



PstI (6)
SdaI (6)

1 CCTGCAGGGCCCACTAGTGGAGCCGAGAGTAATTCATACAAAAGGAGGGGATCGCCTTCGCAAGGGGAGAGCCAGGGACCCTCCCTAAATCTCACAGAC
101 CCAAATCCCTGTAGCCGCCCCACGACAGCGCGAGGAGCATGCGCTCAGGGCTGAGCGGGAGAGCAGAGCACACAAGCTCATAGACCCTGGTCGTGGG
201 GGGGAGGACCGGGGAGCTGGCGCGGGGCAAACCTGGGAAAGCGGTGCTGCTGCTGCTCCGCCCTCTTCCCAGGGTGGGGGAGAACGGTATATAAGTGC
301 GGCAGTGCCTTGGACGTTCTTTTTCGCAACGGTTTGGCGTCAGAACGCAggtgagggggcggtgtggcttccgcgggcccgcgagctggaggctctg
401 tccgagcggggcgggccccgctgtcgtcgccggggttagctgtcgagcattcccgcttcgagttcgggcgccgcgggagcagagctgcagagcctagc
501 gcaaccccgtagcctcgcctcgtgtccggccttaggccttagcgtggtgtccgcgcggccgcgctgactccggccgactctggtcttttttttttt
601 tgttgtttgctccctgctgcttgcgattgcccgttcagcaataggggctaacaaagggagggtgcgggcctgctcggcgagcccgagaggtcatggt
701 tggggagggaatggaggacaggagtgccggtggggcccccgccttcggagacacatgtccgacgcccacctggtggggcagggcctggggtttttccc
801 gaagcaaccaggctggggttagcgtgcccagggccatgtggccccagcaccggcacgatctggctggcgccgcgcttgcctcctccctaactagg

SmaI (953)

901 gtgaggccatcccgccggaccagttgcgtgcgtgaaagatggccgctccgggcccctggtgcaaggagctcaaaatggaggacgaggcagccgggtg
1001 gagcggcggggtgagtcaccacacacaaaggaagagggtcgttccctcaccggctgctgtctcctgtgaccccggtggtcctatcgccgcaatagtcacc
1101 tcgggcttttgagcacggctagtcgcccggggggaggggtgtaatggcggtggagtttggtcacatttggtggggggagactagtcaggccagcctgg
1201 cgctggaagtcattttttggaattgtcccccttagtgggagcgttaattctcgggcttcttagcggtcaaaaggtatcttttaaaccttttttag

BspHI (1333)

NheI (1371)

1301 GTGTTGTGAAAACACCGCTAATTCAAAAGCAATCATGAGCGGTTCTCATCATCATCATCATGATGGTAGCATGACTGGTAGCAGCAAATGGGT
▶▶Me tSer Gl ySer Hi sHi sHi sHi sHi sGl yMe tAl aSer Met Thr Gl yGl yGl nGl nMe tGl y

Bsu36I (1432)

Acc65I (1427)

1401 CGGGATCTGTACGACGATGACGATAAGGTACCTAAGGATCAGCTTGGAGTTGATCCCCTGGTTTTACAACGCTCGTGACTGGGAAAACCTGGCGTTACCC
23▶ ArgAspLeuTyrAspAspAspAspLysValP roLysAspGl nLeuGl yVal AspP roVal Val LeuGl nArgArgAspTrpGl uAsnP roGl yVal Thr G
1501 AACTTAATCGCCTTGCAGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCGCCACCGATCGCCCTCCCAACAGTTGCCAGCCTGAATGG
56▶ l nLeuAsnArgLeuAl aAl aHi sP roP roPheAl aSer TrpArgAsnSer Gl uGl uAl aArgThrAspArgP roSer Gl nGl nLeuArgSer LeuAsnGl

Bsu36I (1669)

1601 CGAATGGCCTTTGCCTGGTTTTCCGGCACCAGAAGCGGTGCCGAAAGCTGGCTGGAGTGCAGTCTTCTGAGCCGATACTGTGCTGCCCTCAAAC
89▶ yGl uTrpArgPheAl aTrpPheP roAl aP roGl uAl aVal P roGl uSer TrpLeuGl uCysAspLeuP roGl uAl aAspThr Val Val Val P roSerAsn
1701 TGGCAGATGCACGGTTACGATGCGCCCATCTACCCAACGTAACCTATCCCATTACGGTCAATCGCCGTTTGTCCACGGAGAATCGACGGGTGTT
123▶ TrpGl nMe tHi sGl yTyrAspAl aP roI leTyrThrAsnVal Thr TyrP roI leThr ValAsnP roP roPheVal P roThr Gl uAsnP roThr Gl yCysT
1801 ACTCGCTCACATTAATGTGTGAAAGCTGGCTACAGGAAGCCAGACCGAATATTTTGTATGGCCTAACCTCGCCGTTTCACTGTGGTGCACGG
156▶ y rSer LeuThr PheAsnVal AspGl uSer TrpLeuGl nGl uGl yGl nThr ArgI l eI l ePheAspGl yVal AsnSerAl aPheHi sLeuTrpCysGAATGG
1901 GCGCTGGGTCGGTTACGGCAGGACAGTCTGTTGCCGCTGAATTTGACCTGAGCGCATTTTTACGGCCGGAGAAAACCGCCTCGCGGTGATGGTGCTG
189▶ yAr gTrpVal Gl yTyrGl yGl nAspSer ArgLeuP roSer Gl uPheAspLeuSerAl aPheLeuArgAl aGl yGl uAsnArgLeuAl aVal Me tVal Leu
2001 CGTTGGAGTAGCGCGGATCTGGAAGATCAGGATATGGCCGATGAGCCGCAATTTCCGCGATCTCGTGTGCATAAACCGACTACAAAATCA
223▶ Ar gTrpSerAspGl Tyr LeuGl uAspGl uAspMe tTrpArgMe tTyr ArgMe tY l l ePheArgAspVal Ser LeuLeuAl aLysP roThr Gl nI l eS
2101 GCGATTTCCATGTTGCCACTCGCTTTAATGATGATTTGACGGCGCTGTACTGGAGGCTGAAGTTCAGATGTGCGCGGAGTTGCGTGACTACCTACGGT
256▶ erAspPheHi sValAl aThr ArgPheAsnAspAspPheSer ArgAl aVal LeuGl uAl aGl uVal Gl nMe tCysGl yGl uLeuArgAspTyrLeuArgVal
2201 AACAGTTTCTTATGGCAGGGTAAACGACGGTCCGACGGCCGACCCGCTTCCGGCGTGAATTCGATGAGCGTGGTGTATGCCATCGCGTC
289▶ I Thr Val SerLeuTrpGl nGl yGl nValAl aSer Gl yThrAl aP roPheGl yGl yGl uI l eI l eAspGl uAspGl yGl yTyrAl aAspArgVal
2301 AACTACGTCTGAACGTCGAAAACCCGAAACGTGGAGCGCCGAAATCCGAATCTCTATCGTGCGTGGTTGAAGTGCACACCGCCGACGGCACGCTGA
323▶ Thr LeuArgLeuAsnVal Gl uAsnP roLysLeuTrpSerAl aGl uI l eP roAsnLeuTyrArgAl aVal Val Gl uLeuHi sThrAl aAspGl yThr LeuI
2401 TTGAAGCAGAAGCCTCGCATGTCGTTTCCGCGAGGTGCGGATTGAAATGGTCTGCTGCTGACGGAACCCGCTGCTGATTGAGGGCTTAACCCG
356▶ I eGl uAl aGl uAl aCysAspVal Gl yPheArgGl uVal Ar gI l eGl uAsnGl yLeuLeuLeuAsnGl yLysP roLeuLeuI l eArgGl yVal AsnAr

EcoRV (2558)

2501 TCACGAGCATCATCCTCTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGAGGATATCCTGCTGATGAAGCAGAACAATTAACGCCGTGCGCTGT
389▶ gHi sGl uHi sHi sP roLeuHi sGl yGl nVal Me tAspGl uGl nThr Me tVal Gl nAspI l eLeuLeuMe tLysGl nAsnAsnPheAsnAl aVal Ar gCys
2601 TCGCATTATCCGAACCATCCGCTGTGGTACAGCTGTGCGACCGTACCGCCGCTGTATGGTGGTGAAGCCCAATTTGAAAACCCACGCGCATGGTGCCAA
423▶ Ser Hi sTyrP roAsnHi sP roLeuTrpTyrThr LeuCysAspArgTyrGl yLeuTyrVal Val AspGl uAl aAsnI l eGl uThr Hi sGl yMe tVal P roM
2701 TGAATCGTCTGACCGATGATCCGCGCTGGCTACGGCGATGAGCGAAGCGCTAACGCGAATGGTGACGCGGATCGTAATCACCCGAGTGTGATCATCTG
456▶ e tAsnArgLeuThrAspAspP roArgTrpLeuP roAl aMe tSer Gl uArgVal Thr ArgMe tVal Gl nArgAspArgAsnHi sP roSerVal I l eI l eTr
2801 TCGCTGGGGAATGAATCAGCGAACCGGCCGCTAATCACGACGGCTGTATCGCTGGAATCAATCTGTCGATCCTTCCCGCCGGTGCAGTATGAAGCCGCG
489▶ pSer LeuGl yAsnGl uSer Gl yHi sGl yAl aAsnHi sAspAl aLeuTyrArgTrpI l eLysSer Val AspP roSer ArgP roVal Gl nTyrGl uGl yGl y
2901 GGAGCCGACACCCGCGCCAGATATTATTTGCCGATGTACGGCGCGCTGGATGAAGACCAGCCTTCCCAGGCTGCGCAATGGTCCATCAAAAAAT
523▶ Gl yAl aAspThr ThrAl aThrAspI l eI l eCysP roMe tTyrAl aArgVal AspGl uAspGl nP roPheP roAl aVal P roLysTrpSerI l eLysLysT
3001 GGCTTTCCGCTAGCTGGAGAGACCGCCCGCTGATGTGAATACGGCCACCGGTAACAGTCTTGGCGGTTCGCGGTTTCGCTAATCTGACAGGCGCTT
556▶ rP leuSer LeuP roGl yGl uThr ArgP roLeuI l eLeuCysGl uTyrAl aHi sAl aMe tGl yAsnSer LeuGl yGl yPheAl aLysTyrTrpGl nAl aPh
3101 TCGTCACTATCCCGTTTACAGGGCGGCTTGTGCTGGACTGGGTGGATCAGTTCGCTGATTAATATGATGAAGCAGGCAACCCGCTGGTGGCTACGCG
589▶ eArgGl nTyrP roArgLeuGl nGl yPheVal TrpAspTrpVal AspGl nSer LeuI l eLysTyrAspGl uAsnGl yAsnP roTrpSerAl aTyrGl y
3201 GGTGATTTGGCGATACCGCAACCGATCCGAGTCTGTATGAACGGTCTGGCTTTTGGCCAGCCGACCGCATCCAGCCTGACGGGAAGCAAAACACC
623▶ Gl yAspPheGl yAspThr P roAsnAspArgGl nPheCysMe tAsnGl yLeuVal PheAl aAspArgThr P roHi sP roAl aLeuThr Gl uAl aLysHi sG
3301 AGCAGAGTTTTTCCAGTCCGTTTATCCGGGCAACCATCGAAGTGAACCGCAATACCTGTTCCGTCATAGCGATAACGAGCTCTGCACTGGATGGT
656▶ I nGl nGl nPhePheGl nArgLeuSer Gl yGl nThr I l eGl uVal Thr Ser Gl uTyrLeuPheArgHi sSerAspAsnGl uLeuLeuHi sTrpMe tVa
3401 GGGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTTGTGATGTCGCTCCACAAGGTAACAAGTTGATTGAAGCTGCTGCAACTCCGACGGCAGG
689▶ I Al aLeuAspGl yLysP roLeuAl aSer Gl yGl uVal P roLeuAspVal Al aP roGl nGl yLysGl nLeuI l eGl uLeuP roGl uLeuP roGl nP roGl u

3501 AGCGCCGGCAACTCTGGCTCACAGTACGCGTAGTGCAACCGAACGCGACCGCATGGTCAAGAAGCCGGGCACATCAGCGCTGGCAGCAGTGGCGTCTGG
723 SerAl aGlyGlnLeuTrpLeuThrValArgValValGlnProAsnAlaThrAlaTrpSerGluAlaGlyHisIleSerAlaTrpGlnGlnTrpArgLeuA
3601 CGGAAAACCTCAGTGTGACGCTCCCGCCGGTCCACGCCATCCCGCATCTGACCACCAGCGAAATGGATTTTTGCATCGAGCTGGTAATAAGCGTTG
756 IAGluAsnLeuSerValThrLeuProAlaAlaSerHisAlaIleProHisLeuThrThrSerGluMetAspPheCysIleGluLeuGlyAsnLysArgTr
3701 GCAATTTAACCGCCAGTCAGGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACAACCTGCTGACGCGCTGCGCGATCAGTTCACCCGTGCACCGCTG
789 pGlnPheAsnArgGlnSerGlyPheLeuSerGlnMetTrpIleGlyAspLysLysGlnLeuLeuThrProLeuArgAspGlnPheThrArgAlaProLeu
3801 GATAACGACATTGGCGTAAGTGAAGCGACCCGATTGACCTAACGCTGGGTGCAACGCTGGAAGCGCGGGCCATTACCAGGCCAAGCAGCGTTGT
823 AspAsnAspIleGlyValSerGluAlaThrArgIleAspProAsnAlaTrpValGluArgTrpLysAlaAlaGlyHisTyrGlnAlaGluAlaLeuL
3901 TGCAGTGCACGCGAGATACACTTGGCTGATGCGGTGCTGATTACGACCGCTCACGCGTGGCAGCATCAGGGGAAAACCTTATTTATCAGCCGAAAACCTA
856 euGlnCysThrAlaAspThrLeuAlaAspAlaValLeuIleThrThrAlaHisAlaTrpGlnHisGlnGlyLysThrLeuPheIleSerArgLysThrTy
4001 CCGGATTGATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATACACCGCATCCGGCGCGGATTGGCCTGAAGTGGCAGTGGCG
889 rArgIleAspGlySerGlyGlnMetAlaIleThrValAlaValAlaSerAspThrProHisProAlaArgIleGlyLeuAsnCysGlnLeuAla
4101 CAGGTAGCAGAGCGGTAACTGGCTCGGATTAGGGCCGCAAGAAAACCTATCCGACCGCTTACTGCGCCTGTTTTGACCGCTGGGATCGCAATTGT
923 GlnValAlaGluArgValAsnTrpLeuGlyLeuGlyProGlnGluAsnTyrProAspArgLeuThrAlaAlaCysPheAspArgTrpAspLeuProLeuS
4201 CAGACATGTATACCCGTAGCTCTCCCGAGCGAAAACGGTCTGCGCTGCGGGAGCGCGAATTGAATTATGGCCACACCAGTGGCGCGCGACTTCCA
956 erAspMetTyrThrProTyrValPheProSerGluAsnGlyLeuArgCysGlyThrArgGluLeuAsnTyrGlyProHisGlnTrpArgGlyAspPheGly
4301 GTTCAACATCAGCCGTCACAGTCAACAGCAACTGATGGAACCCAGCCATCGCCATCTGTCAGCAGCGGAAAGGACATGGTGAATATCAGCGGTTTC
989 nPheAsnIleSerArgTyrSerGlnGlnGlnLeuMetGluThrSerHisArgHisLeuLeuHisAlaGluGluGlyThrTrpLeuAsnIleAspGlyPhe
4401 CATATGGGGATTGGTGGCGACACTCTGGAGCCGCTCAGTATCGCGGAAATTACAGCTGAGCGCCGCTCGCTACCATTACCAGTTGGTCTGGTGTCAAA
1023 HisMetGlyIleGlyGlyAspAspSerTrpSerProSerValSerAlaGluLeuGlnLeuSerAlaGlyArgTyrHisTyrGlnLeuValTrpCysGlnL

NheI (4524)

EcoRI (4518)

4501 AATAATAATCTAGTCGAGAATTCGCTAGCTCGACATGATAAGATACATTGATGAGTTTGGACAAACCAACTAGAATGCAGTGAAAAAATGCTTTATT
1056 ys...
4601 TGTGAAATTTGTGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAA

PacI (4799)

SwaI (4789)

4701 TTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAGATCCATTTAAATGTTAATT
4801 AACTAGCCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCT
4901 GCGCGTAATCTGCTGCTTCAAACAAAAAACACCGCTACCAGCGGTGTTTTGTTTGGCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGC
5001 TTCAGCAGAGCGCAGATACCAAACTACTGTTCTTCTAGTGTAGCCGATGTTAGGCCACCACTTCAAGAACTCTGTAGCACCAGCTACATACCTCGCTCGC
5101 TAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGCTG
5201 AACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAAGTATGAGAAAGCGCCACGCTTCCGAA
5301 GGGGAAAGGCGGACAGGTATCCGGTAAGCGGCGAGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGTCTCG
5401 TCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGCCTTTTTACGGTT

PacI (5539)

AseI (5565)

5501 CCTGGCCTTTTGTGGCCTTTTGTCTCACATGTTCTTAATTAATTTTTCAAAGTAGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATA
5601 CGACTCACTATAGGAGGGCCATCATGGCCAAGTTGACCAAGTGTGCCAGTGTCTCCAGTGTCTCACAGCCAGGGATGTGGCTGGAGCTGTTGAGTCTGGACTGACAG
5701 GTTGGGGTTCTCCAGAGATTTTGTGGAGGATGACTTTGCAAGTGTGGTCCAGAGATGATGTACCCTGTTTCATCTCAGCAGTCCAGGACAGGTGGTGCCT
5801 GACAACCCCTGGCTTGGGTGTGGGTGAGAGGACTGGATGAGCTGTATGCTGAGTGGAGTGGTGGTCTCCACCAACTTCAGGGATGCCAGTGGCCCTG
5901 CCATGACAGAGATTGGAGAGCAGCCCTGGGGAGAGAGTTTGCCTGAGAGACCCAGCAGGCAACTGTGTGCATTTTGTGGCAGAGGAGCAGGACTGAGG

PacI (6055)

6001 ATAAGAATTGAGTTTCAGAAAAGGGGGCCTGAGTGGCCCTTTTTTCAACTTAATTA