



PvuI (7)
SgfI (6) 1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGCAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82) **EcoNI (96)**
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245) **Bsu36I (291)**
Psp1406I (203) **PvuII (239)** **EcoNI (287)**
201 GTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTACCGCGCCCGCCCTACCTGAGGGCC
301 GCCATCCACGCGGTTGAGTCGCGTTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTGGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

NcoI (560) **BstEII (555)** **AgeI (552)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCCTACCTGAGATCAcgggtcaCCATGGCCCCGCGCCCCGGCGCCGCCCGCTGTTCCG
1 M A P R A R R R R P L F A

PstI (631)
601 GCTGCTGCTGCTCTGCGCGCTGCTCGCCCGCTGCAGGTGGCTTTCAGATCGCTCCTCCATGTACCAGTGAAGCATTATGAGCATCTGGGACGGTGC
13 L L L L C A L L A R L Q V A L Q I A P P C T S E K H Y E H L G R C

BspLU11I (725)
701 TGTAACAAATGTGAACCAGAAAGTACATGCTTCTAAATGCACTACTACCTCTGACAGTGTATGCTGCGCCTGTGGCCGGATGAATACTTGGATAGCT
47 C N K C E P G K Y M S S K C T T T S D S V C L P C G P D E Y L D S

SfiI (849) **NgoMIV (868)**
801 GGAATGAAGAAGATAAATGCTTGTGCATAAAGTTTGTGATACAGGCAAGGCCCTGGTGGCCGTGGTGCCTGGCAACAGCAGCACACCCCGGGCGCTGCGC
80 W N E E D K C L L H K V C D T G K A L V A V V A G N S T T P R R C A

ApaLI (900) **Acc65I (911)** **DraIII (951)**
901 GTGCACGGCTGGGTACCACTGGAGCCAGGACTGCGAGTGTGCCCGCGCAACACCGAGTGCAGCGCCGGGCTGGGCGCCAGCACCCGTTGCAGCTCAAC
113 C T A G Y H W S Q D C E C C R R N T E C A P G L G A Q H P L Q L N

DraIII (1005)
1001 AAGGACACAGTGTGCAAACTTGCCTTGCAGGCTACTTCTCTGATGCCTTTTCTCCACGGCAAATGCAGACCTGGACCAACTGTACCTTCTTGGAA
147 K D T V C K P C L A G Y F S D A F S S T D K C R P W T N C T F L G

1101 AGAGAGTAGAACATCATGGGACAGAGAAATCCGATGCGGTTTGCAGTCTTCTCTGCCAGTAGAAAACCACAAATGAACCCATGTTTACTTGCCCGG
180 K R V E H H G T E K S D A V C S S S L P A R K P P N E P H V Y L P G

1201 TTTAATAATTCTGCTTCTTCCGCTGTGGCCCTGGTGGCTGCCATCATCTTTGGCGTTTGTATAGGAAAAAGGAAAGCACTCACAGCTAATTTG
213 L I I L L L F A S V A L V A A I F G V C Y R K K G K A L T A N L

Bsu36I (1349)
1301 TGGCACTGGATCAATGAGGCTTGTGGCCGCTAAGTGGAGATAAGGAGTCTCAGGTGACAGTGTGTGTCAGTACACACACGGCAAACCTTGGTACGACAGG
247 W H W I N E A C G R L S G D K E S S G D S C V S T H T A N F G Q Q

BbrPI (1491)
1401 GAGCATGTGAAGTGTCTTACTGCTGACTCTGGAGGAGAAGACATTTCCAGAAGATATGTGCTACCCAGATCAAGGTGGTGTCTGTCAGGGCACGTGTGT
280 G A C E G V L L L T L E E K T F P E D M C Y P D Q G G V C Q G T C V

1501 AGGAGGTGGTCCCTACGCACAAGGCGAAGATGCCAGGATGCTCTATTGGTCAAGACCGAGATAGAGGAAGACAGCTTCAGACAGATGCCACAGAA
313 G G G P Y A Q G E D A R M L S L V S K T E I E E D S F R Q M P T E

1601 GATGAATACATGGACAGGCCCTCCAGCCACAGACCACTGTTCTCTCAGTGGAGCAATCCACACCTCTTCTGAAACCCCTGGAGG
347 D E Y M D R P S Q P T D Q L L F L T E P G S K S T P P F S E P L E

1701 TGGGGGAGAATGACAGTTAAGCCAGTCTTACGGGGACACAGAGCACAGTGGGTTAGAAAGTGCAACTGCACTGAGCCCTGTGCAGGACTGATTG
380 V G E N D S L S Q C F T G T Q S T V G S E S C N C T E P L C R T D W

1801 GACTCCCATGTCTCTGAAAACCTTGGAAAAGAGGTGGACAGTGGCCATTGCCCGCACTGGGCGAGCCAGCCCACTGGGCGAGATGTCTGC
413 T P M S S E N Y L Q K E V D S G H C P H W A A S P S P N W A D V C

XcmI (1971)
1901 ACAGGCTGCCGAACCTCTGGGGAGGACTGTGAACCCCTCGTGGGTTCCCAAAACGTGGACCTTGGCCAGTGGCCTATGGCATGGGCTTCCCC
447 T G C R N P P G E D C E P L V G S P K R G P L P Q C A Y G M G L P

XcmI (2034)
2001 CTGAAGAAGAAGCCAGCAGGACGGAGGCCAGACCCGAGGATGGGGCTGATGGGAGGCTCCCAAGCTCAGCGAGGGCAGGTGCCGGTCTGGAAG
480 P E E E A S R T E A R D Q P E D G A D G R L P S S A R A G A G S G S

2101 CTCCCTGGTGGCAGTCCCTGCATCTGAAAATGTGACTGAAAACAGTAACCTCACGTTTCTCCAGCGGGCAGGTGATGAACCTCAAGGGCGACATC
513 S P G G Q S P A S G N V T G N S N S T F I S S G Q V M N F K G D I

NcoI (2255) **AscI (2286)**
2201 ATCGTGGTCTACGTGAGCCAGCCTCGCAGGAGGGCGCGCGGCTGCGGAGCCCATGGCCGCGCCGCTGCAAGGAGGACCTGGCGCGCCGAGACT
547 I V V Y V S Q T S Q E G A A A A A E P M G R P V Q E E T L A R R D

XhoI (2373) **StuI (2369)**
2301 CCTTCGCGGGAAACGCCCGCGCTTCCCGACCCGTGCGGCGGCCCGAGGGCTGCGGGAGCCGGAGAAGGCCTCGAGGCGGTTGCAAGGAGCAAGGCGG
580 S F A G N G P R F P D P C G G P E G L R E P E K A S R P V Q E Q G G

NheI (2413)
2401 GGCCAAGGCTTGAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAATAAATGCTTTATTTGTGAA
613 A K A •

HpaI (2551) **MfeI (2562)**
2501 ATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTAAACAACAACATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGG

2601 TGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACCTTTAACCTCCAAATCAAGCCTCTACT
 EcoRI (2647)
 2701 TGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGA

2801 TATAGTGATTTTCCCAAGGTTTGAAGTACTCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATGAAAAATTTCAGAAAATA
 SspI (2886)
 2901 ATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTT
 SmaI (2900)

3001 AGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGAGGGGATGAGTTCCTCAATGGTGG
 141 • N R T Y K L P I L E E I T T
 3101 TTTTGACCAGCTTGCATTTCATCTCAATGAGCACAAGCAGCTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGAT
 SacI (3161) BstXI (3190)
 126 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 P S V V R I
 3201 GGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCA
 93 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 I A I A E

3301 GCACAGACAGTGACCCTGCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTGTTGT
 StuI (3325)
 59 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 V H H K N D

3401 CCTCATAGAGCATGGTATCTTCTCAGTGGCGACTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGATGGCCCTCTATAGTGAGT
 XmnI (3467)
 26 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

3501 CGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCT
 AseI (3533) SacI (3590)

3601 TATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAC
 SpeI (3688)

3701 AAACCTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGG

3801 TAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCCATAAGGTCATGTAAGTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGAC
 SnaBI (3816)

3901 GTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCT
 NdeI (3921)

4001 ATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGC
 SdaI (4099)

4101 CTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGGCGCGTGTGGCGTTTTTCCATAGGCTCCGCCCCCT
 PacI (4107) BspLU11I (4117)

4201 GACGAGCATCACAAAATCGACGCTCAAGTCAAGGTTGGCGAAACCCGACAGGACTATAAAGATACCAGCGTTTTCCCTGGAAGCTCCCTCGTGGCT
 4301 CTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCATAGCTACGCTGTAGGTATCTCAGTTC

4401 GGTGTAGGTCGTTTCGCTCAAGCTGGGCTGTGTGCACGAACCCCCGTTGAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTGAGTCAACCCG
 ApaLI (4431)

4501 GTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTA
 4601 ACTACGGTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAC
 4701 CACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGAC

4801 GCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAATTAATTAACATTTAAATCAGCGGCCGCAATAAAAATATCTTTATTTTTCATTACA
 PacI (4847) SmaI (4856) EagI (4867) NotI (4866)

4901 TCTGTGTGTTGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCGAGTGCA
 5001 AGTGCAGGTGCCAGAACATTTCTCTATCGAA