



EcoRI (23)

XbaI (19)

NotI (2)

EcoRV (17)

SdaI (38)

1 GCGGCCGCGTCGACGATATCTAGAATTCGGATCCTGCAGGATCCGCTGGATGCAACTCAGCTGGGGTCAGCTCAG

76 TCGACTTGGGTAACTGAGTGCCGGCCTTGTCTGTCTTTGAATATCAGATAAATGAGTTGACTTAAAATTTGTT

151 CATTGTACTTTTCATGTCACCTGTGCTTTTCCCTGCCTTCTACCCACAGCCCTGCCAGCTGGCAGGAGGAAGGT

226 CAGCAGAGCTGCTGATAAGAGCCGTATAAAGAGGGTTCCGCTCATGGCAAGGGGCAGTGGTCTACTCTCTCCACA

NcoI (301)

301 CCATGGTTCTGGGGCCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGGGCTGAGGCTACAGCTCTCCCTGGGCATCA

1 M V L G P C M L L L L L L L G L R L Q L S L G I

376 TCCCAGTTGAGGAGGAGAACCCGACTTCTGGAACCGCGAGGCAGCCGAGGCCCTGGGTGCCGCAAGAAGCTGC

25 I P V E E E N P D F W N R E A A E A L G A A K K L

451 AGCCTGCACAGACAGCCGCAAGAACCTCATCATCTTCTGGGCGATGGGATGGGGGTGTCTACGGTGACAGCTG

50 Q P A Q T A A K N L I I F L G D G M G V S T V T A

NdeI (596)

526 CCAGGATCCTAAAAGGGCAGAAGAAGGACAAACTGGGGCCTGAGATACCCCTGGCTATGGACCGCTTCCCATATG

75 A R I L K G Q K K D K L G P E I P L A M D R F P Y

601 TGGCTCTGTCCAAGACATAACAATGTAGACAAACATGTGCCAGACAGTGGAGCCACAGCCACGGCCTACCTGTGCG

100 V A L S K T Y N V D K H V P D S G A T A T A Y L C

676 GGGTCAAGGGCAACTTCCAGACCATTGGCTTGGTGCAGCCGCCGCTTTAACAGTGAACACGACACGCGGCA

125 G V K G N F Q T I G L S A A A R F N Q C N T T R G

751 ACGAGGTCATCTCCGTGATGAATCGGGCCAAGAAAGCAGGGAAGTCACTGGGAGTGGTAACCACCACACGAGTGC

150 N E V I S V M N R A K K A G K S V G V V T T T R V

826 AGCACGCCTCGCCAGCCGGCACCTACGCCACACGGTGAACCGCAACTGGTACTCGGACGCCGACGTGCCTGCCT

175 Q H A S P A G T Y A H T V N R N W Y S D A D V P A

901 CGGCCCGCCAGGAGGGGTGCCAGGACATCGCTACGCAGCTCATCTCCAACATGGACATTGATGTGATCCTGGGTG

200 S A R Q E G C Q D I A T Q L I S N M D I D V I L G

976 GAGGCCGAAAGTACATGTTTTGCATGGGAACCCAGACCCTGAGTACCCAGATGACTACAGCCAAGGTGGGACCA

225 G G R K Y M F R M G T P D P E Y P D D Y S Q G G T

1051 GGCTGGACGGGAAGAATCTGGTGCAGGAATGGCTGGCGAAGCGCCAGGGTGGCCGGTATGTGTGGAACCGCACTG

250 R L D G K N L V Q E W L A K R Q G A R Y V W N R T

1126 AGCTCATGCAGGCTTCCCTGGACCCGTCTGTGACCCATCTCATGGGTCTCTTTGAGCCTGGAGACATGAAATACG

275 E L M Q A S L D P S V T H L M G L F E P G D M K Y

1201 AGATCCACCGAGACTCCACACTGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGCGCCTGCTGAGCAGGA

300 E I H R D S T L D P S L M E M T E A A L R L L S R

SacII (1283)

1276 ACCCCCGCGGCTTCTTCTTCTCGTGAGGGTGGTGCATCGACCACGGTCATCACGAAAGCAGGGCTTACCGGG

325 N P R G F F L F V E G G R I D H G H H E S R A Y R

1351 CACTGACTGAGACGATCATGTTTCGACGACGCCATTGAGAGGGCGGGCCAGCTCACCAGCGAGGAGGACACGCTGA

350 A L T E T I M F D D A I E R A G Q L T S E E D T L

1426 GCCTCGTCACTGCCGACACTCCCACGTCTTCTCCTTCCGAGGGTACCCCTGCGAGGGAGCTCCATCTTCCGGG

375 S L V T A D H S H V F S F G G Y P L R G S S I F G

1501 TGGCCCCTGGCAAGGCCCGGGACAGGAAGGCCTACACGGTCTCCTATACGAAACGGTCCAGGCTATGTGCTCA

400 L A P G K A R D R K A Y T V L L Y G N G P G Y V L

1576 AGGACGGCGCCCGCCGGATGTTACCGAGAGCGAGAGCGGGAGCCCCGAGTATCGGCAGCAGTCAGCAGTGGCCC

425 K D G A R P D V T E S E S G S P E Y R Q Q S A V P

1651 TGGACGAAGAGACCCACGCAGGCGAGGACGTGGCGGTGTTTCGCGCGCGCCCGCAGGCGCACCTGGTTCACGGCG

450 L D E E T H A G E D V A V F A R G P Q A H L V H G

1726 TGCAGGAGCAGACCTTCATAGCGCACGTATGGCCTTCCGCCCTGCTGGAGCCCTACACCGCCTGCGACCTGG

475 V Q E Q T F I A H V M A F A A C L E P Y T A C D L

NheI (1867)

1801 CGCCCCCGCCGGCACCACCGACGCCGCGCACCCGGGGCGGTCCCGGTCCAAGCGTCTGGATTGAAAGCTAGCTGG

500 A P P A G T T D A A H P G R S R S K R L D •

1876 CCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAGTGAATGCAGTGAAAAAATGCTTTATTTGT

1951 GAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATT

2026 CATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATG

2101 GAATTAATTCTAAAATACAGCATAGCAAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGG
▶

2176 ATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGAGCCTCACCTTCTTTCATGGAG
2251 TTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCC

SspI (2346)

2326 ACATTCCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAG
2401 GCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTGTAGTGTGGACTTAGGGAACAAAGGAACCTT
2476 TAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTATCCTCAGTCCTGCTCCTCTGCCACAAAGTGCACGC
125↓ • D Q E E A V F H V C

2551 AGTTGCCGGCCGGGTGCGCAGGGCGAACTCCCGCCCCACGGCTGCTCGCCGATCTCGGTCATGGCCGGCCCGG
114↓ N G A P D R L A F E R G W P Q E G I E T M A P G S

2626 AGGCGTCCCGGAAGTTCGTGGACACGACCTCCGACCACTCGGCGTACAGCTCGTCCAGGCCGCGCACCCACACC
89↓ A D R F N T S V V E S W E A Y L E D L G R V W V W

SgrAI

2701 AGGCCAGGGTGTGTCCGGCACCACTGGTCTGGACCGCGCTGATGAACAGGGTCACGTCGTCCCGGACCACAC
64↓ A L T N D P V V Q D Q V A S I F L T V D D R V V G

2776 CGGCGAAGTCGTCTCCACGAAGTCCCGGGAGAACCCGAGCCGGTCCGTCAGAACTCGACCGCTCCGGCGACGT
39↓ A F D D E V F D R S F G L R D T W F E V A G A V D

2851 CGCGCGGGTGAGCACCGGAACGGCACTGGTCAACTTGGCCATGATGGCTCCTCCTGTGAGGAGAGAAAGAGAA
14↓ R A T L V P V A S T L K A M ←

2926 GAAGTTAGTACAATTGCTATAGTGAGTTGTATTATACTATGCAGATATACTATGCCAATGATTAATTGTCAAAC

3001 TAGGGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTG
▶

3076 CTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC

3151 CCGACAGGACTATAAAGATACCAGGCGTTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCG

3226 CTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTTCTCATAGCTCACGCTGTAGGTATCTC

3301 AGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCAGCCGCTGCGCCTTA

3376 TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG

3451 ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGA

3526 ACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA

3601 CAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAA

3676 GATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCT

3751 AGTTAATTAACATTTAAATCA
