



SgfI (11)

1 GGATCTGCGATCGTCCGGTGCCCGTCAAGTGGGAGAGCGCACATCGCCCACAGTCCCCGAGAAAGTTGGGGGAGGGTTCGCAATTGAACGGTGCCTA

101 GAGAAGTGGCGCGGGTAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCCGC

HindIII (246)

PvuII (242) Bsu36I (293)

201 GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGGCCGCCGCCCTACCTGAGGGC

301 GGCATCCACGCCGTTGAGTTCGGCTTCCGCGCTCCCGCTGTGGTGCCTCCTGAACTGCCTCGCGCTTAGGTAAGTTAAAGCTCAGGTCGAGACC

401 GGGCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTGGCTGACCCTGCTTGTCTCAACTCTACGTTCTTGTTCGTTT

AgeI (553) **BspHI (569)**

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGTAGGAGGCCATCATGATTTGCACAGGCCATCACTTCTTAGC

601 GCCTTCGCATTTCTAGGTGACGACGAGCCAGGACAAAGCTTGTATCTGAAGCACAACTGAAGACCCACGCCAGGCTGCCAGGGCAGGACCTCTCGC

111▶Al aPheAspI l eLeuGlyAl aAl aGlyYl nAspLysLeuLeuTyrLeuLysHi sLysLeuLysThr P roArgP roGly yCysG l nGly nAspLeuLeuH

NcoI (704) **XbaI (750)**

701 ATGCCATGGTTCCTCAGGCTGGGCCAGGAACTGAGGCCAGGATCTCTAGAGGCATTGAAGCCGATGCGGTGGCCCGCTGGTGGCCCGCCAGTG

44▶i sAl aMe tVal LeuLeuLysLeuGlyYl nGly nUthr Gl uAl aArgI l eSer LeuGly uAl aLeuLysAl aAspAl aValAl aArgLeuValAl aArgGly nTr

801 GGCTGGCGTGGACAGACCCAGAGGAGCCAGAGCCAGATGTCCTGGGCTGTGGCGCGCTTAGACCACCTGAGGAGAAGCTGTGCC

77▶pAl aGly yVal AspSer Thr Gl uAspP roGly uGly uP roP roAspVal Ser TrpAl aValAl aArgLeuTyrHi sLeuLeuAl aGly uGly nLysLeuCysP ro

AgeI (995)

901 GCCTCGTGGGACGCTGGCTACAGGAAGCCGTCGCACCCTCAGCTCCAGGACGACCACCGGCTGGGGAATTCAGGATGAGCCCGAAACCGGT

111▶Al aSer LeuArgAspValAl aTyrGly nGly uAl aVal aArgThr LeuSer Ser aArgAspAspHi s aArgLeuGlyYl nLeuGly nAspGly uAl aArgAsnArgC

BamHI (1020) **BspEI (1034)**

1001 GTGGGTGGGACATTGCTGGGATCCAGGGAGCATCCGGACGCTCCAGTCCAATCTGGGCTGCCTCCCACCATCCTCGGCTTGGCCCTGGGACCAGGAG

144▶ysGly yTrpAl l eAl aGly yAspP roGlyYl n l eArgThr LeuGly nSerAsnLeuGly yCysLeuP roP roSer Ser Al aLeuP roSer Gl yThr ArgSe

1101 CCTCCCACGCCCATTTGACGGTGTTCGGACTGGAGCCAGGGTGTCTCCGATCCACTGGCAGCCCTGCCTCCCTGGCCAGCAACTGGAAATCAGC

177▶r LeuP roArgP rol l eAspGly yVal SerAspTrpSer Gl nGly yCysSer LeuArgSer Thr Gl ySer P roAl aSer LeuAl aSerAsnLeuGly u l l eSer

1201 CAGTCCCTACCATGCCCCTTCCAGCCTGCACCGCAGCCACATGGGCCAGCAAGCTCTGTGACGACCCCGAGCCAGCTTGGTGGCCGAGCTGTCC

211▶Gl nSer P roThr Me tP roPheLeuSer LeuHi sArgSer P roHi sGly yP roSer LysLeuCysAspAspP roGly nAl aSer LeuVal l P roGly uP roVal l P

Bsu36I (1319)

1301 CCGTGGCTGCCAGGAGCTGAGGAGATGAGCTGGCCGCATCGGGGAGATTGCCAGCCACCAGAGCTGCAAGCAGCCACCTCCTGGGCTTCCCGA

244▶r oGly yGly yCysGly nGly uP roGly uGly uMe tSer TrpP roP roSer Gly yGly u l l eAl aSer P roP roGly uLeuP roSer Ser P roP roGly yLeuP roGly

