



PstI (6)
SdaI (6) SpeI (13)
1 CCTGCAGGGCCCACTAGTCATGGATGGTTGTGAACCACCATGTGGTTGCTGGGATTGAACCTCCGGACCTTTGGAAAAGAAGTCGGGTGCTCTTACCCAC
101 TGAACCATCTCACCAGCCCCACAGAACTCTTATACAAGAAGAAGAGGGCTGTGTTGAGGGCTAGACTTCATTTTTCTCTCTCTCTCTCTCTCTCT
201 TCTTTCTCTTCT
301 TAGCCCTGCCTATCTGAAACTCACTCTATAGACCAGACTGGCCTCAAACCTCATAGAAATCCACCTGCCTCTGCCTCCCGAGTGCTAGGATTAAGGCAT
401 GACTCACCACCACCCGACTCTGGGCTGTATTTCTCTATTTGTTGGCATACTTTTGACACAAGAACCTCTATGGCTCCCTCCACAGAGCCAACTTTTG
501 ACATGGCTTCTGTACTCCAAGAAGTTCTACTCTGGACCTCTCCAGGAGCCTGGTGTGAGTATGTCTCTGAATTGTCTGTCTTGTCTGTATCTCTCTCT
AgeI (667)
601 AGAGTTAGATAAGGGTTGTCCAGGGCAGAGGAGTCAAGGTCAGATTCATTTAAGTAACAGTTCTCAACCGGTGGGTATCACCATTAGAGAACACATATT
AseI (768)
701 TCCAATGGTCTTAGGAAGTCTGAGACACCTCTCAGTAGCAAAATTACAGTTTTGAAAGTAGCAACAAAAATTAATTTATGTTGGGGTTCTTAAGACCATCA
801 GAAATGTGTGCAGTTGCCACCCACAGGTTGAAAAGTCTGATTGAGAAGGTTGGTGTGCATTTAGCTAAATCCCGAGGCCACTCTTCCAGAGCAAGGC
901 AACACAGGAGACCAGCTGACCTAAGAGATGACCCACAGTGGGACAGAGGGAGTTAGAAACTGGCTGTCCCTAGAAGATTCAGGAGCAACCCCTTCCC
EcoNI (1087)
1001 AGGGGGTCTCTCAAGAACCACAAAGCTGTGCTACCCAGGCAGGACCACCTTCGAGTCTAGCAGACCAGCTGATGAAGCTGCCCTCCCATAGGGTCCCTG
PstI (1183)
1101 GAGGGATGGTGCCTCCCCTGGGTCTCAATTTGCATGGCAGGAAGGGCCTAATGGGGAAGAGCGGAGAGGAGACAAGCTGCAGCCAGGCTGACATACA
NcoI (1251) NheI (1289)
1201 GTTACATGTTTCGCCCCAAACCTCCTTGGGCTCAGAGACAGAGCAGTGCACATGGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACT
Bsu36I (1350)
1301 GGTGGACAGCAAATGGGTCGGGATCTGTACGACGATGACGATAAGGTACCTAAGGATCAGCTGGAGTTGATCCCGTCGTTTTACAACGTCGTGACTGGG
17> Gl yGl yGl nMe tGl yAr gAspLeuTyrAspAspAspLysVal P roLysAspGnGCCAGACGCGAATTATTTTATGGTGGGCTTAACTGAGCCGATACT
1401 AAAACCTGGCGTTACCCAACTTAATCGCCTTGACGACATCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGACCGATCGCCTTCCCAACA
50> l uAsnP roGl yVal Thr Gl nLeuAsnArgLeuAl aAl aHi sP roP roPheAl aSer TrpArgAsnSer Gl uGl uAl aArgThrAspArgP roSer Gl nGl
Bsu36I (1587)
1501 GTTGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTCCGGCACCAAGAAGCGGTGCCGAAAAGCTGGCTGGAGTGCATCTTCTGAGCCGATACT
83> nLeuArgSer LeuAsnGl yGl uTrpArgPheAl aTrpPheP roAl aP roGl uAl aVal P roGl uSer TrpLeuGl uCysAspLeuP roGl uAl aAspThr
1601 GTCGTGTCCTCAAACTGGCAGATGCACGGTTACGATGCGCCATCTACACCAACGTAACCTATCCATTACGGTCAATCCGCGTTTGTCCACGGC
117> Val Val Val P roSerAsnTrpGl nMe tHi sGl yTyrAspAl aP roI l eTyrThrAsnVal Thr TyrP roI l eThrValAsnP roPheVal P roThrG
1701 AGAATCCGAGGGTTGTACTCGCTCACATTAATGTTGATGAAAGTGGCTACAGAAAGCCAGACGCGAATTATTTTATGGTGGGCTTAACTGAGCCGCTT
150> l uAsnP roThr Gl yCysTyrSer LeuThr PheAsnVal AspGl uSer TrpLeuGl nGl uGl yGl nThrArgI l eI l ePheAspGl yVal lAsnSerAl aPh
1801 TCATCTGTGGTCAACGGCGCTGGGTGGTTACGCCAGGACAGTCTGTTCCGCTCTGAATTTGACCTGAGCGCATTTCACGCCCGGAGAAAACCGC
183> eHi sLeuTrpCysAsnGl yAr gTrpVal Gl yTyrGl yGl nAspSer Ar gLeuP roSer Gl uPheAspLeuSerAl aPheLeuArgAl aGl yGl uAsnArg
1901 CTCGCGGTGATGGTCTGCCTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGATTTCCTGACGCTCTGTTGCTGACATA
217> LeuAl aValMe tVal LeuArgTrpSerAspGl ySer TyrLeuGl uAspGl nAspMe tTrpArgMe tSer Gl yI l ePheArgAspVal SerLeuLeuHi sL
2001 AACCGACTACAAAATCAGCGATTCCATGTTGCCACTCGCTTAAATGATGATTTACGCCGCGCTGTACTGGAGGCTGAAGTTGATGTGCGCGGAGTT
250> ysP roThr Thr Gl nI l eSerAspPheHi sValAl aThr Ar gPheAsnAspAspPheSer Ar gAl aVal LeuGl uAl aGl uVal Gl nMe tCysGl yGl uLe
ClaI (2187)
2101 GCGTGACTACCTACGGTAAACAGTTCTTTATGGCAGGGTGAACGCAGGTCGCCAGCGGCCACCGCCCTTTCCGGCGTGAATATTCGATGAGCGTGGT
283> uArgAspTyrLeuArgVal Thr Val SerLeuTrpGl nGl yGl uThr Gl nValAl aSer Gl yThrAl aP roPheGl yGl yGl uI l eI l eAspGl uArgGl y
2201 GGTATGCCGATCGCGTACACTACGCTGAACGTCGAAAACCCGAACTGTGGAGCGCGGAAATCCCGAATCTCTATCGTGGTGGTGAACATGCACA
317> Gl yTyrAl aAspArgVal ThrLeuArgLeuAsnVal Gl uAsnP roLysLeuTrpGl uAl aGl uI l eP roAsnLeuTyrArgAl aVal Val Gl uLeuHi sT
2301 CCGCCAGCGCAGCTGATTGAAGCAGAAGCCTCGATGTCGGTTCCGCGAGGTGCGGATTGAAAATGGTCTGCTGCTGTAACCGGCAAGCCGTTGCT
350> hrAl aAspGl yThrLeuI l eGl uAl aGl uAl aCysAspVal Gl yPheArgGl uVal Ar gl l eGl uAsnGl yLeuLeuLeuLeuAsnGl yLysP roLeuLe
EcoRV (2476)
2401 GATTCGAGGCGTTAACCGTCACGAGCATCTCTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGATCTCTGCTGATGAAGCAGAACAAC
383> ul l eArgGl yVal lAsnArgHi sGl uHi sHi sP roLeuHi sGl yGl nValMe tAspGl uGl nThr Me tVal Gl nAspI l eLeuLeuMe tLysGl nAsnAsn
2501 TTTAACGCCGTGCGTGTTCGCATTATCCGAACCATCCGCTGTGGTACACGCTGTGCGACCCTACGGCTGTATGTGGTGGATGAAGCCAATATTGAAA
417> PheAsnAl aVal Ar gCysSer Hi sTyrP roAsnHi sP roLeuTrpTyrThrLeuCysAspArgTyrGl yLeuTyrVal Val AspGl uAl aAsnI l eGl uT
2601 CCCACGGCATGGTCCAATGAATCGCTGACCGATGATCCGCTGCTACCCGCGTACCCGCGATGAGCGAAGCCGTAACCGCAATGGTGCAGCGCATCGTAATCA
450> hrHi sGl yMe tVal P roMe tAsnArgLeuThrAspAspP roAr gTrpLeuP roAl aMe tSer Gl uAr gVal Thr Ar gMe tVal Gl nAr gAspArgAsnHi
2701 CCCGAGTGTGATCATCTGGTCTGCTGGGAATGAATCAGGCCACGGCGTAATCAGCAGCGCTGTATCGCTGGATCAAATCTGTGATCTTCCCGCCG
483> sP roSer Val I l eI l eTrpSerLeuGl yAsnGl uSer Gl yHi sGl yAl aAsnHi sAspAl aLeuTyrArgTrpI l eLysSerVal AspP roSerArgP ro
BbsI (2880)
2801 GTGCAGTATGAAGCGCGGAGCCGACACCAGCCACCGATATTATTTGCCGATGTACGCGCGCTGGATGAAGCAGCCCTTCCCGCTGTGCCGA
517> Val Gl nTyrGl uGl yGl yAl aAspThr ThrAl aThrAspI l eI l eCysP roMe tTyrAl aAr gVal AspGl uAspGl nP roPheP roAl aVal P roL
2901 AATGGTCCATCAAAAATGGCTTCGCTACCTGGAGAGAGCGCCCGCTGATCCTTTGCGAATACGCCACGCGATGGGTAACAGTCTTGGCGTTTCGC
550> ysTrpSerI l eLysLysTrpLeuSerLeuP roGl yGl uThrAr gP roLeuI l eLeuCysGl uTyrAl aHi sAl aMe tGl yAsnSerLeuGl yGl yPheAl
3001 TAAATCTGGCAGCGTTTCGTCAGTATCCCGTTACAGGCGAGCTTCGCTGGAGTGGGATGATCAGTCTGATTAATATGATGAACAGCGCAAC
583> aLysTyrTrpGl nAl aPheArgGl nTyrP roAr gLeuGl nGl yGl yPheVal TrpAspTrpVal lAspGl nSerLeuI l eLysTyrAspGl uAsnGl yAsn
Eco47III (3198)
3101 CCGTGGTTCGCTTACGGCGGTGATTTGGCGATACGCCAACGATCGCCAGTCTGTATGAACGGTCTGGTCTTTGCCGACCGCACGCCGATCCAGCGC
617> P roTrpSerAl aTyrGl yGl yAspPheGl yAspThr P roAsnAspArgGl nPheCysMe tAsnGl yLeuVal PheAl aAspArgThrP roHi sP roAl aL

3201 TGACGGGAAGCAAACACCAGCAGCAGTTTTTCCAGTTCGGTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGATAACGA
650>euThr Gl uAl aLysHis sGl nGl nGl nPhePheGl nPheArgLeuSer Gl yGl nThr l l eGl uVal Thr Ser Gl uTyrLeuPheArgHis sSerAspAsnGl
3301 GCTCCTGCACTGGATGGTGGCGTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGTCGCTCCACAAGGTAACAGTTGATTGAACTGCCT
683>uLeuLeuHis sTrpMetVal lAl aLeuAspGl yLysP roLeuAl aSer Gl yGl uVal P roLeuAspVal lAl aP roGl nGl yLysGl nLeu l l eGl uLeuP ro
3401 GAACTACCGCAGCCGGAGAGCGCCGGCAACTCTGGCTCACAGTACGCGTAGTGCAACCGAACCGCCGACATGGTCAGAAGCCGGGCACATCAGCGCCT
717>Gl uLeuP roGl nP roGl uSer Al aGl yGl nLeuTrpLeuThr Val lArgVal Val lGl nP roAsnAl aThr Al aTrpSer Gl uAl aGl yHis l l eSer Al aT
3501 GGCAGCAGTGGCGTCTGGCGGAAACCTCAGTGTGACGCTCCCGCGCTCCACGCCATCCCGCATCTGACCACCAGCGAAATGGATTTTTGCATCGA
750>r pGl nGl nTrpArgLeuAl aGl uAsnLeuSer Val Thr LeuProAl aAl aSer His sAl a l l eP roHis sLeuThr Thr Ser Gl uMe tAspPheCys l l eGl
3601 GCTGGGTAATAAGCGTTGGCAATTTAACCGCCAGTCAGGCTTTCTTTACAGATGTGGATTGGCGATAAAAAACAACCTGCTGACGCCGCTGCGCGATCAG
783>uLeuGl yAsnLysArgTrpGl nPheAsnArgGl nSer Gl yPheLeuSer Gl nMe tTrp l l eGl yAspLysLysGl nLeuLeuThr P roLeuArgAspGl n
3701 TTCACCCGTCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCCGATTGACCCTAACGCCTGGGTCCGAACGCTGGAAAGCGCGCGCCATTACC
817>PheThr ArgAl aP roLeuAspAsnAsp l l eGl yVal Ser Gl uAl aThr Arg l l eAspP roAsnAl aTrpVal l Gl uArgTrpLysAl aAl aGl yHis sTyrG
3801 AGGCCGAAGCAGCGTTGTTGCAGTGCACGGCAGATACACTTGCTGATGCGGTGCTGATTACGACCCTCACGCGTGGCAGCATCAGGGGAAAACTTATT
850> l nAl aGl uAl aAl aLeuLeuGl nCysThr Al aAspThr LeuAl aAspAl aVal lLeu l l eThr Thr Al aHis sAl aTrpGl nHis sGl nGl yLysThr LeuPh
3901 TATCAGCCGGAAAACTACCGATTGATGGTGTGATGTTCAAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGGATACACCGCATCCGCGCGGATTGGC
883>e l l eSer ArgLysThr TyrArg l l eAspGl ySer Gl yGl nMe tAl a l l eThr Val lAspVal l Gl uVal lAl aSer AspThr P roHis sP roAl aArg l l eGl y
4001 CTGAAGTCCAGCTGGCGCAGGTAGCAGAGCGGGTAAACTGGCTCGGATTAGGGCCCAAGAAAATATCCCGACCGCTTACTGCCGCTGTTTTGACC
917>LeuAsnCysGl nLeuAl aGl nVal lAl aGl uArgVal lAsnTrpLeuGl yLeuGl yP roGl nGl uAsnTyrP roAspArgLeuThr Al aAl aCysPheAspA

BbsI (4132)

4101 GCTGGGATCTGCCATTGTCAGACATGTATACCCCGTACGCTTCCCGAGCGAAAACGGTCTCGCTGCGGGACGCGCGAATTGAATTATGGCCACACCA
950>r gTrpAspLeuP roLeuSer AspMetTyrThr P roTyrVal PheP roSer Gl uAsnGl yLeuArgCysGl yThr ArgGl uLeuAsnTyrGl yP roHis sGl
4201 GTGGCGCGGCACTTCCAGTTCAACATCAGCCGCTACAGTCAACAGCAACTGATGAAACAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCACATGG
983>nTrpArgGl yAspPheGl nPheAsn l l eSer ArgTyrSer Gl nGl nGl nLeuMe tGl uThr Ser His sArgHis sLeuLeuHis sAl aGl uGl uGl yThr Trp
4301 CTGAATATCGACGGTTCCATATGGGGATTGGTGGCAGCAGCTCTGGAGCCGTCAGTATCGCGGAATTACAGCTGAGCGCGGTGCTACCATTACC
1017>LeuAsn l l eAspGl yPheHis sMe tGl y l l eGl yGl yAspAspSer TrpSer P roSer Val Ser Al aGl uLeuGl nLeuSer Al aGl yArgTyrHis sTyrG

NheI (4442)

EcoRI (4436)

4401 AGTTGGTCTGGTGTCAAAAATAATACTAGTCGAGAATTCGTAGCTCGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAG
1050> l nLeuVal l TrpCysGl nLys ●●●
4501 TGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAA

MfeI (4616)

DraI (4665)

4601 ACAAGTTAACAAACAACAAATTCATTCTATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAGA

DraI (4704)

SwaI (4707)

4701 TCCATTAAATGTTAAATTAAGTACCCATGACCAAAATCCCTTAACGTGAGTTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTC

4801 TTGAGATCCTTTTTTCTGCGGTAATCTGCTGCTTCAAAACAAAAAACCCGCTACCAGCGGTGTTTGTGGCCGATCAAGAGCTACCAACTCTT

4901 TTTCCGAAGGTAAGTGGCTTACGAGAGCGCAGATACCAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCCTTCAAGAACTCTGTAGCACCGC

5001 CTACATACCTCGCTCTGCTAATCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGTTGGACTCAAGACGATAGTTACCGGATAA

5101 GCGCAGCGGTGGGCTGAACGGGGGTTCTGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA

5201 AGCGCCACGCTTCCGAAGGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCCGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAAACGCCT

5301 GGTATCTTTATAGTCTGTGGGTTTTGCCACCTCTGACTTGAGCGTGCATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGAAAAACGCCAGCAA

AseI (5483)

5401 CGCGGCTTTTTACGGTCTCGGCTTTTGTGGCCTTTTGTCCACATGTTCTTAATTAATTTTTCAAAAGTAGTTGACAAATTAATCATCGGCATAGTA

SfiI (5534)

MseI (5545)

5501 TATCGCATAGTATAATACGACTCACTATAGGAGGGCCATCATGGCCAAGTTGACCAAGTGTCCAGTGTCCAGTGTCCAGCCAGGGATGTGGCTGGAGCTG
1>MetAl aLysLeuThr SerAl aVal lP roVal lLeuThr Al aArgAspVal lAl aGl yAl aV

5600 TTGAGTTCTGGACTGACAGGTTGGGTTCTCCAGAGATTTTGTGGAGGATGACTTTGCGAGTGTGGTCAGAGATGATGTCACCCTGTTCTATCTCAGCAGT
20>al Gl uPheTrpThr AspArgLeuGl yPheSer ArgAspPheVal l Gl uAspAspPheAl aGl yVal lVal lArgAspAspVal lThr LeuPhe l l eSer Al aVa

5700 CCAGGACCAGGTGGTGCCTGACAACACCTGGCTTGGGTGGGTGAGAGGACTGGATGAGCTGTATGCTGAGTGGAGTGGTCTCCACCAACTTC
53> l Gl nAspGl nVal lVal lP roAspAsnThr LeuAl aTrpVal lTrpVal lArgGl yLeuAspGl uLeuTyrAl aGl uTrpSer Gl uVal lVal lSer ThrAsnPhe

5800 AGGGATGCCAGTGGCCTGCCATGACAGAGATTGGAGAGCAGCCCTGGGGAGAGAGTTTGCCTGAGAGACCAGCAGGCAACTGTGTGCACTTTGTGG
87>A rgAspAl aSer Gl yP roAl aMe tThr Gl u l l eGl yGl uGl nP roTrpGl yA rgGl uPheAl aLeuArgAspP roAl aGl yAsnCysVal lHis sPheVal lA

SfiI (5943)

5900 CAGAGGAGCAGGACTGAGGATAAGAATTGAGTTTCAGAAAAGGGGCGCTGAGTGCCCTTTTTTCAACTTAATTA

120> l aGl uGl uGl nAsp●●●