



PvuI (7)
SgfI (6)
MfeI (82)

1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203)
HindIII (245)

201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGGTTGAGTCGCGTTCTGCGCCCTCCCGCTGTGGTGCCTCCTGAAGCTCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGCTCAACTCTACGCTTTGTTTCGTTT

NcoI (560)
BstEII (555)

501 TCTGTTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCCTAAGCTGAGATCACCGGTACCATGGTCCCTGCTGAAACCATGGCAACATCACCCGCTC
601 CAAGCGGAGGAGCTGCTTCCAGGACAGGCAAGGACGGGAGCTTCTCGTGCCTGCCAGCGAGTCCATCTCCGGGCATACGCGCTCTGCGTGTGTAT
130 r LysAl aGl uGl uLeuLeuSer ArgThr Gl yLysAspGl ySer PheLeuVal ArgAl aSer Gl uSer I l eSer ArgAl aTyrAl aLeuCysVal LeuTyr
701 CGGAATTGCGTTTACACTTACAGAATTCTGCCAATGAAGATGATAAATCTACTGTTCAAGCATCCGAAGCGCTCTCCATGAGTTTCTCACCAAGCTGG
470 ArgAsnCysVal TyrThr TyrArgI l eLeuProAsnGl uAspAspLysPheThr Val Gl nAl aSer Gl uGl yVal Ser MetArgPhePheThr LysLeuA
801 ACCAGTCTCATCGAGTTTACAAGAAGGAAACATGGGGCTGGTACCCATCTGCAATACCCTGTGCCGCTGGAGGAAGAGGACACAGGCGACGCCCTGA
800 spGl nLeuI l eGl uPheTyrLysLysGl uAsnMetGl yLeuVal Thr Hi sLeuGl nTyrProVal P roLeuGl uGl uGl uAspThr Gl yAspAspProGl
901 GGAGGACACAGAAAGTGTGCTGTCTCCACCCGAGCTGCCCCAAAGAAACATCCCGCTGACTGCCAGCTCCTGTGAGGCCAAGGAGGTTCTTTTTCAAAC
1130 uGl uAspThr Gl uSer Val Val Ser ProP roGl uLeuP roP roArgAsnI l eP roLeuThr Al aSer Ser CysGl uAl aLysGl uVal P roPheSerAsn
1001 GAGAAATCCCGAGCGACCAGACAGCCGGCCGAGCCTCTCCGAGACATTGTTCCAGCGACTGCAAAGCATGGACACCAGTGGCTTCCAGAAGAGCATC
1470 Gl uAsnP roArgAl aThr Gl uThr Ser ArgP roSer LeuSer Gl uThr LeuPheGl nArgLeuGl nSer MetAspThr Ser Gl yLeuP roGl uGl uHi sL
1101 TTAAGGCCATCCAAGATTATTAAGCACTCAGCTCGCCAGGACTCTGAATTTGTGAAGACAGGGTCCAGAGTCTTCCCTCACCTGAAGAAACTGACCCAC
1800 eulysAl a l eGl nAspTyrLeuSer Thr Gl nLeuAl aGl nAspSer Gl uPheVal LysThr Gl ySer Ser Ser LeuP roHi sLeuLysLysLeuThr Th
EcoRI (722)

1201 ACTGCTCTGCAAGGAGCTCTATGGAGAAGTCATCCGGACCCTCCCATCCCTGGAGTCTCTGCAGAGGTTATTTGACCAGCAGCTCTCCCCGGGCTCCGT
2130 r LeuLeuCysLysGl uLeuTyrGl yGl uVal I l eArgThr LeuP roSer LeuGl uSer LeuGl nArgLeuPheAspGl nGl nLeuSer P roGl yLeuArg
BstEII (840)

