



EcoRI (23)

EcoRV (17)

NotI (2) **SalI (9)** **XbaI (19)** **SdaI (38)** **SpeI (45)**

1 **GCGGCCGCGT**CGACGATATCTAGAATTCGGATCCTGCAGGGCCCACTAGTAGCTCTGAGCTCCTCCTCTGCTCG

75 **CCCAATCCTTCCAACCCCTATGGTGGTATGGCTGACACAGAAAATGTCTGCTCCTGTATGGGACATTTGCCCC**

149 **TCTTCTCCAAATATAAGACAGGATGAGGCCTAGCTTTTTGCTGCTCCAAAGTTTTAAAGAACACATTGCACGGC**

223 **ATTTAGGGACTCTAAAGGGTGGAGGAGGAATGAGGGAATTGCATCATGCCAAGGCTGGTCCTCATCCATCACTG**

297 **CTTCCAGGGCCCAGAGTGGCTTCCAGGA**GTATTCTTACAAAGGAAGCCCCGATCTGTAGCTAACACTCAGAGCC

371 **CATTTTCTGCGTTAACCCCTCCCACCTCATATACAGGAGTAACATGATCAGTGACCTGGGGGAGCTGGCCAA**

445 **ACTGCGGGACCTGCCAAGCTGAGGGCCTTGGTGGTCTGCTGGACAACCCCTGTGCCGATGAGACTGACTACCGCC**

ScaI (566)

519 **AGGAGGCCCTGGTGCAGATGGCACACCTAGAGCGCCTAGACAAAGAGTACTATGAGGACGAGGACCGGGCAGAA**

593 **GCTGAGGAGATCCGACAGAGGCTGAAGGAGGAACAGGAGCAAGAAGCTGACCCGGACCAAGACATGGAACCGTA**

SdaI (700)

667 **CCTCCCGCCA**ACTTAGTGGCTCCTCTAGCCTGCAGGGACAGTAAAGGTGATGGCAGGAAGGCAGCCCCGGAGG

741 **t**CAAAGGCTGGGCACGCGGGAGGAGAGGCCAGAGTCAGAGGCTGCGGGTATCTCAGATATGAAGGAAAGATGAG

815 **AGAGGCTCAGGAAGAGGTAAGAAAAGACACAAGAGACCAGAGAAGGGAGAAGAATTAGAGAGGGAGGCAGAGGA**

889 **CCGCTGTCTCTACAGACATAGCTGGTAGAGACTGGGAGGAAGGGATGAACCTGAGCGCATGAAGGGAAGGAGG**

963 **TGGCTGGTGGTATATGGAGGATGTAGCTGGGCCAGGGAAAAGATCCTGC**ACT**aaaa**ATCTGAAGCT**aaaaat**AA

1037 **CAGGACACGGGGTGGAGAGGCGAAAGGAGGGCAGAGTG**agGCAGAGAGACTGAG**ag**GCCTGGGGATGTGGGCAT

1111 **TCCGGTAGGGCACACAGTTC**ACTTGTCTTCTTTTTCCAGGAGGCCAAAGATGCTGACGTCAAGAACTCATAA

1185 **TACCCAGTGGGGACCACCGC**ATTCATAGCCCTGTTACAAGAAGTGGGAGATGTTCTTTTTGTCCAGACTGG

1259 **AAATCC**gTTACATCCCAGGCTCAGGTTCTGTGGTGGTCATCTCTGTGTGGCTTGTCTGTGGGCCTACCTAAA

1333 **GTCCTAAGCACAGCTCTCAAGCAGATCCGAGGCGACTAAGATGCTAGTAGGGTGTCTGGAGAGAAGAGCCGA**

1407 **GGAGGTGGGCTGTGATGGATCAGTTCAGCTTTCAAATAAAAAGGCGTTTTTATATTCTGTGTGAGTTCGTGAA**

1481 **CCCCTGTGGTGGGCTTCTCCATCTGTCTGGGTTAGTACCTGCCACTATACTGGAATAAG**gGACGCCTGCTTCC

XhoI (1555)

1555 **CTCGAGTTGGCTGGACAAGGTTATGAGCATCCGTGTACTTATGGGGTTGCCAGCTTGGTCCTGGATCGCCGGG**

1629 **CCCTCCCCCACC**CGTTCCGGTCCCCACCACCACCCGCGCTCGTACGTGCGTCTCCGCTGCAGCTCTTGACTC

1703 **ATCGGGCCCCCGGGT**CACATGCGCTCGCTCGGCTCTATAGGCGCCGCCCTGCCACCCCCCGCCGCGCTG

HindIII (1816)

1777 **GGAGCCG**CAGCCGCCACTCCTGCTCTCTGCGCC**GAAGCTTCGAGGG**GGCTCGCATCTCTCTT**CACGGC**

1851 **CCGCCGCC**TACCTGAGGCCGCATCCACGCCGGTTGAGTCGCGTCTGCGCCTCCCGCTGTGGTGCCTCCT

1925 **GAACTGCGTCCGCCGT**TAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGGAGC

1999 **CTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGGTCAACTCTACGTCTTTGTTTCGTTTTTC**

NcoI (2101)

2073 TGTTCGCGCCGTTACAGATCAAGCCACCATGGTCTGGGGCCTGCATGCTGCTGCTGCTGCTGCTGGG
1 MetValLeuGlyProCysMetLeuLeuLeuLeuLeuLeuLeuLeu
2147 CCTGAGGCTACAGCTCTCCCTGGGCATCATCCAGTTGAGGAGGAGAACCCGGACTTCTGGAACCGGAGGCAG
15 yLeuArgLeuGlnLeuSerLeuGlyIleIlePProValGluGluuAsnProAspPheTrpAsnArgGluAlaA
2221 CCGAGGCCCTGGGTGCCGCAAGAAGCTGCAGCCTGCACAGACAGCCGCAAGAACCTCATCATCTTCTGGGC
40 laGluAlaLeuGlyAlaAlaLysLysLeuGlnProAlaGlnThrAlaAlaLysAsnLeuIleIlePheLeuGly
2295 GATGGGATGGGGTGTCTACGGTGACAGCTGCCAGGATCTAAAAGGGCAGAAGAAGGACAACTGGGGCCTGA
65 AspGlyMetGlyValSerThrValThrAlaAlaArgIleLeuLysGlyGlnLysLysAspLysLeuGlyProGly

NdeI (2396)

2369 GATACCCTGGCTATGGACCGCTTCCCATATGTGGCTCTGTCCAAGACATACAATGTAGACAAACATGTGCCAG
89 uIlePProLeuAlaMetAspArgPheProTyrValAlaLeuSerLysThrTyrAsnValAspLysHisValProA
2443 ACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGGTCAAGGGCAACTTCAGACCATTGGCTTGAGTGCAGCC
114 spSerGlyAlaThrAlaThrAlaTyrLeuCysGlyValLysGlyAsnPheGlnThrIleGlyLeuSerAlaAla
2517 GCCCCGTTTAAACAGTGCAACACGACACGCGGCAACGAGGTCATCTCCGTGATGAATCGGGCCAAGAAAAGCAGG
139 AlaArgPheAsnGlnCysAsnThrThrArgGlyAsnGluValIleSerValMetAsnArgAlaLysLysAlaGly
2591 GAAGTCAGTGGGAGTGGTAACCACCACACGAGTGCAGCACGCCTCGCCAGCCGGCACCTACGCCACACGGTGA
163 yLysSerValGlyValValThrThrThrArgValGlnHisAlaSerProAlaGlyThrTyrAlaHisThrValA
2665 ACCGCAACTGGTACTCGGACGCCGACGTGCCTGCCTCGGCCCGCAGGAGGGGTGCCAGGACATCGCTACGCAG
188 snArgAsnTrpTyrSerAspAlaAspValProAlaSerAlaArgGlnGluGlyCysGlnAspIleAlaThrGln
2739 CTCATCTCCAACATGGACATTGATGTGATCCTGGGTGGAGGCCGAAAGTACATGTTTCGCATGGGAACCCCAGA
213 LeuIleSerAsnMetAspIleAspValIleLeuGlyGlyGlyArgLysTyrMetPheArgMetGlyThrProAs
2813 CCCTGAGTACCAGATGACTACAGCCAAGGTGGGACCAGGCTGGACGGGAAGAATCTGGTGCAGGAATGGCTGG
237 pProGlyTyrProAspAspTyrSerGlnGlyGlyThrArgLeuAspGlyLysAsnLeuValGlnGlyTrpLeuA
2887 CGAAGCGCCAGGGTCCCCGATGTGTGGAACCGCACTGAGCTCATGCAGGCTTCCCTGGACCCGTCTGTGACC
262 laLysArgGlnGlyAlaArgTyrValTrpAsnArgThrGluLeuMetGlnAlaSerLeuAspProSerValThr
2961 CATCTCATGGGTCTCTTTGAGCCTGGAGACATGAAATACGAGATCCACCAGACTCCACACTGGACCCCTCCCT
287 HisLeuMetGlyLeuPheGluProGlyAspMetLysTyrGluIleHisArgAspSerThrLeuAspProSerLe

SacII (3083)

3035 GATGGAGATGACAGAGGCTGCCCTGCGCCTGCTGAGCAGGAACCCCCGCGGCTTCTTCTCTTCGTGGAGGGTG
311 uMetGlyMetThrGluAlaAlaLeuArgLeuLeuSerArgAsnProArgGlyPhePheLeuPheValGluGlyG
3109 GTCGCATCGACCACGGTCATCACGAAAGCAGGGCTTACCGGGCACTGACTGAGACGATCATGTTTCGACGACGCC
336 lyArgIleAspHisGlyHisHisGlyuSerArgAlaTyrArgAlaLeuThrGluThrIleMetPheAspAspAla
3183 ATTGAGAGGGCGGGCCAGCTCACCAGCAGGAGGACACGCTGAGCCTCGTCACTGCCGACCACTCCACGTCTT
361 IleGluArgAlaGlyGlnLeuThrSerGluGluAspThrLeuSerLeuValThrAlaAspHisSerHisValPh
3257 CTCCTTCGGAGGCTACCCCTGCGAGGGAGCTCCATCTTCGGGCTGGCCCCTGGCAAGGCCCGGACAGGAAGG
385 eSerPheGlyGlyTyrProLeuArgGlySerSerIlePheGlyLeuAlaProGlyLysAlaArgAspArgLysA
3331 CCTACACGGTCTCTTATACGGAACGGTCCAGGCTATGTCTCAAGGACGGCGCCCGCCGGATGTTACCAG
410 laTyrThrValLeuLeuTyrGlyAsnGlyProGlyTyrValLeuLysAspGlyAlaArgProAspValThrGlu
3405 AGCGAGAGCGGGAGCCCCAGTATCGGCAGCAGTCAAGGACGGCGCCCGCCGGATGTTACCAG
435 SerGlyuSerGlySerProGlyTyrArgGlnGlnSerAlaValProLeuAspGlyGluThrHisAlaGlyGlyAs
3479 CGTGGCGGTGTTTCGCGCGGGCCCGCAGGCGCACCTGGTTCACGGCGTGCAGGAGCAGACCTTCATAGCGCACG
459 pValAlaValPheAlaArgGlyProGlnAlaHisLeuValHisGlyValGlnGluGlnThrPheIleAlaHisSV
3553 TCATGGCTTCGCCGCTTGCCTGGAGCCCTACACCGCTGCGACCTGGCGCCCCCGCCGGCACCACCGACGCC
484 alaMetAlaPheAlaAlaCysLeuGluProTyrThrAlaCysAspLeuAlaProProAlaGlyThrThrAspAla

NheI (3667)

3627 GCGCACC CGGGCGGTCCCGGTCCAAGCGTCTGGATTGAAGCTAGCTGGCCAGACATGATAAGATACATTGATG
509 AlaHisProGlyArgSerArgSerLysArgLeuAsp●●●
3701 AGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTA

MfeI (3816)

3775 TTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTACAGGG
3849 GGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTAATTCATAAATACAGCA
3923 TAGCAA AACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCA
3997 GGGGCTGTTGCCAATGTGCATTAGCTGTTTTGCAGCCTCACCTTCTTTTCATGGAGTTTAAGATATAGTGATTTTT
4071 CCAAGGTTTTGAACTAGCTCTTCTTCTTTATGTTTTAAATGCACTGACCTCCACATTTCCCTTTTTAGTAAA
4145 ATATT CAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAAATCCAGATGCTCA
4219 AGGCCCTTCATAATATCCCCAGTTTATAGTATTTGACTTAGGGAACAAGGAACCTTTAATAGAAATTTGGACAG
4293 CAAGAAAGCGAGCTTCTAGCTTATCTCAGTCCTGCTCCTCTGCCACAAAGTGCACGCAGTTGCCGGCCGGGTG
125 ●●●AspGlnGluGluAlaValPheHisValCysAsnGlyAlaProAsp
4367 GCGCAGGGCGAACTCCCGCCCCACGGCTGCTCGCCGATCTCGGTATGGCCGGCCCGAGGCGTCCCGGAAGT
109 AraLeuAlaPheGluAraGlvTrpProGlnGluGlvIleGluThrMetAlaProGlyvSerAlaAspAraPheAs

4441 TCGTGGACACGACCTCCGACCACTCGGCGTACAGCTCGTCCAGGCCGCGCACCCACACCCAGGCCAGGGTGTG
84↓ nThr Ser Val Val Gl uSer TrpGl uAl aTyrLeuGl uAspLeuGl yArgVal TrpVal TrpAl aLeuThrAsnA
SgrAI (4574)

4515 TCCGGCACCACCTGGTCCTGGACCGCGCTGATGAACAGGGTCACGTCGTCCCGGACCACACCGGCGAAGTCGTC
59↓ spProVal Val Gl nAspGl nValAl aSer I l ePheLeuThr Val AspAspArgVal Val Gl yAl aPheAspAsp
4589 CTCCACGAAGTCCCGGGAGAACCCGAGCCGGTCCGTCCAGAACTCGACCGCTCCGGCGACGTCGCGCGCGGTGA
35↓ Gl uVal PheAspArgSer PheGl yLeuArgAspThr TrpPheGl uValAl aGl yAl aVal AspArgAl aThr Le
4663 GCACCGAACGGCACTGGTCAACTTGCCATGATGGCTCCTCCTGT CAGGAGAGGAAAGAGAAGAAGGTTAGTA
10↓ uVal ProValAl aSer Thr LeuLysAl aMet ←

MfeI (4737)

4737 CAATTGCTATAGTGAGTTGTATTATACTATGCAGATATACTATGCCAATGATTAATTGTCAA ACTAGGGCTGCA
GGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTT
4811
4885 TTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG
ACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCG
4959
5033 GATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCG
GTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTT CAGCCCAGCGCTGCGCCTTATCCGG
5107
5181 TAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA
GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACA
5255
5329 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACA
AACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAG
5403
5477 ATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCT
AGTTAATTAACATTTAAATCA
5551