



PstI (6)
SdaI (6) **SpeI (13)** Acc65I (24)

1 CCTGCAGGGCCACTAGTACCCAGGTACCTATGTTCAAAGTGCCTCAATCTAGTTAAACAAGGGCAGAGACCACGAGAAACAACACTGTGTTAGTAGC

101 AACTTAACAACCAGCCAGCAGTTCTGTCCACACACACCACCGGGCATGGTTCCAAAGCTAAAAAGGCACTAATTGCTTTTCTATAAGGAGGTAGAACA

StuI (222)

201 CAGTCCCTCCGTGTTCTTTAGGCCTGATGGCTGCATTATCGGATCTGTTACCGTGTAAATGTTCTGTCTCACACAGCCGGTTTGGGCTTCTCTCTGC

Scal (339) **BsrGI (379)**

301 ATATGTCTGGGATGGTGACGGTTCCTATATAGAGGAGTACTGGGAAGCCTCTGTGTGTGTGTGTGTCCGTGCATATGTACACATGTGTGTAAAAAG

401 CAGCCACACCGCTGAGAATGGTTAACGGGTAGCCAGGCTGTCTGACTCAGGCCCTAAGACTGGCCAGGAAAGGGCCGGGGAGGTGGGCGGGGTGAA

+1 (590)

501 GGTGGAGCGGGCTTGGCTTGTGCTCACTGCCTTTCCCAACAGGAGTACAAATGCTGGAGTGAGTGAGGTGAACCAAGTCGCCCTTAGGAATGGCTGA

601 AAAAGCCACACCTGGAATCACTCCCTCCCTGCTCTCCACGGCAGGTGCATCTGCGAGACGCTTCGGTCATTAGAGGAATGAGCCGGGAGTGAGCAA

NcoI (741) **NheI (779)**

701 TTCACCAGCTCTCCAGCACTTGGTGGAAAGCAGCAGGCAACcATGGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGACAGC

→ MetGlyYSerHisHisHisHisHisHisSGLYMetAlaSerMetThrGlyYGLN

Bsu36I (840)
Acc65I (835)

801 AAATGGGTCGGGATCTGTACGACGATGACGATAAGGTACCTAAGGATCAGCTGGAGTTGATCCCGTGGTTTACAACGTCGTGACTGGGAAAACCTGG

201 ▶ InMetGlyArgAspLeuTyrAspAspAspLysValProLysAspGlnLeuGlyValAspProValValLeuGlnArgArgAspTrpGluAsnProGly

901 CGTTACCAACTAATCGCCTTGCAGCACATCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCGCCACCGATCGCCCTTCCAACAGTTGCGCAGC

53 ▶ ValThrGlnLeuAsnArgLeuAlaAlaHisProProPheAlaSerTrpArgAsnSerGluGluAlaArgThrAspArgProSerGlnGlnLeuArgSer

Bsu36I (1077)

1001 CTGAATGGCAATGGCGCTTTCGCCTGGTTCCGGCCAGCAAGCGGTGCCGGAAGCTGGCTGGAGTGCATCTTCTGAGGCGGATACTGCTGCTGCC

87 ▶ LeuAsnGlyGluTrpArgPheAlaTrpPheProAlaProGluAlaValProGluuSerTrpLeuGluCysAspLeuProGluAlaAspThrValValValP

1101 CCTAAACTGGCAGATGCACGGTACGATGCGCCATCTACACCAACGTAACCTATCCATTACGGTCAATCCGCCGTTTGTCCACGGAGAATCCGAC

120 ▶ roSerAsnTrpGlnMetHisGlyTyrAspAlaProLeuTyrThrAsnValThrTyrProLeuThrValAsnProProPheValProThrGluAsnProTh

1201 GGGTGTACTCGCTCACATTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACCGCAATATTTTGTGATGGCGTTAACTCGGCTTTCATCTGTGG

153 ▶ rGlyCysTyrSerLeuThrPheAsnValAspGluSerTrpLeuGlnGluGlyNThrArgLleLlePheAspGlyValAsnSerAlaPheHisLeuTrp

1301 TGCAACGGCGCTGGGTCGGTTACGGCCAGGACAGTCGTTGCGCTCTGAATTTGACCTGAGCGCATTTTTACGGCCGGAGAAAACCGCTCGCGGTGA

187 ▶ CysAsnGlyArgTrpValGlyTyrGlyGluAspSerArgLeuProSerGluPheAspLeuSerAlaPheLeuArgAlaGlyGluAsnArgLeuAlaValM

1401 TGGTGTGCGTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAGCGGCATTTTCCGTGACGTCCTGTTGCTGCATAAACCGACTAC

220 ▶ etValLeuArgTrpSerAspGlySerTyrLeuGluAspGlnAspMetTrpArgMetSerGlyLlePheArgAspValSerLeuLeuHisLysProThrTh

1501 ACAAACTCAGCGATTCCATGTTGCCACTCGCTTAAATGATGATTTACGGCCGCTGACTGGAGGCTGAAGTTCAGATGTCGGCGAGTTGCGTGACTAC

253 ▶ rGlnLleSerAspPheHisValAlaThrArgPheAsnAspPheSerArgAlaValLeuGluAlaGluValGlnMetCysGlyGluLeuArgAspTyr

ClaI (1677)

1601 CTACGGGTAAACAGTTCTTTATGGCAGGTTGAAACGCAAGTCCGCCAGCGCCAGCCGCTTTCGGCGGTGAAATTATCGATGAGCGTGGTGGTTATGCCG

287 ▶ LeuArgValThrValSerLeuTrpGlnGlyGluThrGlnValAlaSerGlyThrAlaProPheGlyGlyGluLleLleAspGluArgGlyGlyTyrAlaA

1701 ATCGCGTACACTACGCTGAACGTCGAAACCCGAAACTGTGGAGCGCCGAAATCCCGAATCTCTATCGTGGTGGTGAATGCACACCCCGCAGG

320 ▶ spArgValThrLeuArgLeuAsnValGluAsnProLysLeuTrpSerAlaGluLleProAsnLeuTyrArgAlaValValGluLeuHisThrAlaAspGly

1801 CACGCTGATTGAAGCAGAAGCCTGCGATGTCGGTTCCGGCAGGTGCGGATTGAAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATTCCAGGC

353 ▶ yThrLeuLleGluAlaGluAlaCysAspValGlyPheArgGlyValArgLleGluAsnGlyLeuLeuLeuAsnGlyLysProLeuLeuLleArgGly

EcoRV (1966)

1901 GTTAACCGTCACGAGCATCTCTGATGGTCAGGTCTGGATGAGCAGACGATGGTGCAGGATATCTGCTGATGAAGCAGAACAACCTTAACGCCG

387 ▶ ValAsnArgHisGluHisHisHisProLeuHisGlyGlyValMetAspGluGluNThrMetValGluAspLleLeuLeuMetLysGluAsnAsnPheAsnAlaV

SspI (2083)

2001 TGCCTGTTCCGATTATCCGAACCATCCGCTGGTACACGCTGTGCGCACCGCTACGGCCTGTATGGTGGATGAAGCCAATATTGAAACCCACGGCAT

420 ▶ alArgCysSerHisTyrProAsnHisProLeuTrpTyrThrLeuCysAspArgTyrGlyLeuTyrValValAspGluAlaAsnLleGluThrHisGlyYMe

BsaBI (2179)

2101 GGTGCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCGATGAGCGAACCGTAACCGCAATGGTGCAGCGCGATCGTAATCACCCGAGTGTG

453 ▶ tValProMetAsnArgLeuThrAspAspProArgTrpLeuProAlaMetSerGluArgValThrArgMetValGluArgAspArgAsnHisProSerVal

2201 ATCATCTGGTCCGTTGGGAATGAATCAGGCCACGGCGCTAATCAGCAGCGCTGTATCGCTGGATCAAATCTGTCGATCCTCCCGCCGGTGCAGTATG

487 ▶ lleLleTrpSerLeuGlyAsnGluSerGlyHisGlyAlaAsnHisAspAlaLeuTyrArgTrpLleLysSerValAspProSerArgProValGlnTyrG

BssHIII (2351)

2301 AAGGCGGGAGCCGACACCAGGCCACCGATATTATTTGCCGATGTACGGCGCGTGGATGAAGACCAGCCCTTCCCGCTGTGCCGAAATGGTCCAT

520 ▶ lGlyGlyGlyAlaAspThrThrAlaThrAspLleLleCysProMetTyrAlaArgValAspGluAspGlnProPheProAlaValProLysTrpSerLle

2401 CAAAAATGGCTTTCGCTACCTGGAGAGACGCCCGCTGATCCTTTGGCAATACGCCACCGCATGGGTAACAGTCTTGGCGGTTTCGCTAAATACTGG

553 ▶ eLysLysTrpLeuSerLeuProGlyGluThrArgProLeuLleLeuCysGluTyrAlaHisAlaMetGlyAsnSerLeuGlyGlyPheAlaLysTyrTrp

2501 CAGGCGTTTCGTCAGTATCCCGTTTACAGGGCGGCTTCGCTGGGACTGGGTGGATCAGTCGCTGATTAATAATGATGAAAAACGGCAACCCGTTGCTCGG

587 ▶ GlnAlaPheArgGlnTyrProArgLeuGlnGlyGlyPheValTrpAspTrpValAspGlnSerLeuLleLysTyrAspGluAsnGlyAsnProTrpSerA

Eco47III (2688)

2601 CTTACGGCGGTGATTTTGGCGATACGCCGAACGATCGCCAGTCTGTATGAACGGTCTGGTCTTTGCCGACCGCACGCCGCATCCAGCGCTGACGGAAGC

620 ▶ I aTy rGl yGl yAspPheGl yAspThrP roAsnAspA rGlnPheCysMetAsnGl yLeuValPheAl aAspArgThrP roHi sProAl aLeuThrGl uAl

SacI (2793)

2701 AAAACACCAGCAGCAGTTTTTCCAGTTCGGTTATCCGGGCAAACCTCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGATAACGAGCTCCTGCAC

653 ▶ aLysHi sGl nGl nPhePheGl nPheA rGLeuSerGl yGl nThr l l eGl uVal ThrSerGl uTy rLeuPheA rGH i sSerAspAsnGl uLeuLeuHi s

2801 TGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGTCGCTCCACAAGGTAACAGTTGATTGAACCTGCCTGAACCTACCGC

687 ▶ TrpMetValAl aLeuAspGl yLysP roLeuAl aSerGl yGl uValP roLeuAspValAl aP roGl nGl yLysGl nLeu l l eGl uLeuP roGl uLeuP roG

2901 AGCCGGAGAGCGCCGGGCAACTCTGGCTCACAGTACCGGTAGTGAACCGAACCGGACCGCATGGTCAGAAGCCGGGCACATCAGCGCTGGCAGCAGTG

720 ▶ InP roGl uSerAl aGl yGl nLeuT rpLeuThrValA rGValValGl nP roAsnAl aTh rAl aT rpSerGl uAl aGl yHi s l l eSerAl aT rpGl nGl nT r

3001 GCGTCTGGCGGAAACCTCAGTGTGACGCTCCCGCGCGTCCCACGCCATCCCGCATCTGACCACCAGCGAAATGGATTTTGCATCGAGCTGGGTAAT

753 ▶ P ArgLeuAl aGl uAsnLeuSerValTh rLeuP roAl aAl aSerHi sAl a l l eP roHi sLeuTh rTh rSerGl uMetAspPheCys l l eGl uLeuGl yAsn

3101 AAGCGTGGCAATTTAACCGCCAGTACGGCTTCTTTCACAGATGTGGATTGGCGATAAAAAACAACCTGCTGACGCCGTGGCCGATCAGTTACCCCGTG

787 ▶ LysA rGTrpGl nPheAsnA rGl nSerGl yPheLeuSerGl nMetT rp l l eGl yAspLysLysGl nLeuLeuTh rP roLeuA rGAspGl nPheTh rA rG

3201 CACCGCTGGATAACGACATTTGGCGTAAGTGAAGCGACCCGATTGACCTAACGCCTGGTGAACGCTGGAAGGCGGCGGCCATTACCGCCGAAGC

820 ▶ I aP roLeuAspAsnAsp l l eGl yValSerGl uAl aTh rA rGl l eAspP roAsnAl aT rpValGl uA rGTrpLysAl aAl aGl yHi sTy rGl nAl aGl uAl

3301 AGCGTTGTTCAGTGCACGGCAGATACACTTGCTGATGCGGTGCTGATTACGACCCTCACCGTGGCAGCATCAGGGGAAACCTATTTATCAGCCGG

853 ▶ aAl aLeuLeuGl nCysTh rAl aAspTh rLeuAl aAspAl aVal Leu l l eTh rTh rAl aHi sAl aT rpGl nHi sGl nGl yLysTh rLeuPhe l l eSerA rG

3401 AAAACCTACCGGATTTGATGGTAGTGGTCAAAATGGCGATTACCGTTGATGTGAAGTGCCGAGCGATACCCGCATCCGGCGGGATTGGCTGAACCTGCC

887 ▶ LysTh rTy rA rGl l eAspGl ySerGl yGl nMetAl a l l eTh rValAspValGl uValAl aSerAspTh rP roHi sP roAl aA rGl l eGl yLeuAsnCysG

3501 AGCTGGCCAGGTAGCAGAGCGGTAACCTGGCTCGGATTAGGCGCCGAAAGAACTATCCGACCGCTTACTGCCGCTGTTTTCAGCCGCTGGGATCT

920 ▶ InLeuAl aGl nValAl aGl uA rGValAsnT rpLeuGl yLeuGl yP roGl nGl uAsnTy rP roAspA rGLeuTh rAl aAl aCysPheAspA rGTrpAspLe

3601 GCCATTGTACAGACATGTATACCCCGTACGCTTCCCGAGCGAAACGGTCTGCGCTGCGGGACGCGCAATTGAATTATGGCCACACCGAGTGGCGCGGC

953 ▶ uP roLeuSerAspMetTy rTh rP roTy rValPheP roSerGl uAsnGl yLeuA rG CysGl yTh rA rGl uLeuAsnTy rGl yP roHi sGl nT rpA rGGl y

3701 GACTTCCAGTTCAACATCAGCCGCTACAGTCAACAGCAACTGATGGAACAGCCATCGCCATCTGCTGCACGCGGAAGGACACATGGCTGAATATCG

987 ▶ AspPheGl nPheAsn l l eSerA rG Ty rSerGl nGl nLeuMetGl uTh rSerHi sA rGH i sLeuLeuHi sAl aGl uGl uGl yTh rT rpLeuAsn l l eA

3801 ACGGTTTCCATATGGGATTGGTGGCGACGACTCTGGAGCCGTCAGTATCGGCGGAATTACAGCTGAGCGCCGGTCCGTACCATTACAGTTGGTCTG

1020 ▶ spGl yPheHi sMetGl y l l eGl yGl yAspAspSerT rpSerP roSerValSerAl aGl uLeuGl nLeuSerAl aGl yA rG Ty rHi sTy rGl nLeuValT r

NheI (3932)

3901 GTGTCAAAAAATAAATCTAGTCGAGAAATTCGCTAGCTCGACATGATAAGATACATTGATGAGTTTGGCAAAACCACAACCTAGAATGCAGTGAAAAAAT

1053 ▶ pCysGl nLys•••

4001 GCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACATTATAAGCTGCAATAAACAAGTTAAC

MfeI (4106)

SwaI (4197)

4101 AACACAATTCGATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGGTAGATCCATTTAAA

4201 TGTTAATAACTAGCCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCGTCAGAAAAGATCAAAGGATCTTCTTGAGATCCT

4301 TTTTTCTGCGGTAATCTGCTGCTTGGCAACAAAAAACACCCGCTACCAGCGGTGGTTTGGTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGG

4401 TAACTGGCTTCAGCAGAGCCGAGATACCAATACGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTCAAGAACTCTGTAGCACCCTACATACCT

4501 CGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCAGGTTGGACTCAAGACGATAGTTACCGGATAAGGGCCAGCGG

4601 TCGGGCTGAACGGGGGTTCTGTGCACACAGCCGCTGGAGCGAACGACCTACCCGAACCTGAGATACCTACAGCGTGAGCTATGAGAAAGCCGACGC

4701 TTCCCGAAGGGGAGAAAGGCGACAGGTATCCGGTAAGCGCGAGGTCGGAACAGGAGCGCACGAGGGAGCTTCCAGGGGAAACGCTGGTATCTTTA

4801 TAGTCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGAAAAACGCCAGCAACCGCCGCTTT

AseI (4973)

4901 TTACGGTTCCTGGCCTTTTGGCTGGCCTTTGCTCACATGTTCTTAATTAATTTTTCAAAGTAGTTGACAATTAATCATCGGCATAGTATATCGGCATA

MseI (5035)

5001 GTATAATACGACTCACTATAGGAGGGCCATGAGCCAGTTGACCAGTGTGTCCAGTGTCCAGTGTCCAGCCAGGATGTGGCTGGAGCTGTTGAGTCTGG

5101 ACTGACAGGTTGGGTTCTCCAGAGATTTTGTGAGGATGACTTTGCAGGTTGGTCAGAGATGATGTCACCTGTTCTATCTCAGCAGTCCAGGACGAGG

24 ▶ ThrAspA rGLeuGl yPheSerA rGAspPheValGl uAspAspPheAl aGl yValValA rGAspAspValTh rLeuPhe l l eSerAl aValGl nAspGl nV

5201 TGGTGCCTGACAACACCTGGCTTGGTGTGGTGAGAGGACTGGATGAGCTGTGCTGAGTGGAGTGAGTGGTCTCCACCAACTCAGGGATGCCAG

57 ▶ a lValP roAspAsnTh rLeuAl aT rpValT rpValA rGGl yLeuAspGl uLeuTy rAl aGl uT rpSerGl uValValSerTh rAsnPheA rGAspAl aSe

5301 TGGCCCTGCCATGACAGAGATTGGAGAGCAGCCCTGGGGGAGAGAGTTTGCCTGAGAGACCCAGCAGGCAACTGTGTGACCTTTGTGGCAGAGAGCAG

90 ▶ rGl yP roAl aMetTh rGl u l l eGl yGl uGl nP roT rpGl yA rGGl yPheAl aLeuA rGAspP roAl aGl yAsnCysValHi sPheValAl aGl uGl uGl n

5401 GACTGAGGATAAGAATTGAGTTTCAGAAAAGGGGCGCTGAGTGGCCCTTTTTCAACTTAATTA

124 ▶ Asp•••