



PvuI (7)
SgfI (6) 1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA

MfeI (82)
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTATGTCGTGACTGGCTCCGCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCGCTACCTGAGGCC

PvuII (239)
Bsu36I (291) 301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCAGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

AgeI (552) 501 TCTGTTCTGCCCGTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGTAGGAGGCCAACATGTACCTTGAGGTTAGTGAACGTCAAGTGC

BspLU11I (567)
1 M Y L E V S E R Q V
601 TAGATGCCTCGACTTTGCTTTCTTCTGAAAAACAGCACCTCCTCCTACGATTATGGGAAAAACGAGAGCGACTTCTGACTCCCCGCCCTGCCACA

11 L D A S D F A F L L E N S T S P Y D Y G E N E S D F S D S P P C P Q
701 GGATTTGAGCCTGAACCTTGACAGAACCTTCTGCCAGCCCTACAGCCTCCTCTTCTGCTGGGCTGCTAGGCAATGGGCGGTGGCTGTGTGCTA

44 D F S L N F D R T F L P A L Y S L L F L L G L L G N G A V A A V L
801 CTGAGTCAGCGCACTGCCCTGAGCAGCACGGACACCTTCTGCTCCACCTGGCTGTAGCCGATGTTCTGCTGGTGAACCTTCCATTGTGGCAGTGG

HpaI (874)
78 L S Q R T A L S S T D T F L L H L A V A D V L L V L T L P L W A V
901 ATGCTGTGTCCAGTGGGTTTTGCGCCCTGGCCTCTGCAAAGTGGCAGGCGCCTTGTCAACATCAACTTCTATGCAGGGCCTTCTGCTGGCTGTAT

111 D A A V Q W V F G P G L C K V A G A L F N I N F Y A G A F L L A C I
XmaI (1056)
1001 AAGCTTCGACAGATATCTGAGCATAGTGCAGCCACCCAGATCTACCGCAGGGACCCCGGGTACGTGTAGCCCTCACCTGCATAGTTGTATGGGTCTC

144 S F D R Y L S I V H A T Q I Y R R D P R V R V A L T C I V V W G L
1101 TGCTGCTCTTTGCCCTCCAGATTTTCTACCTATCAGCCAACTACGATCAGCGCCTCAATGCCACCCATTGCCAGTACAACCTCCACAGGTGGGTC

178 C L L F A L P D F I Y L S A N Y D Q R L N A T H C Q Y N F P Q V G
PstI (1217)
1201 GCACTGCTCTGCGTGTACTGCAGCTAGTGGCTGGTTTCTGCTGCCCTTCTGGTCTGCGCTACTGCTATGCCCATATCCTAGCTGTTCTGCTGGTCTC

211 R T A L R V L Q L V A G F L L P L L V M A Y C Y A H I L A V L L V S
DraIII (1379)
1301 CAGAGGCCAGAGCGTTTTTCGAGCTATGAGGCTAGTGGTAGTGGTGGTGGCAGCCTTTGCTGTCTGCTGGACCCCTATCACCTGGTGGTCTAGTGGAT

244 R G Q R R F R A M R L V V V V V A A F A V C W T P Y H L V V L V D
BbrPI (1448) **MscI (1458)**
1401 ATCCTCATGGATGTGGGAGTTTTGGCCCGCAACTGTGGTCGAGAAAGCCAGTGGATGTGGCCAAGTCACTCACCTCGGGCATGGGGTACATGCACTGCT

278 I L M D V G V L A R N C G R E S H V D V A K S V T S G M G Y M H C
BsrBI (1576) **Bsp120I (1592)**
1501 GCCTCAATCCGCTGCTCTATGCCTTTGTGGAGTGAAGTTCAGAGAGCAAATGTGGATGTTGTTACGCGCCTGGGCGCTCTGACCAGAGAGGGCCCA

311 C L N P L L Y A F V G V K F R E Q M W M L F T R L G R S D Q R G P Q
StuI (1650) **NheI (1695)**
1601 GCGGCAGCCGTATCTTACGGAGAGAATCATCTGGTCTGAGACAACCTGAGGCTCTACCTGGGCTTGAATTCTGGACTGGAACCTGTAGCCTGTAG

344 R Q P S S S R R E S S W S E T T E A S Y L G L •
MscI (1701)
1701 CTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCT

HpaI (1833) **MfeI (1844)**
1801 TTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACAAATGCATTCAATTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGTTTTTTAA

EcoRI (1929)
1901 GCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAATAACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGG

2001 ATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAA

SapI (2111) 2101 GGTGTTGAACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCTTTTTAGTAAAATATTCAGAAAATATTTAAATACATCATTGC

SspI (2168) **SwaI (2182)**
2201 AATGAAAATAAATGTTTTTATTAGGCAGAAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGACCTT

2301 TAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCAT

141 • N R T Y K L P I L E E I T T K V L K G N
SacI (2443) **BstXI (2472)**
2401 TCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATC

120 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 S R D V E D

2501 AGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAAATGGCAATGGCTTCAGCACAGACAGTGACCCCTG
87 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 A C V T V R

StuI (2607)

2601 CCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCCCGACATGGTGTCTTGTGCTCATAGAGCATGGTGA
53 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 E Y L M T I

BbsI (2753)

XmnI (2749)

2701 TCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGA
20 K E T A V E V L E L D Q Q S I N F T K M

AseI (2815)

SacI (2872)

2801 TATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCCACCG

SpeI (2970)

2901 TACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCA

SnaBI (3098)

3001 ATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATA

3101 CGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACT

NdeI (3203)

3201 TGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGA

PacI (3389)

PstI (3382)

SdaI (3381)

BspLU11I

3301 ACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAA

3401 CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAAT

3501 CGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGC

3601 CGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTC

ApaLI (3713)

3701 CAAGCTGGGCTGTGTGCACGAACCCCCGTTGAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAAGTCCAACCCGGTAAGACACGACTTATCG

3801 CCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAA

3901 GAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACCACCGTGGTAGCGGTGG

4001 TTTTTTGTGGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAAAAC

EagI (4149)

PacI (4129) SwaI (4138) NotI (4148)

4101 TCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAAATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTT

4201 GTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACA

4301 TTTCTCTATCGAA