



PvuI (7)
SgfI (6)
MfeI (82)

1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA

101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGTGCTGTACTGGCTCCGCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)

Psp1406I (203)
PvuII (239)
Bsu36I (291)

201 GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACAGCGCCCGCCGCTACCTGAGGCC

301 GCCATCCACGCCGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)

401 GGGCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCAGCTTTGCTGACCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

XcmI (573)

AgeI (552)
SphI (568)

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTAGGAGGCCAGCATGCCACATACTTTGTGGATGGTGTGGGTC

601 TTGGGGTTCATCATCAGCCTCTCCAAGGAAGTCTCCAATCAGGCTTCTGTCTTGTGACCGCAATGGTATCTGCAAGGGCAGCTCAGGATCTTTAA

111▶ L G V I I S L S K E E S S N Q A S L S C D R N G I C K G S S G S L

Bsu36I (709)

701 ACTCCATTCCCTCAGGCTCACAGAAGCTGAAAAAGCCTTGACCTGTCCAACAACAGGATCACCTACATTAGCAACAGTGACCTACAGAGGTGTGTGAA

44▶ N S I P S G L T E A V K S L D L S N N R I T Y I S N S D L Q R C V N

801 CCTCCAGGCTCTGGTGTGACATCCAATGGAATTAACACAATAGAGGAAGATTCTTTTTCTCCCTGGGAGCTTTGAACATTTAGACTTATCCTATAAT

77▶ L Q A L V L T S N G I N T I E E D S F S S L G S L E H L D L S Y N

AvrII (983)

901 TACTTATCTAATTTATCGTCTTCTGTTCAAGCCCTTTCTCTTTAACTTCTTAACTTACTGGGAAATCCTTACAAAACCTAGGGGAAACATCTC

111▶ Y L S N L S S S W F K P L S S L T F L N L L G N P Y K T L G E T S

1001 TTTTTCTCATCTCACAAAATGCAAATCCTGAGAGTGGGAAATATGGACACCTTCACTAAGATTCAAAGAAAAGATTTTGTGGACTTACCTTCCTTGA

144▶ L F S H L T K L Q I L R V G N M D T F T K I Q R K D F A G L T F L E

BglII (1121)
NdeI (1185)

1101 GGAACCTGAGATTGATGCTTCAGATCTACAGAGCTATGAGCCAAAAGTTTGAAGTCAATTGAGAATGTAAGTCACTGATCCTTCATATGAAGCAGCAT

177▶ E L E I D A S D L Q S Y E P K S L K S I Q N V S H L I L H M K Q H

1201 ATTTTACTGCTGGAGATTTTTGTAGATGTTACAAGTCCGTGGAATGTTTGGAACTGCGAGATACTGATTTGGACACTTTCCATTTTTCAGAACTATCCA

211▶ I L L L E I F V D V T S S V E C L E L R D T D L D T F H F S E L S

1301 CTGGTGAACAAATTCATTGATTAATAAGTTTACATTTAGAAATGTGAAAATCACCGATGAAAGTTTGTTCAGGTTATGAACTTTTGAATCAGATTTT

244▶ T G E T N S L I K K F T F R N V K I T D E S L F Q V M K L L N Q I S

Psp1406I

1401 TGGATTGTTAGAATTAGAGTTTGTGACTGTACCTTAATGGAGTTGGTAATTTTAGAGCATCTGATAATGACAGAGTTATAGATCCAGGTAAGTGGAA

277▶ G L L E L E F D D C T L N G V G N F R A S D N D R V I D P G K V E

BspEI (1510)

HpaI (1502)

1501 ACGTTAACAAATCCGGAGGCTGCATATCCAAGTTTTACTTATTTTATGATCTGAGCACTTTATATTCACTTACAGAAAGAGTAAAAAGAATCACAGTAG

311▶ T L T I R R L H I P R F Y L F Y D L S T L Y S L T E R V K R I T V

1601 AAAACAGTAAAGTTTTCTGGTTCCTGTTTACTTTACAACATTTAAATCATTAGAATACTGGATCTCAGTGAATTTGATGGTTGAAGAATACTT

344▶ E N S K V F L V P C L L S Q H L K S L E Y L D L S E N L M V E E Y L

1701 GAAAAATTCAGCCTGTGAGGATGCCTGGCCCTCTACAACTTAAATTTAAGGCAAAATCATTGGCATATTGAAAAAACCGGAGAGACTTTGCTC

377▶ K N S A C E D A W P S L Q T L I L R Q N H L A S L E K T G E T L L

EcoRV (1824)
SspI (1885)

1801 ACTCTGAAAAACTTGACTAACATTGATATCAGTAAGAATAGTTTTTCTTCTATGCCTGAAACTGTGAGTGGCCAGAAAAGATGAAATATTTGAACTTAT

411▶ T L K N L T N I D I S K N S F H S M P E T C Q W P E K M K Y L N L

1901 CCAGCACAGAAATACACAGTGAACAGGCTGCATTCCTCAAGACTGGAAATTTAGATGTTAGCAACAACATCTCAATTTATTTTCTTTGAATTTGCC

444▶ S S T R I H S V T G C I P K T L E I L D V S N N N L N L F S L N L P

SpeI (2072)

2001 GCAACTCAAAGAATTTATATTTCCAGAAATAAGTTGATGACTCTACCAGATGCCTCCCTCTTACCATGTTACTAGTATTGAAAATCAGTAGGAATGCA

477▶ Q L K E L Y I S R N K L M T L P D A S L L P M L L V L K I S R N A

EcoRI (2187)

2101 ATAACACGTTTTCTAAGGAGCAACTTGACTCATTTACACACTGAAGACTTTGGAAGCTGGTGGCAATAACTTCATTTGCTCCTGTGAATTCCTCTCCT

511▶ I T T F S K E Q L D S F H T L K T L E A G G N N F I C S C E F L S

2201 TCACTCAGGAGCAGCAAGCACTGGCCAAAGTCTTGATTGATTGGCCAGCAAAATACCTGTGTGACTCTCCATCCCATGTGCGTGGCCAGCAGGTTGAGGA

544▶ F T Q E Q Q A L A K V L I D W P A N Y L C D S P S H V R G Q Q V Q D

XcmI (2396)

2301 TGTCGCTCTCGGTGTGCGAATGTCACAGGACAGCACTGGTGTCTGGCATGTGCTGTGCTGTTCTGCTGATCCTGCTCACGGGGTCTGTGCCAC

577▶ V R L S V S E C H R T A L V S G M C C A L F L L I L L T G V L C H

NcoI (2405)
BamHI (2493)

2401 CGTTTCCATGGCCTGTGGTATATGAAAATGATGTGGGCTGGCTCCAGGCCAAAAGGAAAGCCAGGAAAGCTCCAGCAGGAACATCTGCTATGGATCCT

611▶ R F H G L W Y M K M M W A W L Q A K R K P R K A P S R N I C Y G S

2501 ACCCCTATGATGTGCCAGACTACGCC TAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC TAGAATGCAGTGAAAAA
644 ▶ Y P Y D V P D Y A •

2601 AATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAAC AATTGCATTTCATTTATGTT

2701 TCAGGTT CAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTAACTC

2801 CAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTC

2901 TTTTCATGGAGTTTAAAGATATAGTGATTTTTCCCAAGTTTGAAGTCTCTTCATTCTTTATGTTTTAAATGCAGCTGACCTCCACATTCCCTTTTTAG

3001 TAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCCA

3101 GTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGAT
141 ◀ • N R T Y K L P I

3201 GAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAG
132 ◀ L E E I T T K V L K G N M E I L V F C D P A Y D S I L E R C M G C

3301 GGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCAAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACC
98 ◀ P S V V R I S R D V E D S Y P H R V A V I T D F D K Q G N S V A S G

3401 CAATGGCAATGGCTT CAGCACAGACAGTGACCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCCC
65 ◀ I A I A E A C V T V R G I Y A E I H V A S I I E G T K T R I A A G

3501 GACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGTGT
32 ◀ V H H K N D E Y L M T I K E T A V E V L E L D Q Q S I N F T K M ◀

3601 GCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAA AACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTC

3701 ACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCCTT

3801 GATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCA

3901 AAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGC

4001 CATTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGAC

4101 GTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGCGGGGGTCTGTTGGCGGTGAGCCAGGCGGGCCATTAC

4201 CGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC

4301 ATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGG

4401 AAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGC

4501 TGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAACTATC

4601 GTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT

4701 CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCT

4801 TGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCT

4901 TTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACCTACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAAATA

5001 TCTTTATTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAAT

5101 AGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA