



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA **MfeI (82)**
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC **HindIII (245)** **Bsu36I (291)**
301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT **NgoMIV (441)**

501 TCTGTTTGTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCAACATGTGGACCCTGGTGAAGTGGGTGGCCTTAACAGCAGG **KasI (535)** **AgeI (552)** **BspLU11I (560)**
601 GCTGGTGGCTGGAACGCGGTGCCAGATGGTCACTTCTGCCCTGTGGCCTGCTGCTGGACCCCGGAGGAGCCAGCTACAGCTGCTGCCGCTCCCTTCTG **1 M W T L V S W V A L T A G**
13 L V A G T R C P D G Q F C P V A C C L D P G G A S Y S C C R P L L

701 GACAAATGGCCACAACACTGAGCAGGCATCTGGTGGCCCTGCCAGTTGATGCCACTGCTCTGCCGGCCACTCTGCATCTTTACCGTCTCAGGGA **NgoMIV (766)**
47 D K W P T T L S R H L G G P C Q V D A H C S A G H S C I F T V S G

801 CTTCCAGTTGCTGCCCTTCCAGAGGCGGTGGATGGCCATCACTGCTGCCACGGGGCTTCCACTGCAGTGCAGACGGCGATCCTGCTT **SphI (832)** **BsaBI (841)** **MscI (843)** **PstI (873)**
80 T S S C C P F P E A V A C G D G H H C C P R G F H C S A D G R S C F

901 CCAAAGATCAGGTAACAACCTCCGTGGGTGCCATCCAGTGCCTGATAGTCAAGTGCAGGACTTCTCCACGTGCTGTGTTATGGTCGATGGCTCC **BstBI (952)**
113 Q R S G N N S V G A I Q C P D S Q F E C P D F S T C C V M V D G S

1001 TGGGGTGTGCCCATGCCAGGCTTCTGTGTGAAGACAGGGTGCCTGCTGCTCCGACGGTGCCTTCTGCGACCTGGTTCACACCCGCTGCATCA **BbsI (1036)** **ApaLI (1045)**
147 W G C C P M P Q A S C C E D R V H C C P H G A F C D L V H T R C I

1101 CACCCACGGGCACCCACCCCTGGCAAAGAAGCTCCCTGCCAGAGGACTAACAGGGCAGTGGCCTTGTCCAGCTCGGTGATGTGCCGGACGCGGTC **BspEI (1185)**
180 T P T G T H P L A K K L P A Q R T N R A V A L S S S V M C P D A R S

1201 CCGGTGCCCTGATGGTTCTACCTGCTGTGAGCTGCCAGTGGGAGTATGGTGTGCCCAATGCCAACGCCACCTGCTGCTCCGATCACTGCACTGC **213 R C P D G S T C C E L P S G K Y G C C P M P N A T C C S D H L H C**
1301 TGCCCCCAAGACTGTGTGACCTGATCCAGAGTAAAGTGCCTTCCAAGGAGAACGTTACCACGGACTCCTCACTAAGCTGCCTCCGACACAGTGG **247 C P Q D T V C D L I Q S K L C S K E N A T T D L L T K L P A H T V**
1401 GGGATGTAAATGTGACATGGAGGTGAGCTGCCAGATGGCTATACCTGCTGCCGCTACAGTCCGGGGCTGGGGCTGCTGCCCTTTTACCCAGGCTGT **280 G D V K C D M E V S C P D G Y T C C R L Q S G A W G C C P F T Q A V**

1501 GTGCTGTGAGGACCACATACACTGCTGCCGGGGTTTACGTGTGACACGAGGAGGTTACCTGTGAACAGGGGCCACCCAGGTGCCCTGGATGGAG **SacII (1529)** **Acc65I (1558)** **Bsp120I (1573)**
313 C C E D H I H C C P A G F T C D T Q K G T C E Q G P H Q V P W M E

1601 AAGCCCCAGCTCACCTCAGCCTGCCAGACCCACAAGCCTTGAAGAGAGATGTCCCTGTGATAATGTGACGAGCTGTCCCTCCTCCGATACTGTGCC **XcmI (1698)**
347 K A P A H L S L P D P Q A L K R D V P C D N V S S C P S S D T C C

1701 AACTCACGCTGGGGAGTGGGGCTGCTCCAATCCAGAGGCTGTCTGCTCGGACCACAGACTGCTGCTCCAGGGCTACACTGTGTAGTGA **380 Q L T S G E W G C C P I P E A A V C S D H Q H C C P Q G Y T C V A E**
1801 GGGCAGTGTGAGGAGGAGGAGATCGTGGCTGGACTGGAGAAGATGCCTGCCCGCGGGCTTCTTATCCACCCAGAGACATCGGCTGTGACCAG **413 G Q C Q R G S E I V A G L E K M P A R R A S L S H P R D I G C D Q**

1901 CACACCAGCTGCCGGTGGGGCAGACCTGCTGCCGAGCCTGGTGGGAGTGGGCTGCTGCCAGTTGCCCATGCTGTGTGCTGCGAGGATCGCCAGC **XcmI (1904)**
447 H T S C P V G Q T C C P S L G G S W A C C Q L P H A V C C E D R Q

2001 ACTGTGCCCGGCTGGCTACACCTGCAACGTGAAGGCTCGATCCTGCGAGAAGGAAAGTGGTCTCTGCCAGCTGCCACCTCCTGGCCGTAGCCCTCA **BstXI (2075)**
480 H C C P A G Y T C N V K A R S C E K E V V S A Q P A T F L A R S P H

2101 CGTGGGTGTGAAGGACGTGGAGTGTGGGAAGGACTTCTGCCATGATAACAGACTGCTGCCGAGACAACCGACAGGGCTGGGCTGCTGCTCCCTAC **513 V G V K D V E C G E G H F C H D N Q T C C R D N R Q G W A C C P Y**

2201 CGCCAGGGCTGTTGTGCTGATCGGGCCACTGCTGCTGCTGGCTTCCGCTGCCAGCCAGGGGTACCAAGTGTGCGCAGGGAGGCCCGCGCT **KasI (2225)** **Acc65I (2266)**
547 R Q G V C C A D R R H C C P A G F R C A A R G T K C L R R E A P R

2301 GGGACGCCCTTTGAGGGACCCAGCCTTGAGACAGCTGCTGTGAGGGACAGTACTGAAGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTG **SandI (2315)** **SeaI (2349)** **NheI (2360)**
580 W D A P L R D P A L R Q L L •

2401 GACAAACCACAAC TAGAATGCAGTGA AAAAAATGCTTTATTTGTGAAATTTGTATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGT **HpaI (2498)**

2501 TAACAACAACAATTGCATTATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTC **MfeI (2509)** **EcoRI (2594)**

2601 TAAAAACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTGAACTCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAA

2701 TGTGATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTGAACTAGCTCTTCATTTCTTTATGTTTTAA

2801 ATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTGAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCA

2901 GATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTA

3001 GCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAG
141 • 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

3101 TCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAA
108 D S 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

3201 AGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGAT
75 D K 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

3301 CTCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTCTTGTGTCCCTCATAGAGCATGGTGTCTCTCAGTGGCGACCTCCACCAGCTCCAGATCC
42 E G 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

3401 TGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGT
8 Q Q S I N F T K M

3501 GGATGGCTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTGCGTCAATGGGGCGG

3601 AGTTGTTACGACATTTTGAAAGTCCCCTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAA

3701 CCGCTATCCACGCCATTGATGACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAG

3801 GTCATGACTGGGCATAATGCCAGGCGGGCATTACCCTGATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGACTGCCAAGTGGG

3901 CAGTTTACCGTAAATACTCCACCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGGGGGGT

4001 GTTGGGCGTCCAGCCAGGCGGGCATTACCGTAAGTTATGTAACGCCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACC

4101 GTAAAAAGGCCGCTTGTGCGCTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG

4201 ACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTGTGCGCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCTTCG

4301 GGAAGCGTGGCGTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTTCGCTCCAAGCTGGCTGTGTGCACGAACCCCCGTTCCAGC

4401 CCGACCGTGCCTTATCCGGTAACTATCGTCTTGTAGTCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG

4501 AGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA

4601 GTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAA

4701 AAAAAAGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATT

4801 AACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAA

4901 ACAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCGAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA