



PvuI (7)
SgfI (6)
EcoNI (96)

1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
 101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Bsu36I (291)

Psp1406I (203)
PvuII (239)
EcoNI (287)

201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACAGCGCCCGCCCTACCTGAGGCC
 301 GCCATCCACGCGGTTGAGTGCAGTTCGCCCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)

401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

KasI (535)
AgeI (552)
BspHI (560)
XcmI (575)

501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGGCTACCTGAGATCACCGGTCATCATGAGTGAGGTCACCAAGAATTCCTGGAGAAAATCCT
1▶ M S E V T K N S L E K I L

PvuII (604)
Tth111I (643)
XbaI (660)

601 TCCACAGCTGAAATGCCATTTCACTTGGAACTTATTCAAGGAAGACAGTGTCTCAAGGATCTAGAAGATAGAGTGTGTAACAGATTGAATTTTTAAAC
 13▶ P Q L K C H F T W N L F K E D S V S R D L E D R V C N Q I E F L N

BsrGI (719)

701 ACTGAGTTCAAAGCTACAATGTACAACCTGTTGGCCTACATAAAAACCTAGATGGTAACAACGAGGCAGCCCTGGAATGCTTACGGCAAGCTGAAGAGT
 47▶ T E F K A T M Y N L L A Y I K H L D G N N E A A L E C L R Q A E E
 801 TAATCCAGCAAGAATGCTGACCAAGCAGAAATCAGAAGTCTAGTCACTGGGAACTACGCTGGTCTACTATCACTTGGCAGACTCTCAGATGC
 80▶ L I Q Q E H A D Q A E I R S L V T W G N Y A W V Y Y H L G R L S D A
 901 TCAGATTTATGTAGATAAGGTGAAACAAACCTGCAAGAAATTTCAAATCCATACAGTATTGAGTATTCTGAAGTACTGTGAGGAAGGGTGGACACAA
 113▶ Q I Y V D K V K Q T C K K F S N P Y S I E Y S E L D C E E G W T Q
 1001 CTGAAGTGTGGAAGAAATGAAAGGGCAAGGTGTGTTTTGAGAAGGCTCTGGAAGAAAAGCCCAACAACCCAGAATTCCTCTGGACTGGCAATTGCGA
 147▶ L K C G R N E R A K V C F E K A L E E K P N N P E F S S G L A I A
 1101 TGTACCATCTGGATAATCACCCAGAGAAACAGTCTCTACTGATGTTTTGAGCAGGCCATTGAGCTGAGTCTGATAACCAATACGTCAAGTTCTCTT
 180▶ M Y H L D N H P E K Q F S T D V L K Q A I E L S P D N Q Y V K V L L

PstI (1210)

1201 GGGCTGAAATGCAGAAGATGAATAAAGAAGCTGAAGGAGAGCAGTTTGTGGAAGAAGCCTTGAAAAGTCTCCTTGCCAAACAGATGCTCCTCCGAGT
 213▶ G L K L Q K M N K E A E G E Q F V E E A L E K S P C Q T D V L R S

BstEII (1324)

1301 GCAGCCAAATTTTACAGAAGAAAAGGTGACCTAGACAAAGCTATTGAACTGTTTCAACGGGTGTTGGAATCCACACCAAAACAATGGCTACCTCTATCACC
 247▶ A A K F Y R R K G D L D K A I E L F Q R V L E S T P N N G Y L Y H
 1401 AGATTGGGTGCTGTACAAAGCAAAAGTAAGACAAATGCAAGAATACAGGAGAATCTGAAGTGTGGAATAAAGAGATGATTGAAGCACTAAAGCAATA
 280▶ Q I G C C Y K A K V R Q M Q N T G E S E A S G N K E M I E A L K Q Y

BstBI (1512)
NsiI (1550)

1501 TGCTATGGACTATTGCAATAAAGCTCTTGAAGAGGACTGAATCCTCTGAATGCATACTCCGATCTCGCTGAGTTCCTGGAGACGGAATGTTATCAGACA
 313▶ A M D Y S N K A L E K G L N P L N A Y S D L A E F L E T E C Y Q T

Eco47III (1646)

1601 CCATTCATAAAGGAAGTCCCTGATGCTGAAAAGCAACAATCCCATCAGCGCTACTGCAACCTTCAGAAATATAATGGGAAGTCTGAAGACTGCTGTGC
 347▶ P F N K E V P D A E K Q Q S H Q R Y C N L Q K Y N G K S E D T A V
 1701 AACATGGTTTAGAGGGTTTGTCCATAAGCAAAAATCAACTGACAAGGAAGAGATCAAAGACCAACCAGAAATGTATCTGAAAATCTGCTTCCACAAAA
 380▶ Q H G L E G L S I S K K S T D K E E I K D Q P Q N V S E N L L P Q N

AseI (1827)
BglIII (1848)

1801 TGCACCGAATTATTGGTATCTTCAAGGATTAATTCATAAGCAGAATGGAGATCTGCTGCAAGCAGCCAAATGTTATGAGAAGGAAGTGGCCCGCTGCTA
 413▶ A P N Y W Y L Q G L I H K Q N G D L L Q A A K C Y E K E L G R L L
 1901 AGGGATGCCCTTCAGGCATAGGCAGTATTTTCTGTGAGTCTGAGCTTGGAGATGGTGTGAGGAAATGGCCAGGGCGCAGTCAGTCCAGTCCCA
 447▶ R D A P S G I G S I F L S A S E L E D G S E E M G Q G A V S S S P

MscI (2060)
NheI (2054)

2001 GAGAGCTCCTCTAAGTCAAGCAACTGAACTGAGACAGAGGAGGAAAAACAGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAA
 480▶ R E L L S N S E Q L N •

HpaI (2192)

2101 CCACAAC TAGAATGCAGTGA AAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACATTATAAGTGCATAAACAAGTTAACAA
 2201 CAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAT
 2301 ACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCA
 2401 TTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGTATTTTCCCAAGTTTGAAGTACTGCTTTCATTTCTTATGTTTTAAATGCAC

SspI (2527)
SwaI (2541)

2501 TGACCTCCACATTCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAAATCCAGATGCT

EcoO109I (2602)

2601 CAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTAAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTA
141▶ •

2701 GTTCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAG
 140▶ N R T Y K L P I 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

BstXI (2831)

2801 ATGAGCTCTCTGCACATGCCACAGGGGGTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTCT
106 I L E R C M G C 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

StuI (2966)

2901 TCTGCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCC
73 Q G N S V A S G 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

3001 AGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTTGTTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGA
40 T K T R I A A G 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

AseI (3174)

3101 GAGATGTTGAAGGCTTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGG
6 S I N F T K M

3201 CGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTCGTCATGGGGCGGAGTTGT

SpeI (3329)

3301 TACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAACCTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTA

SnaBI (3457)

3401 TCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAAGTCATG

NdeI (3562)

3501 TACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTT

3601 ACCGTAATACTCCACCCATTGACGTCAATGGAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGGTCTTGGG

PacI (3748)

PstI (3741)

SdaI (3740)

BspLU11I (3758)

3701 CGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA
CGCTGCAGGTTAATTAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA

3801 AGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA

3901 AAGATAACAGCGTTTTCCCTGGAAAGCTCCCTCGTGCCTCTCTGTTCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGC

ApaLI (4072)

4001 GTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACC

4101 GCTGCGCCTTATCCGTAACATATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG

4201 GTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC

4301 TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAG

PacI (4488) SmaI (4497)

4401 GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATT

EagI (4508)

NotI (4507)

4501 TAAATCAGCGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAA
4601 CGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTATCGAA