



XbaI (15) SdaI (29)

NotI (2) NsiI (14) EcoRI (19)

1 CCGCGCCGCTATGCATCTAGAATTCTCTGACGGCTTACATAAATTACGGTAAATGGCCCGCTGGTGACCGCCCAACGACCCCGCCATTGACGTCAAT
101 AATGACGTATGTTCCCATAGTAAACGCCAATAGGACTTTCCATTGACGTCAATGGTGGAGTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTG
201 TATCATATGCCAAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCTACTTGGC

SnaBI (311)

301 AGTACATCTACGTATTAGTCATCGCTATTACCATGATGATGCGGTTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG

SpeI (437)

401 TCTCCACCCATTGACGTCAATGGGAGTTGTTTGACTAGTCAGGGCCCCAACCCCCAAGCCCCATTTACAACACGCTGGCGCTACAGGCGCGTG

BspEI (543)

501 ACTTCCCTTGTCTTTGGGGGGGGGGCTGAGACTCTATGTGCTCCGATTGGTCAGGCACGGCTTCGGCCCCCTCTGCCACCAGATTGGCCG

601 TAGGCTCCCCGAGCGCCTGCCTCCGAGGGCCGGCGACCATAAAGAAGCCGCTAGCCACGTCCCTCGCAGTTCCGGCGTCCCGGGGTCTGTCT

HindIII (702)

701 CAAGTTGCCCGCAGAACACAGGtaagtgcctgtgtggttcccgggcctggcctctttacgggttatggccttgcgtgcttgaattacttccatg

801 cccctggtgcagtagctgattcttgcagcctcgggttggagtggtggagagttcgagccttgccttaaggagcccttcgctcgtgct

901 tgagttgagcctggccttggcgctggggcgccgctgctaattctgtggcaccttcgacctgctcgtgcttctgctaagtctctagcattttaa

BglIII (1062)

1001 atttttgataaccagctgcgagccttttttctggcgagatagcttgttaaatgcggccaagatctgcacactggtatttctggttttggggcgccggg

1101 cggcgagggggccctgctcagcgcacatgttcggcgagggggcctgcgagcggccaccgagaatcggaggggtagtctcaaatgctggccg

1201 cctgctcgtgctggcctgcgcccgctgctcggcccgcttggcggaaggtggcccggtcggcaccagttgctgagcgaaagatggccgc

1301 ttcccgccctgctcagggagctcaaatggaggagcggcgcccgggagagcggcggtgagtcaccacacaaaggaagggccttctctctc

XhoI (1461)

1401 atccgtcgtctcatgtgtaaccagggagtagccggcgccgtccaggcactcgattagttctcagccttttgagtagcgtcgtcttttaggtggggggag

1501 gggttttatcgatggagtttaccacactgagtggtggagactgaagagttaggccagcttggcacttgatgtaattctccttggaaatttgcctttt

AgeI (1688)

1601 tgagtttgatcttgcctcattctcaagcctcagacagtggttcaaggttttttcttccatttcagGTGTCGTGAAAATAACCCCTACCGGTTACAGT

BspHI (1711)

1701 AGGAGGCATCATGATTCTGGGGCCCTGCATGCTGCTGCTGCTGCTGCTGGCCCTGAGGCTACAGCTCTCCCTGGGCATCATCCAGTTGAGGAGG

1801 AGAACCCGACTTCTGGAACCGGAGGCGAGCCGAGGCGCTGGGTGCCCAAGAAGTGCAGCCTGCACAGACAGCCGCAAGAACCCTCATCTCTCTCT

30▶ luAsnP roAspPheT rpAsnArg l uA laA laG luA laLeuG l yA laA laLysLysLeuG l nP roA laG l nTh rA laA laLysAsnLeu l l e l l ePheLe

BamHI (1939)

1901 GGGCGATGGGATGGGGTGTCTACGGTACAGCTGCCAGGATCTAAAAGGGCAGAAGAAGGACAACTGGGGCTGAGATACCCCTGGCTATGGACCGC

63▶ uG l yAspG l yMe tG l yVa lSe rTh rVa lTh rA laA laA rG l l eLeuLysG l yG l nLysLysAspLysLeuG l yP roG l u l l eP roLeuA laMe tAspArg

2001 TTCCATATGTGGCTCTGTCCAAGACATACATGTAGACAAACATGTGCCAGACAGTGGAGCCACAGCCAGCCCTACCTGTGCGGGTCAAGGGCAACT

97▶ PheP roTy rVa lA laLeuSe rLysTh rTy rAsnVa lAspLysH isVa lP roAspSe rG l yA laTh rA laTh rA laTy rLeuCysG l yVa lLysG l yAsnP

2101 TCCAGACATTGGCTTGAAGTCAGCCGCGCTTTAACCAGTGCAACACGACAGCAGGAGTTCATCTCCGTGATGAATCGGGCAAGAAAGCAGG

130▶ heG l nTh r l eG l yLeuSe rA laA laA la r gPheAsnG l nCysAsnTh rTh rA r gG l yAsnG l uVa l l eSe rVa Me tAsnA r gA laLysLysA laG l

2201 GAAGTCAGTGGGAGTGTAACACACACAGTGCAGCAGCCTCGCCAGCCGACCTACGCCACAGGTAACCCCAACTGGTACTCGGACGCCGAC

163▶ yLysSe rVa lG l yVa lVa lTh rTh rTh rA r gVa lG l nH isA laSe rP roA laG l yTh rTy rA laH isTh rVa lAsnA r gAsnT rpTy rSe rAspA laAsp

2301 TCCCTGCCTGGCCCGCCAGGAGGGTGCAGGACATCTACGAGCTCAACATGGACATTGATGTATCCTGGTGGGCGCAAGAGTACA

197▶ Va lP roA laSe rA laA r gG l nG l uG l yCysG l nAsp l l eA laTh rG l nLeu l l eSe rAsnMe tAsp l l eAspVa l l l eLeuG l yG l yG l yA r gLysTy rM

2401 TGTTTCGCATGGGAACCCAGACCCTGAGTACCAGATGACTACAGCAAGTGGGACCAGGCTGGACGGGAAGAATCTGGTGCAAGGATGGTGGCGAA

230▶ e tPheA r gMe tG l yTh rP roAspP roG l uTy rP roAspAspTy rSe rG l nG l yG l yTh rA r gLeuAspG l yLysAsnLeuVa lG l nG l uT rpLeuA laLy

2501 GCGCCAGGGTGGCGGTATGTGTGAACCCGACTGAGCTCATGACGGCTTCCCTGGACCCGCTGTGACCCATCTCATGGTCTCTTTGAGCCTGGAGAC

263▶ sA r gG l nG l yA laA r gTy rVa lT rpAsnA r gTh rG l uLeuMe tG l nA laSe rLeuAspP roSe rVa lTh rH isLeuMe tG l yLeuPheG l uP roG l yAsnP

2601 ATGAAATACGAGATCCACCGAGACTCCACACTGCACCCCTCCTGATGGAGATGACAGAGGCTGCCCTGCGCTGTGAGCAGGAACCCCGCGGCTTCT

297▶ Me tLysTy rG l u l l eH isA r gAspSe rTh rLeuAspP roSe rLeuMe tG l uMe tTh rG l uA laA laLeuA r gLeuLeuSe rA r gAsnP roA r gG l yPheP

2701 TCCTCTTGCGAGGGTGGTGCATCGACCAGGTCATCAGAAAGCAGGGCTTACCGGGCACTGACTGAGAGGATCATGTTCCGACGACGCCATTGAGAG

330▶ heLeuPheVa lG l uG l yG l yA r g l l eAspH isG l yH isH isG l uSe rA r gA laTy rA r gA laLeuTh rG l uTh r l l eMe tPheAspAspA la l l eG l uA r

BbsI (2867)

2801 GCGGGCCAGCTCACCAGCGAGGAGGACGCTGAGCCTCGTCACTGCCAGCCTCCACGCTCTTCTCCTCGGAGGCTACCCCTGCGAGGGAGCTCC

363▶ gA laG l yG l nLeuTh rSe rG l uG l uAspTh rLeuSe rLeuVa lTh rA laAspH isSe rH isVa lPheSe rPheG l yG l yTy rP roLeuA r gG l ySe rSe r

2901 ATCTTCGGGCTGGCCCTGGCAAGGCCGGACAGGAAGGCTACACGGTCTCTATACGGAACGGTCCAGGCTATGTGCTCAAGGACGGCGCCCGGC

397▶ l l ePheG l yLeuA laP roG l yLysA laA r gAspA r gLysA laTy rTh rVa lLeuLeuTy rG l yAsnG l yP roG l yTy rVa lLeuLysAspG l yA laA r gP

3001 CGGATGTTACCAGAGCGAGCGAGGGAGCCCGAGTATCGGCAGCAGTCAGCAGTGGCCCTGGACGAAGAGACCCAGCAGGCGAGGACTGGCGGTGT

430▶ roAspVa lTh rG l uSe rG l uSe rG l ySe rP roG l uTy rA r gG l nG l nSe rA laVa lP roLeuAspG l uG l uTh rH isA laG l yG l uAspVa lA laVa lP h

3101 CGCGCGGGCCCGAGGCGACCTGGTTCACGGCGTGCAGGAGCAGACTTCATAGCGCACGTGATGCCCTTCGCCGCTGCCTGGAGCCCTACACCGCC

463▶ eA laA r gG l yP roG l nA laH isLeuVa lH isG l yVa lG l nG l uG l nTh rPhe l l eA laH isVa lMe tA laPheA laA laCysLeuG l uP roTy rTh rA la

NheI (3277)

3201 TGCACCTGGCGCCCCCGCCGACCCACCAGCCGCGCACCAGGGCGTCCCGTCCAAGCGTCTGGATTGAAGCTAGCTGGCCAGACATGATAAGA

497▶ CysAspLeuA laP roP roA laG l yTh rTh rAspA laA laH isP roG l yA r gSe rA r gSe rLysA r gLeuAsp●●●

3301 TACATTGATGAGTTTGGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAA
3401 GCTGCAATAAAACAAGTTAACACAACAATTGCATTCATTTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAA
3501 ATGTGGTATGG AATTAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAG
3601 GCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAGATATAGTGATTTTTCCCAAGGTTTGAAGTAGCTC
SspI (3756)
3701 TTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGT
3801 TTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAAACAAAGGAACCTTTAATAGAAATGGAC
3901 AGCAAGAAAGCGAGCTTCTAGCTTATCCTCAGTCCTGCTCCTGCCACAAAGTGCACGAGTTGCCGGCCGGTCCGGCAGGGCGAACTCCCGCCCCA
125 ◀ ●●●AspG InG luG luA laVa IPheH isVa ICysAsnG lyA laP roAspArgLeuA laPheG luArgG lyT rp
4001 CGGCTGCTCGCGATCTCGGTCTATGGCCGGCCCGAGGCGTCCCGAAGTTCGTGGACACGACCTCCGACCACTCGGCGTACAGCTCGTCCAGGCGCGC
101 ◀ P roG InG luG ly I leG luTh rMe tA laP roG lySe rA laAspArgPheAsnTh rSe rVa Ma IG luSe rT rpG luA laTy rLeuG luAspLeuG lyArgV
SgrAI (4184)
4101 ACCCACACCCAGGCCAGGGTGTGTCGGCACACCTGGTCCTGGACCGCGCTGATGAACAGGGTCAGTCGTCGGACACACCGGCGAAGTCGTCCT
67 ◀ a IT rpVa IT rpA laLeuTh rAsnAspP roVa Ma IG InAspG InVa IA laSe r I lePheLeuTh rVa AspAspArgVa Va IG lyA laPheAspAspG I
4201 CCACGAAGTCCCGGAGAACCCGAGCCGGTCCGAGAACTGCACCGCTCCGGCGACGTCGCGCGGTTGAGCACCGGAACGGCACTGGTCAACTTGGC
34 ◀ uVa IPheAspArgSe rPheG lyLeuArgAspTh rT rpPheG luVa IA laG lyA laVa lAspArgA laTh rLeuVa IP roVa IA laSe rTh rLeuLysA la
4301 CATGATGGCTCCTCCTGT CAGGAGAGGAAAGAGAAGGTTAGTACAATTGCTATAGTGAGTTGTATTATACTATGCAGATATACTATGCCAATGATTA
1 ◀ Me t
4401 ATTGTCAA ACTAGGGTCTCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTCCAT
4501 AGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAA
4601 GCTCCCTCGTGCGCTCTCCTGTTCCGACCTGCGGCTTACC GGATACTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTG
4701 TAGGTATCTCAGTTCGGTGTAGGTGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTT CAGCCGACCGCTGCGCCTTATCCGGTAACTATCTG
4801 CTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGGGTGCTACAGATTCT
4901 TGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTG
5001 ATCCGGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTT
5101 TCTACGGGTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCA