



SdaI (6)
1 CCTGCAGGCGTTACATAA...
NdeI (182)
101 CGCCAATAGGACTTTCCATTGACGTCAATGGTGGAGTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCC
201 TATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC
301 GCTATTACCATGATGATGCGGTTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGATTTCGAAGTCTCCACCCATTGACGTCAATG
SpeI (414)
401 GGAGTTTGTGTTGACTAGTCAGGGCCCCAACCCCCAAGCCCCATTTCACAACACGCTGGCGCTACAGGCGGTGACTTCCCCTTGTGTTGGGGCGGG
BspEI (520)
501 GGGCTGAGACTCCTATGTGCTCCGATTGGTCAGGCACGGCCTTCGGCCCCCTCTGCCACCGAGATTGGCCGCTAGGCCTCCCCGAGCGCCCTGCC
HindIII (679)
601 TCCGAGGGCCGGCGCACCATAAAAGAAGCCGCTAGCCACGTCCTCGCAGTTGGCGGTCCCGGGTCTGTCTCAAGCTTGCCGCCAGAACACAGg
XmnI (763)
701 taagtccgtgtgtggttcccgcgggctggcctctttacgggttatggccttgcgtgcctgaattacttccatgccctggctgcagtacgtgattc
801 ttgatcccagacttccgggttgaagtgggtgggagagttcgaggccttgcgcttaaggagcccttgcctcgtgcttgagttgaggcctggcttggggc
901 ctggggccgcccgtgctaactctggtggcaccttgcgcttctcgtgctttcgctaagtctctagccatttaaaattttgataaccagctgcgagc
BglIII (1039)
1001 ctttttttctggcgagatagcttctgtaaatgcgggccaagatctgcacactgggtatttccggttttggggccgcccggcgagccggcccgctgcgtccc
1101 agcgacacatgctcggcgaggcggggcctgcgagcggccaccaggagaatcggacggggtagtctcaaacggccgctgctggtgcttggcctgcg
EcoNI (1290)
1201 gccgccgtgatcgccccctggcggaaggctggcccgtcggcaccagttgctgtagcggaagatggccgcttcccggcctgctgcagggagc
SmaI (1323)
1301 tcaaaatggaggacgcggcgcccgggagagcgggctgagtcaccacacaaaggaaaaggcccttctcctcatcctgcttcatgtgactcca
XhoI (1438)
1401 cggagtagcgggcccgtccaggcacctcgattagttctcgagcttttgagtagctcgtcttttaggttgggggaggggttttatgcatggagtttcc
1501 ccacactgagtggggtggagactgaagagttagggcagcttggcacttgatgtaattctccttggaaattgcccttttgagtttggatcttgcctcatc
NotI (1673)
1601 tcaagcctcagacagtggttcaaagttttttcttccatttcagGTGTCGTGAAACTACCCTAAAAGCCACCATGGGGTTCTCATCATCATCA
MetGI yGI ySer Hi sHi sHi sHi
NheI (1711) **Acc65I (1767)**
1701 TCATGGTATGGCTAGCATGACTGGTGACAGCAAATGGTCCGGATCTGTACGACGATGACGATAAGGTACCTAAGGATCAGCTTGGAGTTGATCCCGTC
9> sHi sGI yMetAl aSerMetThr GI yGI yGI nGI nMetGI yArgAspLeuTyrAspAspAspLysVal IP roLysAspGI nLeuGI yVal AspProVal
1801 GTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCAACTTAATCGCCTTGACGACATCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCC
43> Val LeuGI nArgArgAspTrpGI uAsnProGI yVal Thr GI nLeuAsnArgLeuAl aAl aHi sProProPheAl aSer TrpArgAsnSer GI uGI uAl aA
FspI (1927)
1901 GCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTCCTGGTTCCGGCACCAGAAGCGGTGCCGAAAGCTGGCTGGAGTG
76> rgThrAspArgProSer GI nGI nLeuArgSer LeuAsnGI yGI uTrpArgPheAl aTrpPheProAl aProGI uAl aVal IP roGI uSer TrpLeuGI uCy
2001 CGATCTTCTGAGCCGATCTGCTGCTCCCTCAAACCTGGCAGATGCACGTTACGATGCGCCATACACCAACGTAACCTATCCCATTACGGTC
109> sAspLeuProGI uAl aAspThr Val Val Val IP roSerAsnTrpGI nMetHi sGI yTyrAspAl aProI l eTyrThrAsnVal Thr TyrProI l eThr Val
2101 AATCCGCGTGTGTTCCACGAGAAATCCGACGGTGTACTCGCTCACATTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTT
143> AsnProProPheVal IP roThr GI uAsnProThr GI yCysTyrSer LeuThr PheAsnVal AspGI uSer TrpLeuGI nGI uGI yGI nThr ArgI l eI l eP
2201 TTGATGGCGTTAACTCGGCCTTTCATCTGTGGTCAACGGCGCTGGGTACGGCCAGGACAGTCGTTTCCGCTGTAATTTGACCTGAGCGCATT
176> heAspGI yVal lAsnSerAl aPheHi sLeuTrpCysAsnGI yArgTrpVal GI yTyrGI yGI nAspSerArgLeuProSer GI uPheAspLeuSerAl aPh
2301 TTTACGCGCCGGAGAAAACCGCCTCGCGTGATGGTGTGCGTTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGATGAGCGCATTTTC
209> eLeuArgAl aGI yGI uAsnArgLeuAl aVal MetVal LeuArgTrpSerAspGI ySer TyrLeuGI uAspGI nAspMetTrpArgMetSer GI yI l ePhe
2401 CGTGACGTCTGTTGCTGCATAAACCGACTACACAAATCAGCGATTTCCATGTTGCCACTCGCTTAAATGATGATTTACGCCGCTGACTGGAGGCTG
243> yGI uI l eI l eAspGI uArgGI yGI yTyrAl aAspArgVal Thr LeuArgLeuAsnVal GI uAsnProLysLeuTrpSerAl aGI uI l eP roAsnLeuTyr
2501 AAGTTCAGATGTGCGCGAGTTGCGTGACTACCTACGGTAACAGTTTCTTTATGGCAGGTTGAAACGCAGGTCGCCAGCGCACCGCCTTTCGGCGG
276> l uVal GI nMetCysGI yGI uLeuArgAspTyrLeuArgVal Thr Val Ser LeuTrpGI nGI yGI uThr GI nVal lAl aSer GI yThrAl aProPheGI yGI
2601 TGAATATCGATGAGCGTGGTGTATGCCGATCGCGTACACTACGCTGAACGTCGAAAACCCGAAACTGTGGAGCGCCGAAATCCCGAATCTCTAT
309> yGI uI l eI l eAspGI uArgGI yGI yTyrAl aAspArgVal Thr LeuArgLeuAsnVal GI uAsnProLysLeuTrpSerAl aGI uI l eP roAsnLeuTyr
2701 CGTGCGGTGGTTGAACTGCACACCGCGACGGCAGCTGATTGAAGCAGAAGCTGCGATGTCGTTTCCGCGAGGTGCGGATTGAAATGGTCTGCTGC
343> ArgAl aVal Val GI uLeuHi sThrAl aAspGI yThr LeuI l eGI uAl aGI uAl aCysAspVal GI yPheArgGI uValArgI l eGI uAsnGI yLeuLeu
EcoRV (2)
2801 TGCTGAACGGCAAGCCGTTGCTGATTGAGGCGTTAACCGTCACGAGCATCATCTCTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGAT
376> euLeuAsnGI yLysProLeuLeuI l eArgGI yVal lAsnArgHi sGI uHi sHi sProLeuHi sGI yGI nVal MetAspGI uGI nThr MetVal GI nAspI l
2901 CCTGCTGATGAAGCAGAACAACCTTAAACCGCTGCGCTGTTCCGATTATCGAACCATCCGCTGTGGTACACGCTGTGCGACCGCTACGGCCTGTATGTG
409> eLeuLeuMetLysGI nAsnAsnPheAsnAl aVal lArgCysSerHi sTyrProAsnHi sProLeuTrpTyrThr LeuCysAspArgTyrGI yLeuTyrVal
3001 GTGGATGAAGCAATATTGAAACCCACGGCATGGTGCCAAATGAATCGTCTGACCGATGATCCGCGTGGCTACCGCGATGAGCGAACCGGTAACCGGAA

443. Val AspGlu Ala Asn1 LeuGlu Thr His Glu Met Val Pro Met Asn Arg Leu Thr Asp Asp Pro Arg Trp Leu Pro Ala Met Ser Glu Arg Val Thr Arg M
3101 TGGTGCAGCGCATCGTAATCACCCGAGTGTGATCATCTGGTCGCTGGGAAATGAATCAGGCCACGGCGTAATCACGACGCGTGTATCGCTGGATCAA
476▶ et Val Glu N Arg Asp Arg Asn His s Pro Ser Val I I e I e Trp Ser Leu Glu y Asn Glu u Ser Glu y Hi s Glu y Al a Asn His s Asp Al a Leu Tyr Arg Trp I e Ly
3201 ATCTGTGATCCTTCCCGCCCGGTGCAGTATGAAGCGCGGAGCCGACACCGCCACCGATATTATTGGCCGATGTACGCGCGGTGGATGAAGAC
509▶ s Ser Val Asp Pro Ser Arg Pro Val Glu N Tyr Glu u Glu y Glu y Al a Asp Thr Thr Al a Thr Asp I e I e Cys P ro Me t Tyr Al a Arg Val Asp Glu u Asp
3301 CAGCCCTTCCCGCTGTGCCGAAATGGTCCATCAAAAAATGGCTTTCGCTACCTGGAGAGACGCGCCCGTGTATCCTTTGCGAATACGCCACGCGCATGG
543▶ Glu N Pro Phe Pro Ala Val P ro Lys Trp Ser I I e Lys Lys Trp Leu Ser Leu P ro Glu y Glu u Thr Arg P ro Leu I e Leu Cys Glu u Tyr Al a Hi s Al a Me t G
3401 GTAACAGTCTTGGCGTTCGCTAAATAC TGGCAGGCGTTCCTGTCAGTATCCCGTTTACAGGGCGGCTTCTGCTGGGACTGGGTGGATCAGTCGCTGAT
576▶ I y Asn Ser Leu Glu y Glu y Phe Al a Lys Tyr Trp Glu N Al a Phe Arg Glu N Tyr P ro Arg Leu Glu N Glu y Glu y Phe Val Trp Asp Trp Val Asp Glu N Ser Leu I
3501 TAAATATGATGAAAACGGCAACCCGTGGTCGGCTTACGGCGGTGATTTTGGCGATACGCCAAGCATCGCCAGTCTGTATGAACGGTCTGGTCTTTGCC
609▶ e Lys Tyr Asp Glu u Asn Glu y Asn P ro Trp Ser Al a Tyr Glu y Glu y Asp Phe Glu y Asp Thr P ro Asn Asp Arg Glu N Phe Cys Me t Asn Glu y Leu Val I Phe Al a

Eco47III (3620)

3601 GACCGCACGCCGCATCCAGCGCTGACGGAAAGCAAACACCAGCAGCAGTTTTTCCAGTTCGGTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATACC
643▶ Asp Arg Thr P ro Hi s P ro Al a Leu Thr Glu u Al a Lys Hi s Glu N Glu N Phe Phe Glu N Phe Arg Leu Ser Glu y Glu N Thr I I e Glu u Val Thr Ser Glu u Tyr L
3701 TGTCCGTCATAGCGATAACGAGCTCTGCAGCTGGATGGTGGCGTGGTAAAGCCGCTGGCAAGCGGTGAAGTCCCTCGGATGTCCGCTCACAAGG
676▶ eu Phe Arg Hi s Ser Asp Asn Glu u Leu Leu Hi s Trp Met Val Al a Leu Asp Glu y Lys P ro Leu Al a Ser Glu y Glu u Val P ro Leu Asp Val Al a P ro Glu N Glu
3801 TAAACAGTTGATTGAACTGCCTGAACTACCGCAGCCGAGAGCGCCGGCAACTCTGGCTCAGATACGGTAGTGCAACCGAACGCGACCGCATGGTCA
709▶ y Lys Glu N Leu I I e Glu u Leu P ro Glu u Leu P ro Glu N P ro Glu u Ser Al a Glu y Glu N Leu Trp Leu Thr Val A Arg Val Val Glu N P ro Asn Al a Thr Al a Trp Ser
3901 GAAGCCGGGCACATCAGCGCTGGCAGCAGTGGCGTCTGGCGAAAACCTCAGTGTGACGCTCCCGCCGCTCCACGCCATCCCGCATCTGACCACA
743▶ Glu Al a Glu y Hi s I I e Ser Al a Trp Glu N Glu N Trp Arg Leu Al a Glu u Asn Leu Ser Val Trp Leu P ro Al a Al a Ser Hi s Al a I e P ro Hi s Leu Thr S
4001 GCGAAATGGATTTTGCATCGAGCTGGTAATAAGCGTTGGCAATTTAACCGCCAGTCAGGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACA
776▶ er Glu u Me t Asp Phe Cys I I e Glu u Leu Glu y Asn Lys Arg Trp Glu N Phe Asn Arg Glu N Ser Glu y Phe Leu Ser Glu N Me t Trp I I e Glu y Asp Lys Lys Glu N Le
4101 GCTGACGCCGCTGCGCATCAGTTACCCGTCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCCGATTGACCCTAACGCCTGGTCCGAACGC
809▶ u Leu Thr P ro Leu Arg Asp Glu N Phe Thr Arg Al a P ro Leu Asp Asn Asp I I e Glu y Val Ser Glu u Al a Thr Arg I I e Asp P ro Asn Al a Trp Val Glu u Arg
4201 TGGAAAGCGCGGGCCATTACAGCCGAAGCAGCGTTGTTGCAGTGACCGCAGATACACTTGCTGATGCGGTGCTGATTACGACCGCTCACGCGTGGC
843▶ Trp Lys Al a Al a Glu y Hi s Tyr Glu N Al a Glu u Al a Al a Leu Leu Glu N Cys Thr Al a Asp Thr Leu Al a Asp Al a Val I Leu I e Thr Thr Al a Hi s Al a Trp G
4301 AGCATCAGGGAAAACCTTATTTATCAGCCGAAACCTACCGGATTGATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATAC
876▶ I n Hi s Glu N Glu y Lys Thr Leu Phe I I e Ser Arg Lys Thr Tyr Arg I I e Asp Glu y Ser Glu y Glu N Me t Al a I I e Thr Val Asp Val Glu u Val Al a Ser Asp Th
4401 ACCGCATCCGGCGCGGATTGGCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTAAACTGGCTCGGATTAGGGCCGCAAGAAAACCTATCCCGACCGC
909▶ r P ro Hi s P ro Al a Arg I I e Glu y Leu Asn Cys Glu N Leu Al a Glu N Val Al a Al a Glu u Arg Val Asn Trp Leu Glu y Leu Glu y P ro Glu N Glu u Asn Tyr P ro Asp Arg
4501 CTTACTGCCGCTGTTTTGACCGCTGGGATCTGCCATTGTCAGACATGTATACCCGCTACGCTCTCCGAGCGAAAACCGTCTGCGCTGCGGGACGCGCG
943▶ Leu Thr Al a Al a Cys Phe Asp Arg Trp Asp Leu P ro Leu Ser Asp Me t Tyr Thr P ro Tyr Al a Phe P ro Ser Glu u Asn Glu y Leu Arg Cys Glu y Thr Arg G
4601 AATTGAATTATGGCCACACCGTGGCGGCGACTTCCAGTCAACCTCAGCCGCTACAGTCAAGCAAGCACTGATGGAALCAGCCATCGCCATCTGCT
976▶ I u Leu Asn Tyr Glu y P ro Hi s Glu N Trp Arg Glu y Asp Phe Glu N Phe Asn I I e Ser Arg Tyr Ser Glu N Glu N Leu Me t Glu u Thr Ser Hi s Arg Hi s Leu Le

NdeI (4742)

4701 GCACGCGGAAGAAGGCACATGGCTGAATATCGACGGTTTCCATATGGGGATTGGTGGCGACGACTCTGGAGCCCGTCACTATCGGGGAATTACAGCTG
1009▶ u Hi s Al a Glu u Glu y Thr Trp Leu Asn I I e Asp Glu y Phe Hi s Me t Glu y I I e Glu y Glu y Asp Asp Ser Trp Ser P ro Ser Val Ser Al a Glu u Leu Glu N Leu

NheI (4864)

EcoRI (4858)

4801 AGCGCCGGTCTACCATACCAGTTGGTCTGGTGTCAAAAAATAATAATCTAGTCGAGAATTCGCTAGCTCGACATGATAAGATACATTGATGAGTTTGG
1043▶ Ser Al a Glu y Arg Tyr Hi s Tyr Glu N Leu Val Trp Cys Glu N Lys ●●●
4901 ACAAAACCAACTAGAATGCAAGTGAAAAAATGCTTTATTTGGAAATTTGTGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTG

MfeI (5038)

5001 TAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACATTCATTATTTATGTTTTAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTA

SwaI (5129)

5101 AAACCTCTACAAATGTGGTAGATCCATTTAAATGTTAATTAAGTACGATGACCAAATCCCTTAACGTGAGTTTTCTGTTCCACTGAGCGTCAGACCCCG
5201 TAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGGTAATCTGCTGCTTGCACAAAACAAAACACCCGCTACCAGCGGTGGTTTGTGTTGCC
5301 GGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCAC
5401 TTCAAGAACTCTGTAGACCCGCTACATACCTCGCTCTGCTAATCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCAGGTTGGACT
5501 CAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGCTGAACGGGGGTTCTGTCACACAGCCAGCTTGGAGCGAACGACTACACCGAACTGAGATA
5601 CCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCCGAACAGGAGAGCGCACAGG
5701 GAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCTGCTGCGGTTTCCGACCTCTGACTTGAGCGTCAATTTTGTGATGCTCGTCAAGGGGGCGGA
5801 GCCTATGGAACCGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGTGCGCTTTTGTCTACATGTTCTTAATTAATTTTTCAAAGTAGTTG

AseI (5905)

MseI (5967)

5901 ACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACTCACTATAGGAGGGCCATCATGGCCAAGTTGACCGAGTGTGCCAGTGCTCACAG
6001 CCAGGGATGTGGCTGGAGCTGTTGAGTTCTGGACTGACAGGTTGGGTTCTCCAGAGATTTGTGGAGGATGACTTTGCAGGTGTGGTCCAGAGATGATGT
13▶ I a Arg Asp Val Al a Glu y Al a Val Glu u Phe Trp Thr Asp Arg Leu Glu y Phe Ser Arg Asp Phe Val Glu u Asp Asp Phe Al a Glu y Val Val A Arg Asp Asp Va
6101 CACCTGTTTCATCTCAGCAGTCCAGGACAGTGGTCCCTGACAACCCCTGGCTGGGTGGGTGAGAGGACTGGATGAGCTGATGCTGAGTGGAGT
46▶ I Thr Leu Phe I I e Ser Al a Val Glu N Asp Glu N Val Val P ro Asp Asn Thr Leu Al a Trp Val Trp Val A Arg Glu y Leu Asp Glu u Leu Tyr Al a Glu u Trp Ser
6201 GAGGTGGTCTCCACCAACTTCAGGATGCGAGTGGCCCTGCCATGACAGAGATTGGAGAGCAGCCCTGGGGGAGAGAGTTTGGCTGAGAGACCCAGCAG
80▶ Glu u Val Val Ser Thr Asn Phe Arg Asp Al a Ser Glu y P ro Al a Me t Thr Glu u I I e Glu y Glu u N P ro Trp Glu y Arg Glu u Phe Al a Leu Arg Asp P ro Al a G
6301 GCAACTGTGCTACTTTGTGGCAGAGGAGCAGGACTGAGGATAAGAATTGAGTTTCAAGAAAAGGGGCGTGGTGGCCCTTTTCAACTTAATTA
113▶ I y Asn Cys Val I Hi s Phe Val Al a Glu u Glu N Asp ●●●