1401 AGTGGCCAGATGCAACCTCCACTGGCTCCCTGATACCCCGCAGCTCCAGAAACAGCACCACCTACCCAGTGGAGTGACCAGGGTCTGCAGGC

277▶uValAl aP roAspAl aThr Ser Thr Gl yLeuP roAspThr P roAl aAl aP roGly uThr Ser ThrAsnTyrP roVal Gl uCysThr Gl uGly ySerAl aGly y

1501 CCCAGTCTCTCCCCTTGCTATTCTGGAGCGGTCAAAACCCCTGCTGTGCAAGACAGAGCCACTCCAACTTTCTGTAAGATACCACCTCTC

311▶P roGly nSer LeuP roLeuP rol l eLeuGly uP roVal l LysAsnP roCysSer Val LysAspGly nThr P roLeuGly nLeuSer Val Gl uAspThr Thr Ser P

1601 CAATACCAAGCGTGCACCTCCACCTCCACCCAGAAACATCCCTCCTCCTCCCTCCTCTCCTCCTCTCCTCCTCCTCCTCTCTCTCTGTTGTTGCTCACTGAC

344▶r oAsnThr LysP roCysP roP roThr P roThr Thr P roGly uThr Ser P roP roP roP roP roP roP roSer Ser Thr P roCysSer Al aHi sLeuTh

1701 CCCCTCCTCCTGTTCCCTCCCTCCCTGGAATCATCATCGGAACAGAAATCTATAAATTTGGTATCTCCACGCCAGGGCAGACAAACACATGCCCTG

377▶r P roSer Ser LeuPheP roSer Ser LeuGly uSer Ser Ser Gl uGly nLysPheTyrAsnPheVal l l eLeuHi sAl aArgAl aAspGly uHi s l l eAl aLeu

SdaI (1898)

1801 CGGTTCCGGGAGAAGCTGGAGGCCCTTGGCTGCCGACGGGCCACCTTCTGCGAGGATTTCCAGGTGCCGGGCGCGGGGAGCTGAGCTGCCTGCAGG

411▶ArgVal l ArgGly uLysLeuGly uAl aLeuGly yVal l P roAspGly yAl aThr PheCysGly uAspPheGly nVal l P roGly yArgGly yGly uLeuSer CysLeuGly nA

1901 ACGCCATAGACCACCTCACTTTCATCCTACTTCTCACTTCCAACTTCGACTGTGCGCTGAGCCTGCACCAGGTGAACCAAGCCATGATGAGCAACT

444▶spAl a l eAspHi sSerAl aPhe l l e l eLeuLeuLeuThr SerAsnPheAspCysArgLeuSer LeuHi sGly nVal l AsnGly nAl aMe tMe tSerAsnLe

2001 CACGGCAGAGGGTCCAGACTGTGCATCCCCCTTCCGCGCTGGAGAGCTCCCGGCCAGCTCAGCTCCGACAGCCGACGCTGCTCTCCGGGCTG

477▶uThr ArgGly nGly ySer P roAspCysVal l l eP roPheLeuP roLeuGly uSer Ser P roAl aGly nLeuSer Ser AspThr Al aSer LeuLeuSer Gl yLeu

BglIII (2123)

2101 GTGCGGCTGGACGAACACTCCAGATCTTCGCCAGGAAGTGGCCAACACTTCAAGCCCCACAGGCTTACGGCCCCGAAAGGCACTGGAGGAAGGAAC

511▶Val aArgLeuAspGly uHi sSer Gl n l ePheAl aArgLysValAl aAsnThr PheLysP roHi sArgLeuGly nAl aArgLysAl aMe tTrpArgLysGly uG

2201 AGGACACCCGAGCCCTGGCGAAACAGAGCCAAACCTGGACGGTGGCGGATGCAGGCGGGCAGCTGAACGCAGCCCTACTCAGCCTACCTCCAGAGCTA

544▶ l nAspThr ArgAl aLeuArgGly uGly nSer Gl nHi sLeuAspGly yGly uArgMe tGly nAl aAl aAl aLeuAsnAl aAl aTyrSerAl aTyrLeuGly nSer Ty

2301 CTTGTCTCAGCAGGACAGATGGAGACTTCCAGTGGCTTTTGGGAGCCACATGTGCATTTGGACTGGGGCGCCTAGCTGAGCCTGGGGCGCTGGGGCTGGG

577▶r LeuSer TyrGly nAl aGly nMe tGly uGly nLeuGly nValAl aPheGly ySer Hi sMe tSer PheGly yThr Gl yAl aP roTyrGly yAl aArgMe tP roPheGly y

2401 GGCCAGGTGCCCTGGAGCCCCGCCACCTTCCACTTGGCGGGGTGCCCGCAGCCGCCACCCCTGCACGCATGGCAGGCTGGCACCCTCCACCCG

611▶Gly yGly nVal l P roLeuGly yAl aP roP roP roPheP roThr TrpP roGly yCysP roGly nP roP roP roLeuHi sAl aTyrGly nAl aGly yThr P roP roP

2501 CTTCCACAGCCAGCAGCTTCCAGACTCACTGCCCTTCCCGAGTCCCGAGCCTTCCCTACGGCCTACCCGACCCCTCAGAGCCAGGGCTGCA

644▶r oSer P roGly nP roAl aAl aPheP roGly nSer LeuP roPheP roGly nSer P roAl aPheP roThr Al aSer P roAl aP roP roGly nSer P roGly yLeuGly

PvuII (2634)

2601 ACCCTCATTATCCACCACGACAGATGGTACAGCTGGGCTGAACAACCATGTTGGAACAGAGAGGTTCCAGGCGCCGAGGACAAGACGAGGAG

677▶n P roLeu l l e l l eHi sHi sAl aGly nMe tVal l Gl nLeuGly yLeuAsnAsnHi sMe tTrpAsnGly nArgGly ySer Gl nAl aP roGly uAspLysThr Gl nGly u

NheI (2710)

2701 GCAGAATGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTGGACAAACCAACTAGAATGCAGTGAATAAATGCTTTATTTGTAATTT

711▶Al aGly u●●●

HpaI (2850)

2801 GTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTG

EcoRI (2944)

2901 GGAGTTTTTTAAAGCAAGTAAACCTCTACAAATGGTATGGAATCTAAAAATACAGCATAGCAAACTTTAACTCCAATCAAGCCTCTACTTGA

3001 TCCTTTTCTGAGGATGAATAAGGCATAGGCATCAGGGCTGTGGCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTAAAGATATA

SwaI (3200)

3101 GTGATTTTTCCAAGTTTGAACCTAGCTCTTTCATTTCTTATGTTTTAAATGCAGCTGACCTCCACATTCCCTTTTGTAAATATTCAGAAATAATTT

3201 AAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGG
3301 AACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTATCCTCAGTCTGCTCTCTGCCACAAAGTGCACGCAGTTGCCGGCC
125◀●●●AspGlnGluAlaValPheHisValCysAsnGlyAlaP
3401 GGGTCGCGCAGGGCGAACTCCCGCCCCACGGTGTCTGCCGATCTCGGTATGGCCGGCCGGAGGCGTCCCGAAAGTTCGTGGACACGACCTCCGACC
110◀roAspArgLeuAlaPheGluArgGlyTrpProGlnGlyIleGluThrMetAlaProGlySerAlaAspArgPheAsnThrSerValValGluSerTr
3501 ACTCGGCGTACAGCTCGTCCAGGCCGCGCACCCACCCAGGCCAGGGTGTGTCCGGCACCACCTGGTCTGGACCGCGTGATGAACAGGGTCAAGTCC
77◀pGluAlaTyrLeuGluAspLeuGlyArgValTrpValTrpAlaLeuThrAsnAspProValValGluAspGlnValAlaSerIlePheLeuThrValAsp
SmaI (3641)
3601 GTCCCGGACCACCCGGCAAGTCTCTCCAGGAAGTCCCGGAGAACCCGAGCCGGTGGTCCAGAAGTCCGACCGTCCGGCGACGTCCGGCCGGTG
44◀AspArgValValGluAlaPheAspAspGluValPheAspArgSerPheGlyLeuArgAspThrTrpPheGluValAlaGluAlaValAspArgAlaThrL
AseI (3788)
3701 AGCACCGBAACGGCACTGGTCAACTTGGCCATGATGGCCCTCTATAGTGAGTGTATTATACATATGCCGATATACTATGCCGATGATTAATTGTCAA
10◀euValProValAlaSerThrLeuLysAlaMet
3801 CAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTCGCTCAATG
SpeI (3942)
3901 GGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCCATTTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGA
4001 GTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCC
4101 CATAAGGTCACTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCA
4201 AGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGG
SdaI (4358) PaeI (4365)
4301 GGGGTGCTTGGGCGGTACGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCA
4401 GGAACCGTAAAAAGCCCGTGTGCTGGCGTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCC
4501 GACAGGACTATAAAGATACCAGGCGTTTCCCGTGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCTTTCTC
4601 CCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCG
4701 TTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCCACTGGTAACAGGAT
4801 TAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTG
4901 AAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGC
5001 GCAGAAAAAAGGATCTCAAGAAGATCTTTGATCTTTCTACGGGCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAG
PaeI (5105) SmaI (5113) NotI (5121)
5101 TTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTTGTGTAATCGTAACTAACATACGCTCTCC
5201 ATCAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCAAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA