



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

101 GAGAAAGTGGCGGGGTAACCTGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGGC **PvuII (239)** **Bsu36I (291)**

301 GCCATCCAGCGCGGTTGAGTCGCGTTTCTGCCCGCTCCCGCTGTGGTGCCTCCTGAAGCTCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT **NgoMIV (441)**

501 TCTGTTTGTGCGCGGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCAACATGTCGTGCGGCCCTCCGCGCCGTGACTTCCCCGCTG **KasI (535)** **AgeI (552)** **BspLU11I (560)**
1 M S S G L R A A D F P R W

601 GAAGCGCCACATCTCGGAGCAACTGAGGCGCCGGACCGGCTGCAGAGACAGGCGTTCGAGGAGATCATCTGCAGTATAACAAATTGCTGGAAAAGTCA **KasI (626)** **BglII (699)**
13> K R H I S E Q L R R R D R L Q R Q A F E E I I L Q Y N K L L E K S
701 GATCTTCATTAGTGTGGCCAGAAAACAGGCTGAAAAGCATGACGTACCAACAGGCACGAGATAAGTCCCGGACATGATGGCAGATGGAATGACA
47> D L H S V L A Q K L Q A E K H D V P N R H E I S P G H D G T W N D
801 ATCAGCTACAAGAAATGGCCCAACTGAGGATTAAGCACCAAGAGAACTGACTGAATTACACAAGAAACGTGGGGAGTTAGCTCAACTGGTGATTGACCT
80> N Q L Q E M A Q L R I K H Q E E L T E L H K K R G E L A Q L V I D L
901 GAATAACCAATGCAGCGGAAGGACAGGAGATGCAGATGAATGAAGCAAAAATTGCAGAATGTTGCAGACTATCTGTACCTGGAGACGGAGTGCCTA
113> N N Q M Q R K D R E M Q M N E A K I A E C L Q T I S D L E T E C L

1001 GACCTGCGCACTAAGCTTTGTGACCTTGAAGAGCCAACAGACCCTGAAGGATGAATATGATGCCCTGCAGATCACTTTTACTGCCTTGGAGGGAAAAC **FspI (1004)**
147> D L R T K L C D L E R A N Q T L K D E Y D A L Q I T F T A L E G K

1101 TGAGGAAAACACTACGGAAGAGAACCAGGAGCTGGTCAACAGATGGATGGCTGAGAAAAGCCAGGAAGCCAATCGGCTTAATGCAGAGAATGAAAAAGACTC **BstXI (1136)**
180> L R K T T E E N Q E L V T R W M A E K A Q E A N R L N A E N E K D S
1201 CAGGAGCGGCAAGCCCGGCTGCAGAAAAGCTTGCAAGCAGCAAAGAACCTCTACCAGTCGAACAGGATGATGACATTGAGGCTATTGTGGATGAA
213> R R R Q A R L Q K E L A E A A K E P L P V E Q D D D I E V I V D E

1301 ACTTCTGATCACACAGAAGACCTCTCTGTGCGAGCCATCAGCAGAGCAGCCACTAAGCGACTCTCGAGCCGTGCTGGAGGCTTCTGGATTCTATCA **StuI (1380)**
247> T S D H T E E T S P V R A I S R A A T K R L S Q P A G G L L D S I

1401 CTAATATCTTTGGGAGACGCTCTGTCTCTTCTCCAGTCCCCAGGACAATGTGGATGCTCATCCTGGTTCTGGTAAAGAAGTGGGGTACCAGCTAC **BsaBI (1456)** **Acc65I (1488)**
280> T N I F G R R S V S S F P V P Q D N V D A H P G S G K E V R V P A T

1501 TGCCTTGTGTCTTCGATGCACATGATGGGGAAGTCAACGCTGTGCAGTTCAGTCCAGGTTCCCGGTTACTGGCCACTGGAGGCATGGACCGCAGGGTT **XcmI (1574)**
313> A L C V F D A H D G E V N A V Q F S P G S R L L A T G G M D R R V
1601 AAGCTTTGGGAAGTATTTGGAGAAAAATGTGAGTTCAGGGTCCCTATCTGGCAGTAATGCAGGAATTACAAGCATTGAATTTGATAGTGTGGATCTT
347> K L W E V F G E K C E F K G S L S G S N A G I T S I E F D S A G S

1701 ACCTCTTAGCAGCTTCAATGATTTTGAAGCCGAATCTGGACTGTGGATGATTATCGATTACGGCACACACTCACGGGACACAGTGGGAAAGTGTGTGTC **ClaI (1754)**
380> Y L L A A S N D F A S R I W T V D D Y R L R H T L T G H S G K V L S
1801 TGCTAAGTTCCTGCTGGACAATGCAGGATTGTCTCAGGAAGTACGACCGGACTCTCAAACCTCTGGATCTACGCAGCAAAGCTGCATAAAGACAGTG
413> A K F L L D N A R I V S G S H D R T L K L W D L R S K V C I K T V

1901 TTTGCAGGATCCAGTTGCAATGATTTGTCTGCACAGAGCAATGTGTAATGAGTGGACATTTTGACAAGAAAATTCGTTTCTGGGACATTCGATCAGAGA **BamHI (1906)**
447> F A G S S C N D I V C T E Q C V M S G H F D K K I R F W D I R S E
2001 GCATAGTTCGAGAGATGGAGCTGTTGGGAAAGATTACTGCCCTGGACTTAAACCCAGAAAGGACTGAGCTCCTGAGCTGCTCCCGTGTGACTTGCTAAA
480> S I V R E M E L L G K I T A L D L N P E R T E L L S C S R D D L L K

2101 AGTTATTGATCTCCGAACAAATGCTATCAAGCAGACATTCAGTGCACCTGGGTTCAAGTGCAGGCTCTGACTGGACCAGAGTTGTCTTTCAGCCCTGATGGC **ApaLI (2141)**
513> V I D L R T N A I K Q T F S A P G F K C G S D W T R V V F S P D G

2201 AGTTACGTGGCAGCAGGCTCTGCTGAGGGCTCTCTGTATATCTGGAGTGTGCTCACAGGGAAAGTGGAAAAGGTTCTTTCAAAGCAGCACAGCTCATCCA **XmnI (2266)** **XcmI (2297)**
547> S Y V A A G S A E G S L Y I W S V L T G K V E K V L S K Q H S S S

2301 TCAATGCGGTGGCGTGGTGCCTCTGGCTCGCACGTTGTGTCAGTGTGGACAAAGGATGCAAGCTGTGCTGTGGGCACAGTACTGACGGGCTCTCAGGG **ScaI (2378)**
580> I N A V A W S P S G S H V V S V D K G C K A V L W A Q Y •

2401 CTGGGAGGACCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTGTGAAA **EcoO109I (2405)** **MscI (2418)**

2501 **HpaI (2550)** MfeI (2561)
TTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACCAAGTTAAACAACAACAAATTGCATTCATTTTATGTTTCAGGTTACAGGGGAGGT

2601 **EcoRI (2646)**
GTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTAACTCCAATCAAGCCTCTACTT

2701 GAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTTTTTCATGGAGTTAAGAT

2801 **SspI (2885)** SwaI (2899)
ATAGTGATTTTTCCAAGTTTGAAGTACTGCTCTTCATTCTTTATGTTTTAAATGCAGCTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAA

2901 **EcoO109I (2960)**
TTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTA

3001 **BstXI (3189)**
GGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGT
141 • N R T Y K L P I L E E I T T

3101 TTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATG
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 GATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAG
92 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 A

3301 **StuI (3324)**
CACAGACAGTGACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGTG
59 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 **XmnI (3466)** **BspHI (3474)**
CTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGATGGCCCTCTATAGTGAGTC
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 **AseI (3532)**
GTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTT

3601 **SpeI (3687)**
ATATAGACCTCCCACGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAC

3700 AAACCTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGG

3800 **SnaBI (3815)**
TAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTCAATTGAC

3900 **NdeI (3920)**
GTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCT

4000 **SdaI (4098)**
ATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGTCCAGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGC

4100 **PacI (4106)** **BspLU11I (4116)**
CTG C A G G T T A A T T A A G A A C A T G T G A G C A A A A G G C C A G A A A A G G C C A G A A C C G T A A A A G G C C G T T G C T G G C G T T T T C C A T A G G C T C C G C C C C

4198 CTGACGAGCATCAGAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCG

4298 CTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAGCGTGGCGTTTCTCATAGCTCACGCTGTAGGTATCTCAGT

4398 **ApaLI (4430)**
TCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACC

4498 CGGTAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCC

4598 TAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAA

4698 ACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGCTGCTG

4798 **EagI (4866)** **NotI (4865)**
PacI (4846) SwaI (4855) AGCGGCCGCAATAAAATATCTTTATTTTCATTA

4898 CATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTG

4998 CAAGTGCAGGTGCCAGAACATTTCTCTATCGAA