1301 CCACGTCCTCAGGTTCTGGTGAGGCAATCCCATCAACATGGTGTCCAAGCTCAGCCAAGTACAAGCTGTTGTGCTCCATTGAAGACAAGGTCAAGG
2470 P roArgP roGl nVal P roGl yGl uAl aAsnP roI l eAsnMetVal Ser LysLeuSer Gl nLeuThr Ser LeuLeuSer Ser I l eGl uAspLysVal LysA
1401 CCTTGCTGCACGAGGGTCTGAGTCTCCGACCCGGCCCTCCCTATCCCTCCAGTACCTTTGAGGTGAAGGCAGAGTCTCTGGGATTCCTCAGAAAAT
2800 I l eLeuLeuHi sGl uGl uP roGl uSer P roHi sArgP roSer LeuI l eP roP roVal Thr PheGl uVal LysAl aGl uSer LeuGl yI l eP roGl nLysMe
SaI (1510)

1501 GCAGCTCAAAGTGCAGCTTGAGTCTGGGAAACTGATCATTAAAGAGTCCAAGGATGGTCTGAGGACAAGTCTACAGCCACAAGAAAATCCTGCAGCTG
3130 tGl nLeuLysVal AspVal Gl uSer Gl yLysLeuI l eI l eLysLysSer LysAspGl ySer Gl uAspLysPheTyrSer Hi sLysLysI l eLeuGl nLeu
1601 ATTAAGTACAGAAATTTCTGAATAAGTTGGTATCTTGGTGGAAACAGAGAAGGAGAAGATCCTGCGGAAGGAATATGTTTTGCTGACTCCAAAAGA
3470 I l eLysSer Gl nLysPheLeuAsnLysLeuVal I l eLeuVal Gl uThr Gl uLysGl uLysI l eLeuArgLysGl uTyrVal PheAl aAspSer LysLysA
BsaBI (1767)

1701 GAGAAGGCTTCTGCCAGCTCCTGCAGCAGATGAAGAACAAGCACTCAGAGCAGCCGGAGCCGACATGATCACCATCTTCATCGGCACCTGGAACATGGG
3800 r gGl uGl yPheCysGl nLeuLeuGl nGl nMetLysAsnLysHi sSer Gl uGl nP roGl uP roAspMetI l eThr I l ePheI l eGl yThr TrpAsnMetGl
1801 TAACGCCCCCTCCCAAGAAGTACAGTCTCTGGTTCTCTCCAAGGGCAGGAAAGACGCGGGACGACTCTGCGGACTACATCCCCATGACATTTAC
4130 yAsnAl aP roP roLysLysI l eThr Ser TrpPheLeuSer LysGl yGl nGl yLysThr ArgAspAspSer Al aAspTyrI l eP roHi sAspI l eTyr
1901 GTGATCGGCACCAAGAGGACCCCTGAGTGAAGAGGAGTGGTGGAGATCCTCAAACACTCCCTGCAAGAAATCACCAGTGTGACTTTTAAACAGTGC
4470 Val I l eGl yThr Gl nGl uAspP roLeuSer Gl uLysGl uTrpLeuGl uI l eLeuLysHi sSer LeuGl nGl uI l eThr Ser Val Thr PheLysThr ValA
MscI (2036)

2001 CCATCCACACGCTCTGGAACATCCGCATCGTGGTCTGGCCAAAGCTGAGCAGGAGAACGGATCAGCCACATCTGTACTGACAACGTGAAGACAGGCAT
4800 I al l eHi sThr LeuTrpAsnI l eArgI l eVal Val LeuAl aLysP roGl uHi sGl uAsnArgI l eSer Hi sI l eCysThr AspAsnVal LysThr Gl yI l
2101 TGCAAAACACTGGGGAACAAGGAGCCGTGGGGGTGCTGTTTCATGTTCAATGGAACCTCCTTAGGGTTCGTAACAGCCACTGACTTCAGGAAGTAA
5130 eAl aAsnThr LeuGl yAsnLysGl yAl aVal Gl yVal Ser PheMetPheAsnGl yThr Ser LeuGl yPheVal AsnSer Hi sLeuThr Ser Gl ySer Gl u
2201 AAGAAACTCAGGCAAAACAAAATATATGAACATTCTCCGGTCTCGGCCCTGGGCGACAAGAAGCTGAGTCCCTTTAACATCACTACCCGCTTACGC
5470 LysLysLeuArgArgAsnGl nAsnTyrMetAsnI l eLeuArgPheLeuAl aLeuGl yAspLysLysLeuSer P roPheAsnI l eThr Hi sArgPheThrH
2301 ACCTCTTCTGGTTTGGGATCTTAACACTACCGTGTGGATCTGCCTACCTGGGAGGCAGAAACCATCATCCAGAAAATCAAGCAGCAGCAGTACCGAGACCT
5800 i sLeuPheTrpPheGl yAspLeuAsnTyrArgVal AspLeuP roThr TrpGl uAl aGl uThr I l eI l eGl nLysI l eLysGl nGl nGl nTyrAl aAspLe
XmnI (2439)

2401 CCTGTCCCACGACAGCTGCTCACAGAGAGGAGGAGCAGAAGTCTTCTACACTTCGAGGAGGAAGAAATCACGTTTGCCCCAACCTACCGTTTTGAG
6130 uLeuSer Hi sAspGl nLeuLeuThr Gl uArgArgGl uGl nLysVal PheLeuHi sPheGl uGl uGl uI l eThr PheAl aP roThr TyrArgPheGl u
2501 AGACTGACTCGGGACAAATACGCCTACACCAAGCAGAAAGCAGAGGATGAAGTACAACCTGCCTTCTGGTGTGACCGAGTCTCTGGAAGTCTTATC
6470 ArgLeuThr ArgAspLysTyrAl aTyrThr LysGl nLysAl aThr Gl yMetLysTyrAsnLeuP roSer TrpCysAspArgVal LeuTrpLysSer TyrP
DraIII (2608)

2601 CCCTGGTGACGTGGTGTGTCAGTCTTATGGCAGTACCAGCGACATCATGACGAGTGACCACAGCCCTGTCTTTGCCACATTTGAGGCAGGAGTCACTTC
6800 r oLeuVal Hi sVal Val CysGl nSer TyrGl ySer Thr SerAspI l eMetThr SerAspHi sSer P roVal PheAl aThr PheGl uAl aGl yVal Thr Se
BspHI (2645)

2701 CCAGTTTGTCTCCAAGAACGGTCCCGGACTGTTGACAGCCAAGGACAGATTGAGTTTCTCAGGTGCTATGCCACATTGAAGACCAAGTCCAGACAAA
7130 r Gl nPheVal Ser LysAsnGl yProGl yThr Val AspSer Gl nGl yGl nI l eGl uPheLeuArgCysTyrAl aThr LeuLysThr LysSer Gl nThr Lys
XcmI (2795)

XhoI (2817)

2801 TTCTACCTGGAGTTCACCTCGAGCTGCTTGGAGAGTTTTGTCAAGAGTCAGGAAGGAGAAAATGAAGAAGGAAGTGAGGGGGAGCTGGTGGTGAAGTTTG
747▶PheTyrLeuGI uPheHi sSer Ser CysLeuGI uSer PheVal I LysSer GI nGI uGI yGI uAsnGI uGI uGI ySer GI uGI yGI uLeuVal I Val LysPheG
2901 GTGAGACTCTTCCAAAGCTGAAGCCATTATCTCTGACCCCTGAGTACCTGTAGACCAGCACATCCTCATCAGCATCAAGTCCCTGACAGCCGACGAATC
780▶IyGI uThrLeuP roLysLeuLysP roI l eI l eSerAspP roGI uTyrLeuLeuAspGI nHi s l l eLeu l l eSer l l eLysSer SerAspSerAspGI uSe

NcoI (3075)

3001 CTATGGCGAGGGCTGCATTGCCCTTCGGTTAGAGGCCACAGAAACGCAGCTGCCATCTACACGCCCTCTACCCACCATTGGGGAGTTGACAGGCCACTTC
813▶r TyrGI yGI uGI yCysI l eAl aLeuArgLeuGI uAl aThr GI uThr GI nLeuP roI l eTyrThr P roLeuThr Hi sHi sGI yGI uLeuThr GI yHi sPhe
3101 CAGGGGGAGATCAAGCTGCAGACCTCTCAGGGCAAGACGAGGGAGAAGCTCTATGACTTTGTGAAGACGGAGCGTGATGAATCCAAGTGGGCCAAAGACCC
847▶GI nGI yGI u l l eLysLeuGI nThr Ser GI nGI yLysThr ArgGI uLysLeuTyrAspPheVal I LysThr GI uArgAspGI uSer Ser GI yProLysThr L
3201 TGAAGAGCCTCACCAGCCACGACCCATGAAGCAGTGGGAAGTCACTAGCAGGGCCCCCTCCGTGACGTGGCTCCAGCATCACTGAAATCATCAACCCCAA
880▶euLysSer LeuThr Ser Hi sAspP roMeT LysGI nTrpGI uVal Thr Ser ArgAl aP roP roCysSer GI ySer Ser l l eThr GI u l l eI l eAsnP roAs

SfiI (3314)

3301 CTACATGGGAGTGGGGCCCTTGGGCCACCAATGCCCTGCACGTGAAGCAGACCTTGTCCCTGACCAGCAGCCACAGCCTGGAGCTACGACCAGCCG
913▶nTyrMetGI yVal GI yProPheGI yProP roMeT P roLeuHi sVal I LysGI nThr LeuSer P roAspGI nGI nP roThr Al aTrpSer TyrAspGI nP ro

SdaI (3422)

3401 CCCAAGGACTCCCCGCTGGGGCCCTGCAGGGGAGAAAAGTCTCCGACACCTCCCGGCCAGCCGCCATATCACCCAAGAAGTTTTTACCCTCAACAGCAA
947▶P roLysAspSer P roLeuGI yP roCysArgGI yGI uSer P roP roThr P roP roGI yGI nP roP roI l eSer P roLysLysPheLeuP roSer Thr Al aA
3501 ACCGGGGTCTCCCTCCAGGACACAGGAGTCAAGGCCAGTGCCTGGGAAGAACGCAGGGGACACGCTGCCTCAGGAGGACCTGCCCTGACGAAGCC
980▶snArgGI yLeuP roP roArgThr GI nGI uSer ArgP roSerAspLeuGI yLysAsnAl aGI yAspThr LeuP roGI nGI uAspLeuP roLeuThr LysP r

SmaI (3625)

SacII (3685)

3601 CGAGATGTTTGAGAACCCCTGTATGGGTCCCTGAGTTCCTTCCCTAAGCCTGCTCCAGGAAGGACCAGGAATCCCCAAAATGCCCGGAAGAACCC
1013▶oGI uMetPheGI uAsnP roLeuTyrGI ySer LeuSer Ser PheP roLysP roAl aP roArgLysAspGI nGI uSer P roLysMetP roArgLysGI uP ro
3701 CCGCCTGCCCGAACCCGGCATTTGTGCGCCAGCATCTGCTACCAAAGCCAGGAGGCTGATCGCGCGAGGGGCCGGCAAGCAGGTGCCCGCGC
1047▶P roP roCysP roGI uP roGI y l l eLeuSer P roSer l l eVal LeuThr LysAl aGI nGI uAl aAspArgGI yGI uGI yP roGI yLysGI nVal I P roAl aP

PshAI (3879)

3801 CCCGGCTGCGCTCCTTACGTGCTCATCTCTGCGAGGGCAGGGCGCGGGGACAAGAGCCAAGGAAGCCCAAGACCCGGTACAGCTCCAGGC
1080▶r oArgLeuArgSer PheThr CysSer Ser Ser Al aGI uGI yA rgAl aAl aGI yGI yAspLysSer GI nGI yLysP roLysThr P roVal Ser Ser GI nAl

NotI (3975)

SacII (3973)

3901 CCCGGTCCCGCCAAAGAGGCCCATCAAGCCTTCCAGATCGAAATCAACCAGCAGACCCGCCACCCCGACGCCGGCCGCGCGCTGCCAGTCAAGAGC
1113▶aP roVal P roAl aLysArgP roI l eLysP roSer ArgSer GI u l l eAsnGI nGI nThr P roP roThr P roThr P roArgP roP roLeuP roVal LysSer

SfiI (4091)

4001 CCGCGGTGCTGCACCTCCAGCACTCCAAGGGCCGCGACTACCCGACAAACACCGAGCTCCCGCATCACGGCAAGCACCGGCCGAGGAGGGGCCACCA
1147▶P roAl aVal LeuHi sLeuGI nHi sSer LysGI yArgAspTyrArgAspAsnThr GI uLeuP roHi sHi sGI yLysHi sArgP roGI uGI uGI yP roP roG
4101 GGCCTTAGGCAGGACTGCCATGCAGTGAAGCCTTCAAGTGCAGTGCCTGAGTGGGAGCCAGAGGAACGGGTGAAGCCACTGGACCCTCTCCCGGG
1180▶IyP roLeuGI yA rgThr Al aMetGI n●●●

MscI (4250)

NheI (4244)

4201 ACCTCCTGCTGGCTCCTCCTGCCAGCTTCTATGCAAGGCTTGTCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAG

HpaI (4382) MfeI (4393)

4301 AATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAAACAACAAATTGC

EcoRI (4478)

4401 ATTCATTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAAATCTAAAATACAGCATAGC

4501 AAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTT

4601 GCAGCCTCACCTTCTTTCATGGAGTTTAAGATATAGTGTATTTTCCAAAGTTTGAAGTACTGCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCA

SspI (4717)

SwaI (4731)

4701 CATTCCCTTTTTAGTAAAATATTAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTT

4801 CATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTTGTTG

141◀●●●AsnArgThr T

4901 TACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTC

136◀yrLysLeuP roI l eLeuGI uGI u l l eThr Thr LysVal I LeuLysGI yAsnMetGI u l l eLeuVal I PheCysAspP roAl aTyrAspSer l l eLeuGI uAr

5001 TGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTT

103◀gCysMetGI yCysP roSer Val Val ArgI l eSer ArgAspVal I GI uAspSer TyrP roHi sArgValAl aVal l l eThrAspPheAspLysGI nGI yAsn

StuI (5156)

5101 GCTCACAGCAGACCAATGGCAATGGCTTACGACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTGCTC

70◀Ser ValAl 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 sValAl aSer l l e l l eGI uGI yThr LysThr A

XmnI (5298)

5201 CTGATGGCCGCCCGACATGGTCTTGTTCCTCATAGAGCATGGTATCTTCTCAGTGGCAGCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGA

36◀r gl l eAl aAl aGI yVal Hi sHi sLysAsnAspGI uTyrLeuMetThr l l eLysGI uThr Al aVal I GI uVal I LeuGI uLeuAspGI nGI nSer l l eAsnPh

BspHI (5306)

AseI (5364)

5301 AGGCTTTCATGATGGCCCTCTATAGTGCAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGC

3◀eThrLysMet

5401 T TATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTGCGTCAATGGGGCGGAGTTGTTACGACATTT

5501 TGGAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCCATTTGACGTCAATGGGGTGGAGACTTGAAAATCCCGTGAGTCAAACCGCTATCCACGCC

SpeI (5519) ←

5600 ATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCA

SnaBI (5647)

5700 TAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTGTACTGCCAAGTGGGCAGTTTACCGTAAAT

NdeI (5752)

5800 ACTCCACCCATTGACGTCAATGGAAAGTCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGCC

5900 AGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAA TTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGC

SdaI (5930) PacI (5938) BspLU11I (5948) ←

5998 GTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATA

6098 CAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGCGC

6198 TTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGC

ApaLI (6262)

6298 CTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTA

6398 GCGGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA

6498 AAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCA

6598 AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC A

PacI (6678) SmaI (6687)

6698 GCGGCCGAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACA

NotI (6697)

6798 AAACAAACTAGCAAAATAGGCTGTCCCCAGTGAAGTGCAAGTGCCAGGTCAGAACATTTCTCTATCGAA