



EcoRI (23)

NotI (2) XbaI (19) SdaI (38) SpeI (45)

1 CGGGCCGCGTCGACGATATCTAGAATTCGGATCCTGCAGGCCACTAGTCGCCTTGCTGTGCCACTTTGGACTTCCCTCCCTAGCCTGAGCTTCAGTT

Eco47III

101 TTCCTGCCTGTTAGGCAGCCCCATGTCAACTGCACCTTAGTAGGCCGGTTTATGTCGCCGACAAGACGTGAAGTGGTGGAGGTGGGCAGGATCCCAGCGCT
201 ACCATCTTCTTGAACCAGTGATCTCAACACATCGGATTTCTGTTTCTCATCTGCAAAATGGGATCAGTGAGCTCAGGTGGGTCAAAAATCTACAGGAA
301 CTACTTTAGCCAAGCCCGCCCCCTGAAAGTTCCTCGGTGGGCTGTTAGGGTATTGTTTTCATCTGTGGGGCTCCCTGATGCGTCCCACCCACCAGC
401 CTTGGAGAGGGTGGGATGGGAGGGTGGGTGCTTGGGGAGACAAGCCTAGAGCCTGGGCCCTCCACCCCACTGCCTCCCCCATCCAGGGCCCCCAC
501 CCAGTGACAAAGCCCGTGGCACTTCTCTACCCGTTGGCAGGCGCCTGGCCAGCCCCTTCTCTAAGGAAGCGCATTTCTGCCTCCCTGGCGCGCC
601 GGGCTGGATGAGCCGGGAGCTCCCTGCTGCCGGTCAACACAGCCTTTCATCTGCCCTGGGGCCAGGACTGCTGCTGACTGCCATCCATTGGAGCC
701 CAGCACCCCTCCCCGCCATCCTTCGGACAGCAACTCCAGCCAGCCCCGGTCCCTGTGTCCACTTCTCTGACCCTCGGCCGCCACCCAGAAGGC

Tth111I (813)

801 TGGAGCAGGGACGCCGTCGCTCCGGCCGCTGCTCCCTCGGGTCCCGTGGAGCCACGCCGGCCCCGGTGCCGCCCGCAGCCCTGCCACTGGACAC

NcoI (937)

901 AGGATAAGGCCAGCGCACAGGCCCCACGTGGACACCATGGTTCTGGGGCCCTGCATGCTGCTGCTGCTGCTGCTGGCCCTGAGGCTACAGCTCTC
1001 CCTGGGCATCATCCCAGTTGAGGAGGAGAACCCGGACTTCTGGAACCGCAGGACCGGAGCCCTGGGTGCCGCAAGAAGCTGCAGCCTGCACAGACA
210 L G I 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 Q P A Q T
1101 GCCCCAAGAACCTCATCATCTTCTGGCGATGGGATGGGGTGTCTACGGTGACAGCTGCCAGGATCCTAAAAGGCGAGAAGAAGGACAAACTGGGGC
550 A A K N L I I F L G D G M G V S T V T A A R I L K G Q K K D K L G

NdeI (1232)

1201 CTGAGATACCCTGGCTATGGACCGCTTCCCATATGTGGCTCTGTCCAAGACATAACAATGTAGACAAACATGTGCCAGACAGTGGAGCCACAGCCACGGC
880 P E I P L A M D R F P Y V A L S K T Y N V D K H V P D S G A T A T A
1301 CTACCTGTGCGGGTCAAGGGCACTTCCAGACATTGGCTTGTAGTGCAGCCGCGCTTAAACAGTGAACACGACAGCGGCAACGAGGCTCATCTCC
1210 Y L C D 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 N E V I S
1401 GTGATGAATCGGGCAAGAAAGCAGGGAAGTCAGTGGGAGTGGTAACCACACAGAGTGCAGCAGCCCTGCCAGCCGGCACCTACGCCACACGGTGA
1550 V M N R A K K A G K S V G V V T T T R V Q H A S P A G T Y A H T V
1501 ACCGCAACTGGTACTCGGACCGCAGCTGCCTGCCTCGGCCCGCAGGAGGGGTGCCAGGACATCGTACGAGCTCATCTCCAACATGGACATGTATGT
1880 N R N W Y S D A D V P A S A R Q E G C Q D I A T Q L I A T S N M D I D V
1601 GATCTGGGTGGAGGCCAAAGTACATGTTTCGCATGGGAACCCAGACCTGAGTACCCAGATGACTACAGCAAGGTGGACCGAGCTGGACGGGAAG
2210 I L G 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 R L D G K
1701 AATCTGGTGCAGGAATGGCTGGCGAAGCGCCAGGGTGGCCGGTATGTGTGGAACCCACTGAGCTCATGCAGGCTTCCCTGGACCCGCTGTGACCCATC
2550 N L V Q E W L A K R Q G A R Y V W N R T E L M Q A S L D P S V T H
1801 TCATGGGTCTCTTTGAGCCTGGAGACATGAAATACGAGATCCACCGAGACTCCACACTGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGGCCT
2880 L M G L F E P G D M K Y E I H R D S T L D P S L M E M T E A A L R L

SacII (1919)

1901 GCTGAGCAGGAACCCCGCGGCTTCTTCTCTTCTGAGGGTGGTGCATCGACCACGGTTCATCACGAAAGCAGGGCTTACCGGGCACTGAGTACGAGC
3210 L S R 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 A L T E T
2001 ATCATGTTGACGACGCCATTGAGAGGGCGGCGAGTACCAGCAGGAGGACAGCTGAGCCTCGTCACTGCCAGCCATCCACGCTTCTTCTCTTCC
3550 I M F D D A I E R A G Q L T S E E D T L S L V T A D H S V F S F
2101 GAGGCTACCCCTGCGAGGGAGCTCCATCTTGGGCTGGCCCTGGCAAGCCCGGGACAGGAAGCCTACAGGCTCCTCTATACGGAAACGGTCCAGG
3880 G G Y P L R G S S I F G L A P G K A R D R K A Y T V L L Y G N G P G
2201 CTATGTGCTCAAGGACGGCCCGCCGGATGTTACCGAGAGCGAGAGCGGGAGCCCGAGTATCGGCAGCAGTCAAGTCCGCTGGACGAAAGACG
4210 Y V L 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 L D E E T
2301 CACGACGGCAGGACGTTGGCGGTGTTGCGCGCGGCGCAGGCGCACCTGGTTCAGGCGTGCAGGAGCAGACTTATAGCGCACGTCATTGGCCTTCG
4550 H A G E D V A V F A R G P Q A H L V H G V Q E Q T F I A H V M A F
2401 CCGCTGCCTGGAGCCCTACACCGCTGCGACCTGGCGCCCGCGCGCACCCAGCGCGCACCCGGGGCGGTCCCGGTCCAAGCGTCTGGATTG
4880 A A C L E P Y T A C D L A P P A G T T D A A H P G R S R S K R L D

NheI (2503)

2501 AAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGC
5210

MfeI (2652)

2601 TATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAAACAACATTCATTATTTATGTTTCAGGTTTCAGGGGGAGGTGGGGAGTT
2701 TTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCC

2801 TTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTTTCATGGAGTTTAAAGATATAGTG
2901 TATTTTCCAAGGTTTGAAGTCTTCTTCTTATGTTTAAATGCAGTACCTCCACATTCCTTTTATGAAAAATTCAGAAATAATTTAAA
3001 TACATCATTGCAATGAAAAATAAATGTTTTTATTAGGCAGAAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTATGATGTTGACTTAGGGAAC
3101 AAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTATCCTCAGTCTGCTCTGCGCAAAAGTGCACGCAGTTGCCGGCCGGG

1250 • D Q E E A V F H V C N G A P

3201 TCGCGCAGGGCGAAGTCCCGCCCCACGGCTGCTCGCCGATCTCGGTCATGGCCGGCCGGAGGCGTCCCGGAAGTTCGTGGACACGACCTCCGACCACT
1090 D R L A F E R G W P Q E G I E T M A P G S A D R F N T S V V E S W E
3301 CGGCGTACAGCTCGTCCAGCGCGCACCCACCCAGGCGAGGGTGTGTCGGCACCTGCTGCTGGACCGCTGATGAACAGGGTGCAGCTGCT
760 A Y L E D L G R V W V W A L T N D P V V Q D Q V A S I F L T V D D

SgrAI (3410)

3401 CCGGACACACCGCGGAAGTCTGCTCCACGAAGTCCCGGGAGAACCAGCGGGTCCGTCGAGTCCAGAACTCGACCGCTCCGGCGACGTCGCGCGGGTGGAGC
430 R V V G 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 R A T L

3501 ACCGGAACGGCACTGGTCAACTGGCCATGATGGCTCCTCCTGTCAGGAGAGGAAAGAGAAGAAGGTTAGTACAATTGCTATAGTGAGTTGTATTACT
9 V P V A S T L K A M
3601 ATGCAGATACTATGCCAATGATTAATTGTCAAAGTAGGGCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAA
3701 AGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA
3801 AAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGC
3901 GTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACC
4001 GCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG
4101 GTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC
4201 TTCGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTTGAAGCAGCAGATTACGCGCAGAAAAAAG
4301 GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATT
4401 TAAATCA