



PstI (6)
SdaI (6)
1 CCTGCAGGTCATAGTTCTCCAGCTGACTTTGTCAAGACAGTGATGCTGTGTCCAGCAGTTGTTCTGAGTATCCTTTTCATTATCCACTGCTCTTTC
101 TTCTTAAATTCACCCCAACATTGTAATAGTCTTTCTTAAACTCTGTTCAAAGAACCAGCTTGAGTGTGTACAGTCTTCTGCTGGGGTCTGGC

SpeI (206) **AseI (277)**
201 AAACCACTAGTGACCTTTATTCATAAGAGATGATGATTCTTGATACTTCTCATTTGCAAAATCCAATTATTATAATTTTCATATCAATTAGAATAA

HindIII (386)
301 TATATCTTCTTCAATTTAGTTACCTCACTATGGGCTATGTACAAACTCCAAGAAAAAGTTAGTCATGTGCTTTGCAGAAGATAAAAAGCTTAGTGTAATA
401 CAGGCTGAGAGTATTGTAGTAAGAAGGGGAGTGGTTATATAGTCTTAGCCAAAACATGTGATAGTCACTCCAGGGTGTCTGAAAAGAAGTCTGTGA

NcoI (546) **NheI (584)**
501 CACTCATTAACCTATTGGTGCAGAAATTTGAATGATCTAAAGGAGACCATGGGGGCTTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGG
MetGl yGl ySer Hi sHi sHi sHi sHi sHi sGl yMeAl aSer MeTThr Gl yGl

Acc65I (640)
601 ACAGCAAATGGGTGGGATCTGTACGACGATGACGATAAGGTACCTAAGGATCAGCTTGGAGTTGATCCCCTGTTTTACAACGTCGTGACTGGGAAAA
18> yGl nGl nMeTGl yArgAspLeuTyrAspAspAspAspLysVal P roLysAspGl nLeuGl yVal AspP roVal Val LeuGl nArgArgAspTrpGl uAsn
701 CCTGGCGTTACCCAACTTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCATCGCCCTTCCCAACAGTTGC
52> P roGl yVal Thr Gl nLeuAsnArgLeuAl aAl aHi sP roP roPheAl aSer TrpArgAsnSer Gl uGl uAl aArgThrAspArgP roSer Gl nGl nLeuA
801 GCAGCCTGAATGGCGAATGGCGCTTTCCTGGTTCCGGCACCAGAAGCGGTGCCGGAAGCTGGCTGGAGTGCATCTTCTGAGGCCGATACTGTCGT
85> r gSer LeuAsnGl yGl nLeuPheAl aTrpPheP roAl aP roGl uAl aVal P roGl uSer TrpLeuGl uCysAspLeuP roGl uAl aAspThr Val Va
901 CGTCCCCTCAAACCTGGCAGATGCACGGTTACGATGCGCCATCTACACCAACGTAACCTATCCCATTACGGTCAATCCGCCGTTTGTCCACGGAGAAT
118> l Val P roSerAsnT rpGl nMeTHi sGl yTyrAspAl aP ro l l eTyrThrAsnVal Thr TyrP ro l l eThr Val l AsnP roP roPheVal P roThr Gl uAsn
1001 CCGACGGGTTGTTACTCGCTCACATTAATGTTGATGAAAGCTGGCTACAGGAAGCCAGACGCGAATATTTTTGATGGCGTAACTCGGCCGTTTCATC
152> P roThr Gl yCysTyrSer LeuThr PheAsnVal AspGl uSer TrpLeuGl nGl uGl yGl nThr Arg l l e l l ePheAspVal l AspP roAl aPheHi sL
1101 TGTGGTGCAACGGCGCTGGGTGCGTTACGGCCAGGACAGTCGTTTGCCTGTAATTTGACCTGAGCGCATTTTTACGCGCCGAGAAAAACCGCTCGC
185> euTrpCysAsnGl yArgT rpVal l Gl yTyrGl yGl nAspSer ArgLeuP roSer Gl uPheAspLeuSer Al aPheLeuArgAl aGl yGl uAsnArgLeuAl
1201 GGTGATGGTGTGCTGGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGCATTTTCCGTGACGCTCTGTTGCTGCATAAACCG
218> aValMeTVal LeuArgT rpSer AspGl ySer TyrLeuGl uAspGl nAspMeT rpArgMeT Ser Gl y l l ePheArgSpVal Ser LeuLeuHi sLysP ro
1301 ACTACAAAATCAGCGATTTCCATGTTGCCACTCGCTTAATGATGATTTACGGCCGCTGACTGGAGGCTGAAATTCAGATGTGCGCGGATTTGCCGTG
252> Thr Thr Gl n l l eSerAspPheHi sVal Al aThr ArgPheAsnAspAspPheSer ArgAl aVal LeuGl uAl aGl uVal l Gl nMeT CysGl yGl uLeuArgA
1401 ACTACCTACGGGTAACAGTTTCTTTATGGCAGGTTGAAACGCGAGGTGCCAGCGCCGCGCCCTTTCGGCGGTGAAATATTCGATGAGCGTGGTGGTTA
285> spTyrLeuArgVal l Thr Val Ser LeuT rpGl nGl yGl uThr Gl nVal Al aSer Gl yThr Al aP roPheGl yGl yGl u l l e l l eAspGl uArgGl yGl yTy
1501 TGCCGATCGCGTACACTAGCTGAACGTCGAAAACCGGAACTGTGGAGCGCGAAATCCCGAATCTCTACTGTCGGTGGTTGAACTGCAACCGCC
318> r Al aAspArgVal l Thr LeuArgLeuAsnVal l Gl uAsnP roLysLeuT rpSer Al aGl u l l eP roAsnLeuTyrArgAl aVal Val l Gl uLeuHi sThr Al a
1601 GACGGCAGCTGATTTGAAGCAGAAGCTGCGATGTGCGTTTCCGCGAGGTGCGGATTTGAAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATTC
352> AspGl yThr Leu l l eGl uAl aGl uAl aCysAspVal l Gl yPheArgGl uVal l Arg l l eGl uAsnGl yLeuLeuLeuAsnGl yLysP roLeuLeu l l eA

EcoRV (1771)
1701 GAGGCGTTAACCGTCACGAGCATCATCTCTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGATATCTGCTGATGAAGCAGAACAACCTTAA
385> r gGl yVal l AsnArgHi sGl uHi sHi sP roLeuHi sGl yGl nVal MeTAspGl uGl nThr MeTVal l Gl nAsp l l eLeuLeuMeT LysGl nAsnAsnPheAs
1801 CGCCGTGGCTGTTTCGATTATCCGAACCATCCGCTGTTGGTACACGCTGTGCAGCCGCTACGGCTGTATGTTGGTGGATGAAGCCAATATTGAAACCCAC
418> nAl aVal ArgCysSer Hi sTyrP roAsnHi sP roLeuT rpTyr Thr LeuCysAspArgTyrGl yLeuTyrVal l AspGl uAl aAsn l l eGl uThr Hi s
1901 GGCATGGTGCAATGAATCGTCTGACCAGTATCCGCGTGGCTACCGCGATGAGCGAACCGGTAACCGCAATGGTGACGCGATCGTAATCACCCGA
452> Gl yMeTVal P roMeTAsnArgLeuThrAspAspP roArgT rpLeuP roAl aMeT Ser Gl uArgVal l Thr ArgMeTVal l Gl nArgAspArgAsnHi sP roS
2001 GTGTGATCATCTGGTTCGCTGGGGAATGAATCAGGCCACGGCTCAATCACGACGCGTATCGCTGGATCAAATCTGTCGATCCTTCCCGCCCGGTGCA
485> er Val l l e l l eT rpSer LeuGl yAsnGl uSer Gl yHi sGl yAl aAsnHi sAl aLeuTyrArgT rp l l eLysSerVal l AspP roAl aP roVal l Gl
2101 GTATGAAGGCCGCGGAGCCGACACCAGCCACCGATATTATTTGCCGATGTACGCGCGTGGATGAAGACCAGCCCTTCCGGCTGTGCCGAAATGG
518> nTyrGl uGl yGl yAl aAspThr Thr Al aThrAsp l l e l l eCysP roMeT TyrAl aArgVal l AspGl uAspGl nP roPheP roAl aVal P roLysT rp
2201 TCCATCAAAAAATGGCTTTCGCTACCTGGAGAGACGCGCCGCTGATCCTTTGCGAATACGCCACGCGATGGGTAACAGTCTTGGCGGTTTGCCTAAAT
552> Ser l l eLysLysT rpLeuSer LeuP roGl yGl uThr ArgP roLeu l l eLeuCysGl uTyrAl aHi sAl aMeTGl yAsnSer LeuGl yGl yPheAl aLysT
2301 ACTGGCAGGCGTTTCGTCAGTATCCCGTTTACAGGCGGCTTTCGCTGGGATGGTGGATCAGTCGCTGATTAATATGATGAAAACGGCAACCCGTG
585> y rT rpGl nAl aPheArgGl nTyrP roArgLeuGl nGl yGl yPheVal l TrpAspT rpVal l AspGl nSer Leu l l eLysTyrAspGl uAsnGl yAsnP roT r
2401 GTCGGCTTACGGCGGTGATTTTGGCGATACGCCGACGATCGCCAGTCTGTATGAACGGTCTGGTCTTTCGCCAGCCGACGCCGATCCAGCGCTGACG
618> pSer Al aTyrGl yGl yAspPheGl yAspThr P roAsnAspArgGl nPheCysMeTAsnGl yLeuVal l PheAl aAspArgT rpHi sP roAl aLeuTh
2501 GAAGCAAAACACAGCAGCAGTTTTTCCAGTTCGTTTTTCCGGCAAACCTCGCTGAGTGAACGACCGCAATACCTGTCGATGAGCGGTAACGAGCTCC
652> Gl uAl aLysHi sGl nGl nPhePheGl nPheArgLeuSer Gl yGl nThr l l eGl uVal l Thr Ser Gl uTyrLeuPheArgHi sSerAspAsnGl uLeuL
2601 TGCACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTGGATGTGCTCCACAAGGTAAACAGTTGATTGAACCTGCCTGAAC
685> euHi sT rpMeTVal l Al aLeuAspGl yLysP roLeuAl aSer Gl yGl uVal l P roLeuAspVal l Al aP roGl nGl yLysGl nLeu l l eGl uLeuP roGl uLe
2701 ACCGCAGCCGAGAGCGCCGCAACTTGCCTCACAGTACGCGTAGTGCAACGAAACGCGACCGCATGGTCAAGAACCGCGCACATCAGCGCTGGCAG
718> uP roGl nP roGl uSer Al aGl yGl nLeuT rpLeuThr Val l ArgVal l Val l Gl nP roAsnAl aThr Al aT rpSer Gl uAl aGl yHi s l l eSer Al aT rpGl n
2801 CAGTGGCGTCTGGCGAAAACCTCAGTGTGACGCTCCCGCCCGCTCCACGCCATCCCGCATCTGACCACAGCGAAATGGATTTTGCATCGAGCTGG
752> Gl nT rpArgLeuAl aGl uAsnLeuSer Val l Thr LeuP roAl aAl aSer Hi sAl a l l eP roHi sLeuThr Thr Ser Gl uMeTAspPheCys l l eGl uLeuG
2901 GTAAATAGCGTTGGCAATTAACCGCAGTCAAGCTTTTTCACAGATGTGGATGGCGGATAAAAAACAACCTGCTGACCGCGTCCGCGATCGTTAC
785> l yAsnLysArgT rpGl nPheAsnArgGl nSer Gl yPheLeuSer Gl nMeT rp l l eGl yAspLysLysGl nLeuLeuThr P roLeuArgAspGl nPheTh
3001 CCGTGCACCGCTGGATAACGACATTTGGCGTAAGTGAAGCGACCCGATTTGACCTAACGCTGGTTCGAACGCTGGAAGCGCGCGGCCATTACAGGCC
818> rArgAl aP roLeuAspAsnAl l eGl yVal l Ser Gl uAl aThrArg l l eAspP roAsnAl aT rpVal l Gl uArgT rpLysAl aAl aGl yHi sTyrGl nAl a
3101 GAAGCAGCGTTGTTGTCAGTGCACGCGAGATACACTTGCTGAGCGGTGCTGATTACGACCCTACCGCTGAGCAGCAGTGCAGGCAAAACCTTATTTATCA
852> Gl uAl aAl aLeuLeuGl nCysThr Al aAspThr LeuAl aAspAl aVal l Leu l l eThr Thr Al aHi sAl aT rpGl nHi sGl nGl yLysThr LeuPhe l l eS
3201 GCCGAAAACCTACCGATTGATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATACCCGCATCCGGCGCGGATTGGCTGAA
885> erArgLysThr TyrArg l l eAspGl ySer Gl yGl nMeTAl a l l eThr Val l AspVal l Gl uVal l Al aSerAspThr P roHi sP roAl aArg l l eGl yLeuAs
3301 CTGCCAGTCGGCAGAGTACGAGCGGGTAACTGGCTCGGATAGGGCCGCAAGAAAACCTACCCGACCGCTTACTCCGCTGTTTTGACCGCTGG
918> nCysGl nLeuAl aGl nVal l Al aGl uArgVal l AsnT rpLeuGl yLeuGl yP roGl nGl uAsnTyrP roAspArgLeuThr Al aAl aCysPheAspArgT rp
3401 GATCTGCCATTGTACAGCATGTATACCCGTACGCTTCCGAGCGAAAACGGTCTGCGCTGCGGGACCGCGAATGAATATGCCCCACACAGTGGC
952> AspLeuP roLeuSerAspMeT TyrThr P roTyrVal l PheP roSer Gl uAsnGl yLeuArgCysGl yThr ArgGl uLeuAsnTyrGl yP roHi sGl nT rpA

3501 GCGGCGACTTCCAGTTCAACATCAGCCGCTACAGTCAACAGCAACTGATGGAAACCAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCACATGGCTGAA
985▶ rGlyAspPheGlnPheAsnIleSerArgTyrSerGlnGlnGlnLeuMetGluThrSerHisArgHisLeuLeuHisAlaGluGluGlyThrTrpLeuAs
3601 TATCGACGGTTCCATATGGGGATTGGTGGCGACGACTCCTGGAGCCCGTCAGTATCGGCGGAATTACAGCTGAGCGCCGGTCGCTACCATTACCAGTTG
1018▶ nIleAspGlyPheHisMetGlyIleGlyGlyAspAspSerTrpSerProSerValSerAlaGluLeuGlnLeuSerAlaGlyArgTyrHisTyrGlnLeu

NheI (3737)

EcoRI (3731)

3701 GTCTGGTGTCAAAAATAATAATCTAGTCGAGAATTCGCTAGCTCGACATGATAAGATACATTGATGAGTTGGACAAACCACAAC TAGAATGCAGTGAAA
1052▶ ValTrpCysGlnLys•••

3801 AAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAG

SwaI (4002)

3901 TTAACAACAACAATTGCATTCATTTTATGTTTCAGGTT CAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTAGATCCAT

PacI (4012)

4001 TTA AATGTTAATTAAGTACGCCATGACCAAAATCCCTAACCTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAG

4101 ATCCTTTTTTCTGCGCGTAATCTGCTGCTTGAACAACAAAAACCACCGCTACCAGCGGTGGTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCC

4201 GAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACCTCAAGAAGCTCTGTAGCACCCCTACA

4301 TACCTCGCTCTGTAATCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGTTGGACTCAAGACGATAGTTACCGGATAAGGGCC

4401 AGCGGTGGGCTGAACGGGGGTTCTGTGCACACAGCCAGCTTGAGCGAACGACCTACACCGAAGTACAGGATACCTACAGCGTGAGCTATGAGAAAGCGC

4501 CACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTGGAAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCCTGGTAT

4601 CTTTATAGTCTGTGGGTTTTGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGGCGG

PacI (4752)

AseI (4778)

4701 CCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTAATTAATTTTTCAAAGTAGTTGACAATTAATCATCGGCATAGTATATCG

4801 GCATAGTATAATACGACTCACTATAGGAGGGCCATCATGGCCAAGTTGACCAGTGTGTCCAGTGTCCACAGCCAGGGATGTGGCTGGAGCTGTTGAGT

4901 TCTGGACTGACAGGTTGGGGTCTCCAGAGATTTGTGGAGGATGACTTTGCAGGTGTGGTCAGAGATGATGTACCCTGTTTCTCAGCAGTCCAGGA

5001 CCAGGTGGTGCCTGACAACCCCTGGCTTGGGTGTGGTGAGAGGACTGGATGAGCTGTATGCTGAGTGGAGTGAGGTGGTCTCCACCAACTTCAGGGAT

5101 GCCAGTGGCCCTGCCATGACAGAGATTGGAGAGCAGCCCTGGGGGAGAGAGTTGCCCTGAGAGACCCAGCAGGCAACTGTGTGCACTTTGTGGCAGAGG

PacI (5268)

5201 AGCAGGACTGAGGATAAGAATTGAGTTTCAGAAAAGGGGCGCTGAGTGGCCCTTTTTTCAACTTAATTAA