



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82)
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGCC
HindIII (245)
Bsu36I (291)
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCCGCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGTCAAGTGCAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

KasI (535) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCCTAGTGCAGTACCCGGTCCGGTGCAGCAGTCTTCTCACCTGAACCTGGA
BstEII (555)
AgeI (552) 1▶ M A A A A A A S H L N L D
601 TGCCCTCCGGAAAGTGCTAGAATGTCCCATCTGCATGGAGTCTTCACTGAAGAGCAGCTGCGACCCAAGCTGCTGCACTGTGGCCATACCATCTGCCG
13▶ A L R E V L E C P I C M E S F T E E Q L R P K L L H C G H T I C R
701 CAGTGTCTGGAGAAGCTCCTGGCCAGCAGCATCAATGGCGTCCGCTGCCCTTTTGCAGCAAGATTACTCGCATCACCAGCCTGACCCAGCTGACCGACA
47▶ Q C L E K L L A S S I N G V R C P F C S K I T R I T S L T Q L T D
801 ACCTGACGGTGTGAAGTACATGACACAGCTGGGCTCAGTGAAGCCGCTCCGCTGCTCATGTGCCGAGGCTGTGCCCGGGCTGCCTCGCCAGTTCTG
80▶ N L T V L K I I D T A G L S E A V G L L M C R G C G R R L P R Q F C

BsrBI (994) 901 CCGAAGCTGTGGTGTGGTGTGTGTAACCTGCCGGAGGCAGATCACCAACCCCTGGCCACTGCACACTTCCGGTCAAGGAGGCAGCTGAGGAGCGG
113▶ R S C G V V L C E P C R E A D H Q P P G H C T L P V K E A A E E R
1001 CCGAGGACTTCCGGGAGAAGTTGACTCGTCTAAGGAACTTACTGGAGACTGCAGAGGAGGAAAGGCAGCCTTGGAGGGCTCCAGGGACTTCCAGG
147▶ R R D F G E K L T R L R E L T G E L Q R R K A A L E G V S R D L Q

BsrBI (1164) 1101 CAAGTATAAGGCTGTTCTTCAAGAATATGGCCATGAGGAACCCGCATCCAGGAAGAGCTAGCCCGCTCCGGAAGTCTTTCACAGGCTCCTTGGCTGA
NheI (1158) 180▶ A R Y K A V L Q E Y G H E E R R I Q E E L A R S R K F F T G S L A E
1201 GGTGAGAAGTCCAACAGTCAAGTGGTAGAGGAGCAGACTACCTACTCAACATTGCTGAGGTGCAGGCCGTGCTCGCTGTGACTACTTTCTAGCGAAG
213▶ V E K S N S Q V V E E Q S Y L L N I A E V Q A V S R C D Y F L A K

XmaI (1373) 1301 ATCAAGCAAGCTGATGTAGCCCTCCTGGAGGAGACAGCGGATGAGGAGGAGCCCGAGTCACTGCCAGCCTACCCGGAGCTTACCCTGCAAGATGTGG
247▶ I K Q A D V A L L E E T A D E E E P E L T A S L P R E L T L Q D V

HpaI (1465) 1401 AGCTCCTAAGGTAGGACACGTTGGTCTCTGCAAAATTGGCCAGGCTGTTAAGAAGCCCGGACAGTTAACATGGAAGATTCTGGGAGGGGAGGAGGG
280▶ E L L K V G H V G P L Q I G Q A V K K P R T V N M E D S W A G E E G

BstAPI (1504) 1501 AGCAGCATCTTCTGCCTCAGCCTCGTAACCTTTAGAGAGATGGACATGAGCCCTGAGGAAGTAGCTCCAGCCCTAGGGCTTCCCCCGGAACAGCGG
313▶ A A S S A S A S V T F R E M D M S P E E V A P S P R A S P A K Q R
1601 AGTTCTGAGGACGTTCCGGTATCCAGCAGTGTCTGTTTCTCAAGAAGTGGGGGCGAAAGGCAGCACTCCCGGCATGTTCAATCTTCCAGTCACTCT
347▶ S S E A A S G I Q Q C L F L K K M G A K G S T P G M F N L P V S L
1701 ATGTGACCAGTCAAGTGTGGTGTGGTGGCCAGCCGGGCAACTATCGTATCCAAGTGTTCACCCGAAAGGCTTTTGAAGGAGATCCGCCAGGCC
380▶ Y V T S Q S E V L V A D R G N Y R I Q V F N R K G F L K E I R R S P

HindIII (1824) 1801 CAGCGCATTGATAGCTTCGTGCTAAGCTTCTTGGAGCCGACTTGCCAATCTCACTCCACTTTTCACTGGCCATGAAGTCCATGGGCTGATTGGTGT
413▶ S G I D S F V L S F L G A D L P N L T P L S V A M N C H G L I G V

DraIII (1945) 1901 ACTGATAGCTACGACAACCTCCCTTAAAGTCTATACCATGGATGCCACTGTGTGGCCTGTACAGGAGCCAGCTGAGCAAACCATGGGCATCACAGCCC
447▶ T D S Y D N S L K V Y T M D G H C V A C H R S Q L S K P W G I T A
2001 TACCATTGGCCAGTTTGGTGTGACTGACGTGGAAGCGGGGAAAGCTCTGGTGTTCCTGTTAGACCGAGGAGCAGGAGTGGTCAAATACAGCTGCCTCG
480▶ L P S G Q F V V T D V E G K L W C F T V D R G A G V V K Y S C L C
2101 CAGTGTGTGAGGCCAAGTTTGTACCTGTGATGCTGAAGGCACAGTCTACTTACCCAAAGGCTTGGGGCTCAATGTGGAACCCGACAGAATGAACAC
513▶ S A V R P K F V T C D A E G T V Y F T Q G L G L N V E N R Q N E H
2201 CACCTGGAGGGTGGCTTCTCATCGTTTCTTGGCCCGATGGGCAGCTGGCCGGCAAATCAGCCACTTCTTCTGAGAATGAAGATTTCCGCTGCA
547▶ H L E G G F S I G S V G P D G Q L G R Q I S H F F S E N E D F R C
2301 TCGCTGGCATGTGTGTAGATGCTCGGGCGACCTATTGTGGCAGATGACAGCCGCAAGGAAATTCTCCATTTTCCAAAGGGCGGTGGCTACAGCGTCT
580▶ I A G M C V D A R G D L I V A D S S R K E I L H F P K G G G Y S V L

EcoO109I (2408) 2401 TATTCGAGAGGGCCTTACCTGTCCAGTGGGCATCGCCCTCACACCAAGGGGCGAGTGTGGTCTTGGACTGTTGGGATCACTGCGTCAAGATCTACAGC
613▶ I R E G L T C P V G I A L T P K G Q L L V L D C W D H C V K I Y S

BglIII (2489)
SpeI (2541) 2501 TATCATCTGAGAAGATATCCACCCCTTAAAGATGAGGAGACTAGTCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAACCACT
647▶ Y H L R R Y S T P •

HpaI (2684) 2601 AGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAAT
MfeI (2695)

EcoRI (2780) 2701 GCATTATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGATTTGAATTTAAATACAGCATA
2801 GCAAACTTTAACCTCCAATCAAGCCTTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGCTGTGGCAATGTGCATTAGCTGT

2901 TTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTTGAAGCTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCC

3001 CACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCC **SspI (3019)** **Swal (3033)** **EcoO109I (3094)**

3101 TTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGG **141** • N R T

3201 TGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTC
137 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 I L E

3301 TCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCG
104 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 Q G

3401 TTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGG **StuI (3458)**
70 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 T K T

3501 TCCTGATGGCCGCCCGACATGGTGTCTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTT
37 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 S I N

3601 GAAGTCTTCATGGTGGCCCTCCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACGCTGGATGGCGTCTCCA **BbsI (3604)** **AseI (3666)**
4 F T K M

3701 GCTTATCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTCGCTCAATGGGGCGGAGTTGTTACGACAT

3801 TTTGAAAGTCCCCTTGTATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGTGGAGACTTGAAATCCCCTGAGTCAAACCGCTATCCACGCC **SpeI (3821)**

3901 CATTGATGTAAGTCCAAAACCGCATCATCATGGAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTGATGTAAGTGGG **SnaBI (3949)**

4001 ATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTAAGTGGGCGAGTTTACCGTAAA **NdeI (4054)**

4101 TACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTCAGC

4201 CAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCGCG **PacI (4240)** **SdaI (4232)** **BspLU11I (4250)**

4301 TTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACC

4401 AGGCGTTTTCCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGCT

4501 TTCTCATAGCTCACGCTGAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCCGACCGCTGCGCC **ApaLI (4564)**

4601 TTATCCGGTAACTATCGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAA

4701 GCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAA

4801 AAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACCGCGAGAAAAAAGGATCTCAA

4901 GAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAG **PacI (4980)** **Swal (4989)** **NotI (4999)**

5001 CGGCCGAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAA **EagI (5000)**

5101 AACAAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA