCATTTAAAGAGAAACTCTAGAAAATAATGACAAGGATTCAAACCTATTGTCAAATGAGTCCAAATGTGAGATTTTGGAACTCAACCCATGAAC
547▶GluAspProPheLysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluG

AseI (2398)

2301 AATGGGCCATTCAAATGGAAAAAAGCTGCAAAAAGAGGAAATCGCAAAAAGACGTGTTTGTGCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAAT
580▶InTrpAlaIleGlnMetGluLysLysAlaAlaLysGluGlyAsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle

BstBI (2411)
AsuII (2411)
Bst1107I (2427)

2401 TAATGACACAATTCGAATGATAGATGCGTATACTCATCTTGAACCTTTCTATAATGAAGAGAAAAGATAAGAAGTTTGCAGTCATAGAAGATGATAGTGAT
613▶eAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspAspSerAsp
2501 GAGGGTGGTGTGATGATGAGTATTGTGATGGTGTGAAAGATGAGGATGATTTAAAGAAAACCTTTGAAACTGGATGAACAGATAGATTTCTCATGACTTTAT
647▶GluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuP

2601 TTTTGGAAAACAATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAATATGAAAATGAAAAGCTGACCAAATTAAGAAATACCATAATGGAGCAATATAC
680▶hePheGI uAsnAsnLysMe tLeuLysArgLeuAl aGI uAsnP roGI uTyrGI uAsnGI uLysLeuThr LysLeuArgAsnThr I l eMe tGI uGI nTyrTh
NdeI (2753)

2701 TAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGATATGCGCTTCCAGTGGATTACTGAAAATGAAAAATTTGCTGAA
713▶r ArgThr GI uGI uSer Al aArgGI yI l eI l ePheThr LysThrArgGI nSer Al aTyrAl aLeuSer GI nT rpl l eThr GI uAsnGI uLysPheAl aGI u
XcmI (2814)

2801 GTAGGAGTCAAAGCCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCATGACACAGAATGAACAAAAAGAGTCATTAGTAAATTTGCGCA
747▶Val GI yVal LysAl aHi sHi sLeuI l eGI yAl aGI yHi sSer Ser GI uPheLysP roMe tThr GI nAsnGI uGI nLysGI uVal I l eSer LysPheArgT
2901 CTGGAAAAATAAATCTGCTTATCGCTACACAGTGGCAGAAAGGCTGGATATTAAGAATGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGA
780▶hr GI yLysI l eAsnLeuLeuI l eAl aThr Thr Val Al aGI uGI uGI yLeuAspI l eLysGI uCysAsnI l eVal I l eArgTyrGI yLeuVal ThrAsnGI
NcoI (3005)

3001 AATAGCCATGGTCCAGGCCCGTGGTTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAGACAGTT
813▶ul l eAl aMe tVal GI nAl aArgGI yArgAl aArgAl aAspGI uSer Thr TyrVal l eLeuVal Al aHi sSer GI ySer GI yVal I l eGI uHi sGI uThr Val
3101 AATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGTGTTCAAATATGAAACAGAGGAGTATGCTCATAAGATTTTGAATTACAGATGCAAA
847▶AsnAspPheArgGI uLysMe tMe tTyrLysAl a l eHi sCysVal GI nAsnMe tLysP roGI uGI uTyrAl aHi sLysI l eLeuGI uLeuGI nMe tGI nS
SspI (3232) NheI (3297)

3201 GTATAATGGAAAAGAAAATGAAAACCAAGAGAAAATATTGCCAAGCATTACAAGAATAACCCATCACTAATAACTTTCTTTGCAAAAACTGCAGTGTGCT
880▶er I l eMe tGI uLysLysMe tLysThr LysArgAsnI l eAl aLysHi sTyrLysAsnAsnP roSer LeuI l eThr PheLeuCysLysAsnCysSer Val Le
EcoRV (3316) Ppu10I (3337)
Eco32I (3316) NsiI (3337)

3301 AGCCTGTTCTGGGGAAGATATCCATGTAATTGAGAAAATGCATCACGTCAATATGACCCAGAATTCAGGAACCTTACATTGTAAGAGAAAACAAAGCA
913▶uAl aCysSer GI yGI uAspI l eHi sVal I l eGI uLysMe tHi sHi sVal l eAsnMe tThr P roGI uPheLysGI uLeuTyrI l eVal l eArgGI uAsnLysAl a
MscI (3453) Ball (3453) ApaLI (3478)

3401 CTGCAAAAGAAAGTGTGCCACTATCAAATAAATGGTGAATCATCTGCAAATGTGGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGC
947▶LeuGI nLysLysCysAl aAspTyrGI nI l eAsnGI yGI uI l eI l eCysLysCysGI yGI nAl aT rpl GI yThr Me tMe tVal l eHi sLysGI yLeuAspLeuP
3501 CTTGTCTCAAATAAGGAATTTGTAGTGGTTTTCAAATAAATTCAACAAAGAAACAATACAAAAAGTGGGTAGAATTACCTATCACATTTCCAATCT
980▶r oCysLeuLysI l eArgAsnPheVal Val l ePheLysAsnAsnSer Thr LysLysGI nTyrLysLysT rpl Val GI uLeuP roI l eThr PheP roAsnLe
MscI (3666) Ball (3666) AvrII (3654)

3601 TGACTATTAGAATGCTGTTTATTAGTGATGAGGATTAGCACTTGATTGAAGACCTAGGACTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTG
1013▶uAspTyrSer GI uCysCysLeuPheSerAspGI uAsp●●●

3701 GACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAACAAGT
HpaI (3798)

3801 TAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTC
3901 TAAAATACAGCATAGCAAACTTAACTCCAAATCAAGCCTCTACTGAACTCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAA
4001 TGTGCATTAGCTGTTTGCAGCCTCACCTCTTTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGGTTGAACTAGCTCTTCATTTCTTTATGTTTTAA
4101 ATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCA
4201 GATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAGGAACCTTTAATAGAATTTGGACAGCAAGAAAGCGAGCTTCTA
4301 GCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCTCAATGAGCACAAAGCAGTCAGGAGCATAG
141▶●●●AsnArgThr TyrLysLeuP roI l eLeuGI uGI uI l eThr Thr LysVal l eLeuLysGI yAsnMe tGI uI l eLeuVal l ePheCysAspP roAl aTyrA
4401 TCAGAGATGAGCTCTCGACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAA
108▶spSer I l eLeuGI uArgCysMe tGI yCysP roSer Val Val A rgl l eSer ArgAspVal GI uAspSer TyrP roHi sArgVal Al aVal I l eThr AspPh
StuI (4572) Eco147I (4572)

4501 AGTCCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAGATGAT
75▶eAspLysGI nGI yAsnSer Val Al aSer GI yI l eAl aI l eAl aGI uAl aCysVal Thr Val l eArgGI yI l eTyrAl aGI uI l eHi sVal Al aSer I l eI l e
4601 CTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGTGTTGTCCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCC
42▶GI uGI yThr LysThrArgI l eAl aAl aGI yVal l eHi sHi sLysAsnAspGI uTyrLeuMe tThr I l eLysGI uThr Al aVal GI uVal l eLeuGI uLeuAspG
XmnI (4714) AseI (4780)

4701 TGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAAACAGCGT
8▶l nGI nSer I l eAsnPheThr LysMe t
4801 GGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCCAATTTGCGTCAATGGGGCGG

4901 AGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCCATTTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAA
SpeI (4935)

5000 ACCGCTATCCAGCCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAA
SnaBI (5063) Eco105I (5063)

5100 GGTCATGACTGGGCATAATGCCAGGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGACTGCCAAGTGG
NdeI (5168)

5200 GCAGTTTACGTAATACTCCACCCATTGACGTC AATGGAAAGTCCCTATTGGCGT TACTATGGGAACATACGTCATTATTGACGTC AATGGGCGGGGT

5300 CGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCC T G C A G G T T A A T T A A G A A C A T G T G A G C A A A A G G C C A G C A A A A G G C C A G G A

5398 ACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATC A C A A A A A T C G A C G C T C A A G T C A G A G G T G G C G A A A C C C G A C

5498 AGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCT

5598 TCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTT C

5698 AGCCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAG

5798 CAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAG

5898 CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCA

5998 GAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAAC T C A C G T T A A G G G A T T T T G G T C A T G G C T A G T T A

PacI (6094)

EagI (6114)

SwaI (6103) **NotI (6113)**

6098 ATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATC

6198 AAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA