



**SdaI (6)**  
1 CCTGCAGGGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCTGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAA  
101 AGTCCCCAGGCTCCCCAGCAGGCAGAAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAAG

**NotI (276)**  
201 TATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCACTAGTTCGCCAGAGCGCGGAGGGCCTCCAGCGGCCGCCCTCCCCACAGCAGGG  
301 GCGGGGTCCCCGCCACCAGGAAAGGAGCGGGCTCGGGCGGGCGGCGCTGATTGGCCGGGGCGGCTGACGCCGACGCGGTATAAGAGACCACAAGCG

**HindIII (428)** **SacII (472)**  
401 ACCCGCAGGGCCAGAGCTTCTTCGCCAAGCTTCCCGTCAGAACGCAGGTGAGGGGGGGTGTGGCTTCCGCGGGCCGCCGAGCTGGAGTCTGCTCCG  
501 AGCGGGCCGGGCCCTGTCTGCGGGGATTAGCTGCGAGCATTCCCGTTCGAGTTGCGGGCGGCGGGAGGCAGAGTGCAGGCCTAGCGGCAA  
601 CCCCAGTGCCTCGCTCGTTCGGCTTAGAGCCTAGCGTGTGTCCGCGCCGCCCGCGTGTACTCCGGCCGCACTCTGGTCTTTTTTTTTTTTTGTT  
701 GTTGTTCCTGCTGCCTTCGATTGCCGTTACGCAATAGGGGTAACAAAGGAGGGTGCGGGGCTTGTCTGCCCGGAGCCCGAGAGGTCATGGTTGGG  
801 GAGGAATGGAGGGACAGGAGTGGCGGCTGGGGCCCGCCGCTTCCGAGCACATGTCGACGCCACTGGATGGGGCGAGGCCTGGGTTTTTCCCGAAG  
901 CAACCAGGCTGGGTTAGCGTGCAGGCCATGTGGCCCCAGCACCGCAGCATCTGGCTTGGCGGCGCGGTTGCCCTGCCCTAACCTAGGGTGA

**SmaI (1049)**  
1001 GGCCATCCCGTCCGGCACCAGTTGCGTGCCTGAAAGATGGCCGCTCCCGGGCCTGTTGCAAGGAGCTCAAATGGAGGACGCGGCAGCCCGGTGGAGC  
1101 GGGCGGGTGTAGTACCACACAAAGGAAGAGGGCCTGGTCCCTACCGGCTGCTGCTTCTGTGACCCCGTGGTCTATCGGCCAATAGTACCTCGG  
1201 GCTTTTGTAGCAGCGCTAGTCGGCGGGGGAGGGGATGTAATGGCGTTGGAGTTTGTTCACATTTGGTGGTGGAGACTAGTACGGCCAGCTGGCGCT

**EcoNI (1391)**  
1301 GGAAGTCATTTTTGGAATTTGTCCCTTGAGTTTTGAGCGGAGCTAATTCTCGGGCTTCTTAGCGGTTCAAAGGTATCTTTAAACCCTTTTTAGGTTG

**BspHI (1429)** **NheI (1467)**  
1401 TGTGAAAACCACCGCTAATCAAAGCAATCATGACGGTTCATCATCATCATCATCATGGTATGGTACGATGACTGGTGGACAGCAATGGTCCGGG  
1501 ATCTGTACGACGATGACGATAAGGTACCTAAGGATCAGCTTGGAGTTGATCCCGTGTGTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCA  
24> spLeuTyrAspAspAspAspLysValProLysAspGlnLeuGlyValAspProValValLeuGlnArgArgAspTrpGluAsnProlGlyValThrGlnLe  
1601 TAATCGCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCATCGCCCTCCCAACAGTTGCGCAGCCTGAATGGCGAA  
57> uAsnArgLeuAlaAlaHisProProPheAlaSerTrpArgAsnSerGluGluAlaArgThrAspArgProSerGlnGlnLeuArgSerLeuAsnGlyGlu  
1701 TGGCGCTTTGCTGGTTCCGGCACCAGAAGCGGTGCCGAAAGCTGGCTGGAGTGCATCTTCTGAGGCGGATACTGTCTGCTCCCTCAAAGTGGC  
91> TrpArgPheAlaTrpPheProAlaProGluAlaValProGluSerTrpLeuGluCysAspLeuProGluAlaAspThrValValValProSerAsnTrpG  
1801 AGATGCACGGTTACGATGCGCCATCTACACCAACGTAACCTATCCATTACGGTCAATCCGCCGTTTGTCCACGGAGAATCCGACGGGTTGTTACTC  
124> lNMetHisSglyTyrAspAlaProlIeTyrThrAsnValThrTyrProlIeThrValAsnProPheValProThrGluAsnPProThrGlyCysTyrSe  
1901 GCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCAATATTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGAACGGGCGC  
157> rLeuThrPheAsnValAspGluSerTrpLeuGlnGluGlyNThrArgIleIlePheAspGlyValAsnSerAlaPheHisLeuTrpCysAsnGlyArg  
2001 TGGTTCGGTTACGGCCAGGACAGTCTGTTGCCGTTGAAATTTGACCTGAGCGCATTTTTACGCGCCGAGAAAACCGCCTCGCGGTGATGGTCTGCGTT  
191> TrpValGlyTyrGlyGlnAspSerArgLeuProSerGluPheAspLeuSerAlaPheLeuArgAlaGlyGluAsnArgLeuAlaValMetValLeuArgT  
2101 GGAGTACGGCAGTTATCTGGAAGATCAGGATGTGGCGGATGAGCGGCATTTTCCGTCAGTCTGTTGTCATAAACCGACTACACAAATCAGCGA  
224> rPserAspGlySerTyrLeuGluAspGluAspMetTrpArgMetSerGlyIlePheArgAspValSerLeuLeuHisLysProThrGlnIleSerAs  
2201 TTTCCATGTTGCCACTCGCTTAATGATGATTTACGCCGCTGTACTGGAGGCTGAAGTTCAGATGTGGCGGAGTTGCGTGACTACCTACGGGTAACA  
257> pPheHisValAlaThrArgPheAsnAspPheSerArgAlaValLeuGluAlaGluValGlnMetCysGlyGluLeuArgAspTyrLeuArgValThr  
2301 GTTCTTTATGGCAGGGTAAACGACGGTCCGACGGCACCAGCGCCTTTCGGCGGTGAAATTCATGATGAGCGTGGTGGTTATGCCGATCGCGTCACAC  
291> ValSerLeuTrpGlnGlyGluThrGlnValAlaSerGlyThrAlaProPheGlyGlyGluIleIleAspGluArgGlyGlyTyrAlaAspArgValThrL  
2401 TACGTCTGAACGTCGAAAACCGAAACTGTGGAGCGCCGAAATCCGAACTCTATCTGTCGGTGGTTGAACTGCACACCGCCGACGCGCAGTATTGA  
324> euArgLeuAsnValGluAsnProLysLeuTrpSerAlaGluIleProAsnLeuTyrArgAlaValValGluLeuHisThrAlaAspGlyThrLeuIleG  
2501 AGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTGCGGATTGAAATGGTCTGCTGCTGCTGAACGGCAAGCGGTTGCTGATTCGAGGCGTTAACCGTAC  
357> uAlaGluAlaCysAspValGlyPheArgGluValArgIleGluAsnGlyLeuLeuLeuLeuAsnGlyLysProLeuLeuIleArgGlyValAsnArgHis

**EcoRV (2654)**  
2601 GAGCATCATCCTCTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGATATCTGCTGATGAAGCAGAACAACCTTAAACGCCGTGGCTGTTCCG  
391> GluHisHisSProLeuHisSglyGlyNValMetAspGluGlnThrMetValGlnAspIleLeuLeuMetLysGlnAsnAsnPheAsnAlaValArgCysSerH  
2701 ATTATCCGAACCATCCGCTGTGGTACAGCTGTGCGACCGCTACGGCCTGTATGTGGTGGATGAAGCAAATTTGAAACCCACGGCATGGTCCAATGAA  
424> iS TyrProAsnHisSProLeuTrpTyrThrLeuCysAspArgTyrGlyLeuTyrValValAspGluAlaAsnIleGluThrHisSglyYMetValProMeTAs  
2801 TCGTCTGACCGATGATCCGCGCTGGCTACCGCGGATGAGCGAACCGTAACGCGAATGGTGCAGCGCGATCGTAATACCCGAGTGTGATCATCTGGTCCG  
457> nArgLeuThrAspAspProArgTrpLeuProAlaMetSerGluArgValThrArgMetValGlnArgAspArgAsnHisSProSerValIleIleTrpSer  
2901 CTGGGGAATGAATCAGGCCACGGCCTAATCACGACCGCTGTATCGCTGGATCAAATCTGTGATCCTTCCCGCCCGTGCAGTATGAAGCGCGGGAG  
491> LeuGlyAsnGluSerGlyHisSglyAlaAsnHisAspAlaLeuTyrArgTrpIleLysSerValAspProSerArgProValGlnTyrGluGlyGlyYA  
3001 CCGACACACGGCCACCGATATTATTTGCCGATGTACGCGCGGCTGGATGAGAACCGCCCTTCCCGGCTGTGCCGAAATGGTCCATCAAAAAATGGCT  
524> lAspThrThrAlaThrAspIleIleCysProMetTyrAlaArgValAspGluAspGlnProPheProAlaValProLysTrpSerIleLysLysTrpLe  
3101 TTCGCTACCTGGAGAGACGCGCCGCTGATCCTTTGCGAATACGCCACGCGATGGGTAACAGTCTTGGCGGTTTCGCTAAATACTGGCAGGGCTTTCGT  
557> uSerLeuProGlyGluThrArgProLeuIleLeuCysGlyTyrAlaHisAlaMetGlyAsnSerLeuGlyPheAlaLysTyrTrpGlnAlaPheArg  
3201 CAGTATCCCGTTTACAGGCGGCTTCGCTGGGATGGGATCAGTCTGATTAAATGATGAAAACGGCAACCCGTTGGTTCGGCTTACGGCGGCTG  
591> GlnTyrProArgLeuGlnGlyGlyPheValTrpAspTrpValAspGlnSerLeuIleLysTyrAspGluAsnGlyAsnPProTrpSerAlaTyrGlyGlyYA

**Eco47III (3376)**  
3301 ATTTTGGCGATACGCCAACGATCGCCAGTTCTGTATGAACGGTCTGGTCTTTGCCGACCGCACGCCGATCCAGCGCTGACGGAAGCAAACACCGACA

3401 GCAGTTTTCCAGTTCCGTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGATAACGAGTCCTGCACTGGATGGTGGCG  
657 nGlnPhePheGlnPheArgLeuSerGlyGlnThrIleGluValThrSerGlyTyrLeuPheArgHisSerAspAsnGluLeuLeuHisTrpMetValAla  
3501 CTGGATGGTAAGCCGTGGCAAGCGGTGAAGTGCTCTGGATGTCGCTCCACAAGGTAAACAGTTGATTGAAGTCCCTGAACTACCGCAGCCGAGAGCG  
691 LeuAspGlyLysProLeuAlaSerGlyGluValProLeuAspValAlaProGlnGlyLysGlnLeuIleGluLeuProGluLeuProGlnProGluSerA  
3601 CCGGGCAACTCTGGCTCACAGTACCGTAGTGCAACCGAACGCGCCGATGGTCAGAAGCCGGGCACATCAGCGCTGGCAGCAGTGGCGTCTGGCGGA  
724 IAGlyGlnLeuTrpLeuThrValArgValValGlnProAsnAlaThrAlaTrpSerGluAlaGlyHisIleSerAlaTrpGlnGlnTrpArgLeuAlaGly  
3701 AAACCTCAGTGTGACGCTCCCGCCGCTCCACGCCATCCCGCATCTGACCACCAGCGAAATGGATTTTGCATCGAGCTGGTAATAAGCGTTGGCAA  
757 uAsnLeuSerValThrLeuProAlaAlaSerHisAlaIleProHisLeuThrThrSerGluMetAspPheCysIleGluLeuGlyAsnLysArgTrpGln  
3801 TTTAACCGCCAGTCAGGCTTCTTTCACAGATGTGGATTGGCGATAAAAAACAACCTGCTGACGCCGCTGCGCGATCAGTTCACCCGTGCCCGCTGGATA  
791 PheAsnArgGlnSerGlyPheLeuSerGlnMetTrpIleGlyAspLysLysGlnLeuLeuThrProLeuArgAspGlnPheThrArgAlaProLeuAspA  
3901 ACGACATTGGCGTAAGTGAAGCGACCCGATTGACCCTAACGCCCTGGTTCGAACGCTGGAAGCGGGCGGCCATTACCAGGCCGAAGCAGCGTTGTTGCA  
824 snAspIleGlyValSerGluAlaThrArgIleAspProAsnAlaTrpValGluArgTrpLysAlaAlaGlyHisTyrGlnAlaGluAlaAlaLeuLeuGly  
4001 GTGCACGGCAGATACACTTGTGTATGCGGTGCTGATTACGCCGCTCACGCTGGCAGCATCAGGGGAAAACCTTATTTATCAGCCGAAAACCTACCGG  
857 nCysThrAlaAspThrLeuAlaAspAlaValLeuIleThrThrAlaHisAlaTrpGlnHisGlnGlyLysThrLeuPheIleSerArgLysThrTrpArg  
4101 ATTGATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATACCCGCATCCGGCGCGGATTGGCCTGAAGTCCAGCTGGCGCAGG  
891 IleAspGlySerGlyGlnMetAlaIleThrValAspValGluValAlaSerAspThrProHisProAlaArgIleGlyLeuAsnCysGlnLeuAlaGlnV  
4201 TAGCAGAGCGGGTAAACTGGCTCGGATTAGGGCCGCAAGAAAATATCCCGACCGCCTTACTGCCGCTGTTTTGACCGCTGGATCTGCCATTGTCAGA  
924 AlaAlaGluArgValAsnTrpLeuGlyLeuGlyProGlnGluAsnTyrProAspArgLeuThrAlaAlaCysPheAspArgTrpAspLeuProLeuSerAs  
4301 CATGTATACCCGTCAGTCTTCCCGAGCGAAAACGCTGCGCTGCGGGACGCGCAATTAATTGATGGCCACACCAAGTGGCCGCGGCTCCAGTTC  
957 pMetTyrThrProTyrValPheProSerGluAsnGlyLeuArgCysGlyThrArgGluLeuAsnTyrGlyProHisGlnTrpArgGlyAspPheGlnPhe

NdeI (44)

4401 AACATCAGCCGCTACAGTCAACAGCAACTGATGGAAACCGCCATCGCATCTGCTGCACGCGGAAGAAGGCACATGGCTGAATATCGACGGTTCCATA  
991 AsnIleSerArgTyrSerGlnGlnGluMetGluThrSerHisArgHisLeuLeuHisAlaGluGlyThrTrpLeuAsnIleAspGlyPheHisM  
4501 TGGGGATTGGTGGCGAGCTCCTGGAGCCGTCAGTATCGGCGAATTACAGCTGAGCGCCGCTGCTACCATTACCAGTTGGTCTGGTGTCAAAAATA  
1024 eTgylleGlyGlyAspAspSerTrpSerProSerValSerAlaGluLeuGlnLeuSerAlaGlyArgTyrHisTyrGlnLeuValTrpCysGlnLys••

NheI (4620)

EcoRI (4614)

4601 ATAATCTAGTCGAGAATTCGCTAGCTCGACATGATAAGATACATTGATGAGTTTGGACAAAACCACAAGTGAATGCAAGTGAATAAATGCTTTATTTGTG  
1057 •  
4701 AAATTTGTGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGC

SwaI (4885)

4801 ATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTAGATCCATTTAAATGTTAATTAAC  
4901 AGCCATGACAAAAATCCCTAACGTGAGTTTTCTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGC  
5001 GTAATCTGCTGCTTGCAAACAAAAAACCCGCTACCAGCGGTGTTTGTGTTGCCGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCA  
5101 GCAGAGCGCAGATACAAACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCCTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAAT  
5201 CCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCAGGTTGGACTCAAGACGATAGTTACCGGATAAAGCGCAGCGGTGGGCTGAACG  
5301 GGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACCCGAAGTACAGCTACAGCGTATGAGAAAGCGCCAGCTTCCGAAGGGA  
5401 GAAAGGCGACAGGTATCCGTAAGCGGCAGGGTCCGAACAGGAGAGCGACGAGGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGTCTGTCCG  
5501 GTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCTATGGAAAACGCCAGCAACGCGCCTTTTTACGGTCTCTG

AseI (5661)

5601 GCCTTTTGTGGCCTTTTGTCTCACATGTTCTTAATTAATTTTTCAAAGTAGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGAC

MscI (5723)

5701 TCACTATAGGAGGCCATCATGGCCAAGTTGACCAAGTGTCCAGTGTCTCACAGCCAGGGATGTGGCTGGAGCTGTTGAGTTCTGGACTGACAGGTTG  
1 MeTalaLysLeuThrSerAlaValProValLeuThrAlaArgAspValAlaGlyAlaValGluPheTrpThrAspArgLeu  
5801 GGGTTCTCCAGAGATTTTGTGGAGGATGACTTTGCAAGTGTGGTCAGAGATGATGTACCCCTGTTTCATCTCAGCAGTCCAGGACAGGTGGTGCCTGACA  
28 GlyPheSerArgAspPheValGluAspAspPheAlaGlyValValArgAspAspValThrLeuPheIleSerAlaValGlnAspGlnValValProAspA  
5901 ACACCCTGGCTTGGGTGTGGGTGAGAGGACTGGATGAGCTGTATGCTGAGTGGAGTGGAGTGGTCTCCACCAACTTCAGGGATGCCAGTGGCCCTGCCAT  
61 snThrLeuAlaTrpValTrpValArgGlyLeuAspGluLeuTyrAlaGluTrpSerGluValValSerThrAsnPheArgAspAlaSerGlyProAlaMe  
6001 GACAGAGATTGGAGAGCAGCCCTGGGGAGAGAGTTTGCCTGAGAGACCCAGCAGGCAACTGTGTGCACCTTTGTGGCAGAGGAGCAGGACTGAGGATAA  
94 tThrGluIleGlyGluGlnProTrpGlyArgGluPheAlaLeuArgAspProAlaGlyAsnCysValHisPheValAlaGluGluGlnAsp•••  
6101 GAATTGAGTTTCAGAAAAGGGGGCTGAGTGGCCCTTTTTTCAACTTAATTA