



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

101 GAGAAAGTGGCGGGGTAACGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCGCTACCTGAGGCC **HindIII (245)** **Bsu36I (291)**

301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT **NgoMIV (441)**

501 TCTGTTTGTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCAACATGTGTACCAGCAGTTGGTCATCTCTGGTTTTCCCT **KasI (535)** **AgeI (552)** **BspLU11I (560)**
 1 M C H Q Q L V I S W F S L **BspEI (675)**

601 GGTTTTCTGGCATCTCCCTCTGGCCATGGGAACCTAAGAAAGATGTTTATGTCGTAGAATTGGATTGGTATCCGGATGCCCTGGAGAAATGGTG **NdeI (627)** **BsaBI (667)**
 13 V F L A S P L V A I W E L K K D V Y V V E L D W Y P D A P G E M V
 701 GTCCTCACCTGTGACACCCTGAAGAAGATGGTATCACCTGGACCTTGACCAGAGCAGTGAGGTCTTAGGCTCTGGCAAACCTGACCATCCAAGTCA
 47 V L T C D T P E E D G I T W T L D Q S S E V L G S G K T L T I Q V
 801 AAGAGTTTGGAGATGCTGGCCAGTACACCTGTCAAAAGGAGGCGAGTTCTAAGCCATTCGCTCCTGCTGCTTCAAAAAAGGAAGATGGAATTTGGTC
 80 K E F G D A G Q Y T C H K G G E V L S H S L L L L H K K E D G I W S
 901 CACTGATATTTAAAGACCAGAAAGAACCCAAAAATAAGACCTTTCTAAGATGCGAGGCCAAGAATTATTCTGGACGTTTACCTGTGCTGGTGCAGC
 113 T D I L K D Q K E P K N K T F L R C E A K N Y S G R F T C W W L T

1001 ACAATCAGTACTGATTTGACATTCAGTGTCAAAGCAGCAGAGGCTTCTTGACCCCAAGGGTGACGTGCGGAGCTGTACACTCTCTGCAGAGAGAG **SapI (1043)**
 147 T I S T D L T F S V K S S R G S S D P Q G V T C G A A T L S A E R

1101 TCAGAGGGGACAACAAGGAGTATGAGTACTCAGTGGAGTGCCAGGAGCAGTGCCTGCCAGCTGCTGAGGAGAGTCTGCCATTGAGGTCATGGTGA **XcmI (1181)**
 180 V R G D N K E Y E Y S V E C Q E D S A C P A A E E S L P I E V M V D
 1201 TGCCGTTTCAAGCTCAAGTATGAAAACACACAGCAGCTTCTCATCAGGGACATCATCAAACCTGACCACCAAGAACTTGACGCTGAAGCCATTA
 213 A V H K L K Y E N Y T S S F F I R D I I K P D P P K N L Q L K P L
 1301 AAGAATTTCTGGCAGGTGGAGGTCAGCTGGGAGTACCCTGACACCTGGAGTACTCCACATTCTACTTCTCCCTGACATTCTGCTTCCAGGTCAGGGCA
 247 K N S R Q V E V S W E Y P D T W S T P H S Y F S L T F C V Q V Q G

1401 AGAGCAAGAGAGAAAAGAAAGATAGAGTCTTACGGACAAGACCTCAGCCACGGTCATCTGCCGAAAAATGCCAGCATTAGCGTGGGGCCAGGACCG **Bsp120I (1487)**
 280 K S K R E K K D R V F T D K T S A T V I C R K N A S I S V R A Q D R

1501 CTAATCAGTACTGATTTGACATTCAGTGTCAAAGCAGCAGAGGCTTCTTGACCCCAAGGGTGACGTGCGGAGCTGTACACTCTCTGCAGAGAGAG **Ace65I (1560)** **KasI (1573)**
 313 Y Y S S S W S E W A S V P C S V P G V G V P G V G A R N L P V A T

1601 CCAGACCCAGGAATGTTCCCATGCCTTACCACCTCCCAAAACCTGCTGAGGGCCGTGACGCAACATGCTCCAGAAGGCCAGACAAACTCTAGAATTTTACC **XbaI (1686)**
 347 P D P G M F P C L H H S Q N L L R A V S N M L Q K A R Q T L E F Y

EcoRV (1729) 1701 CTTGCACTTCTGAAGAGATTGATCATGAAGATATCACAAAAGATAAAACCAGCACAGTGGAGGCTGTTTACCATTGGAATTAACCAAGAATGAGAGTTG **BspHI [m] (1720)** **BstXI (1748)** **StuI (1760)**
 380 P C T S E E I D H E D I T K D K T S T V E A C L P L E L T K N E S C
 1801 CCTAAATTCAGAGAGACCTTTTATACTAATGGGAGTTGCTGCGCTCCAGAAAAGACCTTTTTATGATGGCCCTGTGCTTAGTAGTATTTATGAA
 413 L N S R E T S F I T N G S C L A S R K T S F M M A L C L S S I Y E

1901 GACTTGAAGATGTACCAGGTGGAGTCAAGACCATGAATGCAAAGCTGCTGATGGATCCTAAGAGGCAGATCTTTCTAGATCAAAACATGCTGGCAGTTA **BamHI (1953)** **BglII (1967)**
 447 D L K M Y Q V E F K T M N A K L L M D P K R Q I F L D Q N M L A V
 2001 TTGATGAGCTGATGCAGCCCTGAATTTCAACAGTGTGAGCTGTGCCACAAAAATCCTCCCTTGAAGAACCAGGATTTTATAAACTAAAATCAAGCTCTG
 480 I D E L M Q A L N F N S E T V P Q K S S L E E P D F Y K T K I K L C

2101 CATACTTCTTCTGCTTTTCAAGATTCGGGAGTACTATTGATAGAGTGTGAGCTATCTGAATGCTTCTTAAAAAGCGAGGTCCTCCGCTAGCTGGCC **NheI (2189)**
 513 I L L H A F R I R A V T I D R V M S Y L N A S •

2201 AGACATGATAAGATACATTGATGAGTTTGACAAACCAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTT

2301 GTAACCATTATAAGCTGCAATAAACAAGTTAACAAACAATTGCATTCATTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGT **HpaI (2327)** **MfeI (2338)**

2401 AAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAAT

2501 AAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTG

2601 AACTAGCTCTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATGATAAATATTAGAAATAATTTAAATACATCATTGCAATGAA **SapI (2605)** **SspI (2662)** **SwaI (2676)**

2701 AATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAG
2801 AAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTTCATCT
141 • N R T Y K L P I L E E I T T K V L K G N M E
2901 CAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTA
118 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 S Y
3001 GGGTGCCTGACAGCCAAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATG
85 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 G I
3101 TAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTTGTTCCTCATAGAGCATGGTGTATCTTCT
51 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 K E
3201 CAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCCTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACT
18 T A V E V L E L D Q Q S I N F T K M
3301 ATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAACGAGCTCTGCTTATATAGACCTCCACCGTACAG
3401 CCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGG
3501 TGGAGACTTGAAATCCCGTGAGTCAAACCGTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGA
3601 TGTACTGCCAAGTAGGAAAGTCCATAAGTTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCAT
3701 ATGATACACTTGATGTAAGTCCGAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATAC
3801 GTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAAATTAAGAACATGTG
3901 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGC
4001 TCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCTGCCGCTTA
4101 CCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCT
4201 GGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
4301 GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAG
4401 TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT
4501 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGT
4601 TAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTGTGTGA
4701 ATCGTAACATAACATACGCTCTCCATCAAAAACAAACGAAACAAACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGACAGGTGCCAGAACATTTCTC
4801 TATCGAA