



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

Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCGCTACCTGAGGCC **HindIII (245)** **Bsu36I (291)**
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTCTAGGTAAGTTAAAGTCAAGTGCAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGTCTTTGTTTCGTTT **NgoMIV (441)**

501 TCTGTTTGGCGCGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTAGGAGGGCCACCATGGCGGGAACTGTGGGCCCCGCGCGCG **BssHII (595)**
AgeI (552) **NcoI (568)** **SacII (590)**
Bsp120I (586) 1 M A G N C G A R G A

601 CTGTCGGCGCACACGCTGCTGTTTCGACCTGCCGCCGCGTGTCTGGAGAGCTCTGCGTGTCTGGACAGCTGCGACGCGCGCTGGCTGGCGCGGCC **BstAPI (608)** **BssHII (679)**
11 L S A H T L L F D L P P A L L G E L C A V L D S C D G A L G W R G

701 TGGCAGAGAGACTTTCAAGCAGCTGGCTGGATGTTCTGTCATATTGAAAAGTATGTAGACCAAGGTAAGTGAACAAGAGAATTACTTTGGTCTGGGC **XcmI (758)**
44 L A E R L S S S W L D V R H I E K Y V D Q G K S G T R E L L W S W A

801 ACAGAAAAACAAGACCATCGGTGACCTTTACAGGTCCTCCAGGAGATGGGACATCGTCGAGCTATTCATTTAATTACAACTATGGAGCAGTGTGAGT **Tth111I (850)**
77 Q K N K T I G D L L Q V L Q E M G H R R A I H L I T N Y G A V L S

901 CCTTCAGAGAAGAGTTATCAGGAAGGTGGATTTCCAAATATATTATTCAAGGAAACAGCCAATGTCCCGTGGATAATGTTCTTATTCTGAACATAATG **XcmI (958)**
111 P S E K S Y Q E G G F P N I L F K E T A N V T V D N V L I P E H N

1001 AAAAAGGAGTACTGCTTAAATCTTCCATCAGCTTTCAAATATCATAGAAGGAACTAGAAATTTCCACAAAGACTTCTAATTGGAGAAGGAGAGATTTT **ScaI (1007)**
144 E K G V L L K S S I S F Q N I I E G T R N F H K D F L I G E G E I F

1101 TGAGGTATACAGAGTGGAGATTCAAACCTAACATATGCTGTCAAATATTTAAACAGGAGAAAAAATGCAGTGAAGAAGCATTGGAAGAGGTTTTTA **Bst1107I (1104)** **NdeI (1132)**
177 E V Y R V E I Q N L T Y A V K L F K Q E K K M Q C K K H W K R F L

1201 TCTGAGCTTGAAGTTTTACTACTGTTTCATCACCACAAACATACTAGAGTTGGCTGCATATTTACAGAGACTGAGAAGTTCTGTCTGATTATCCATACA **BstBI (1366)**
211 S E L E V L L L F H H P N I L E L A A Y F T E T E K F C L I Y P Y

1301 TGAGAAATGGAACACTTTTTGACAGATTGCAGTGTGATGGTGACACGGCCCACTCCCTTGGCACATTGCAATCGGTATATTAATAGGAATATCCAAAGC **BstXI (1350)** **Bsp119I (1366)** **AseI (1379)**
244 M R N G T L F D R L Q C V G D T A P L P W H I R I G I L I G I S K A

1401 CATTCACTACCTGCACAACGTTCAACCATGCTCGGTCTGTGGCAGTATATCAAGTCAAACATCCTTTTGGATGATCAGTTTCAACCCAACTAACT **Psp1406I (1416)**
277 I H Y L H N V Q P C S V I C G S I S S A N I L L D D Q F Q P K L T

1501 GATTTTGCATGGCACACTTCCGGTCCACCTAGAACATCAGAGTTGTACCATAAATATGACCAGCAGCAGCAGTAAACATCTGTGGTACATGCCAGAAG **NcoI (1507)**
311 D F A M A H F R S H L E H Q S C T I N M T S S S S K H L W Y M P E

1601 AGTACATCAGACAGGGGAACTTTCCATTAACAGATGTCTACAGCTTTGGAATTGTAATAATGGAAGTTCTAACAGGATGTAGAGTAGTGTAGATGA **XbaI (1775)**
344 E Y I R Q G K L S I K T D V Y S F G I V I M E V L T G C R V V L D D

1701 TCCAAAACATATCCAGCTGCGGGATCTCCTTAGAGAATTGATGGAGAAGAGAGCCTGGATTGATGCTCTCTCATTCTAGATAAGAAAGTGCCTCCCTGC **SapI (1820)**
377 P K H I Q L R D L L R E L M E K R G L D S C L S F L D K K V P P C

1801 CCTCGGAATTTCTCTGCCAAGCTTCTGTTTGGCAGGCGGTGTGCTGCAACGCGGGCAAAGTTAAGACCATCAATGGATGAAGTTTTAAATACTCTTG **ScaI (1902)**
411 P R N F S A K L F C L A G R C A A T R A K L R P S M D E V L N T L

1901 AAAGTACTCAAGCCAGCTTGTATTTTGGTGAAGATCCTCCACATCACTAAAGTCTTTCAGGTGCTTCTCCTCTATTCTGGAGAATGTACCAAGTAT **Bsu36I (2054)**
444 E S T Q A S L Y F A E D P P T S L K S F R C P S P L F L E N V P S I

2001 TCCAGTGGAAAGATGATGAAAAGCCAGAATAACAATTTACTACCTTCTGATGAAGGCTGAGGACAGATAAGACTCAGAAAACCTCTTTTGAATGCAGC **HindIII (2158)**
477 P V E D D E S Q N N N L L P S D E G L R I D R M T Q K T P F E C S

2101 CAGTCTGAGGTTATGTTTCTGAGCTTGGACAAAAAGCCAGAGAGCAAGAGAAATGAGGAAAGCTTGAACATGCCAGTCTTCTTGTGAAGAAAGTTGGT **XmnI (2355)** **BstAPI (2386)**
511 Q S E V M F L S L D K K P E S K R N E E A C N M P S S S C E E S W

2201 TCCCAAAGTATATAGTTCCATCCAGGACTTAAGGCCCTATAAGGTAATATAGATCTTCTCAGAAGCTCCAGGGCATTCTTGAGGAGCAGGCCAGT **MscI (2415)**
544 F P K Y I V P S Q D L R P Y K V N I D P S S E A P G H S C R S R P V

2301 GGAGAGCAGCTGTTCTCCAAATTTTCTGGGATGAATATGAACAGTACAAAAAGAAATAATTTCTACCAGAAGATAAGAAAAAGCAAGTATTGCATA **NheI (2409)**
577 E S S C S S K F S W D E Y E Q Y K K E •

2401 GGCACCTGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTAAAAAATGCTTTATTTGTGAAATTT

2501 **HpaI (2547)** MfeI (2558) GTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACCAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTG

2601 **EcoRI (2643)** GGAGGTTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTTAACCTCCAATCAAGCCTCTACTTGAA

2701 TCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTTAAGATATA

2801 **SapI (2825)** **SspI (2882)** **SwaI (2896)** GTGTATTTTCCCAAGTTTGAAGTAGCTCTTCATTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTT

2901 AAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGG

3001 AACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTCTCGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTT
141 • N R T Y K L P I L E E I T T K

3101 **BstXI (3186)** GACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGAT
125 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 S

3201 CTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCAC
91 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 C

3301 AGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGCTTGTGTCTC
58 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 E

3401 **BbsI (3467)** **XmnI (3463)** ATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCTATAGTGAGTCGTA
25 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 (3529)** TTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATA

3601 **SpeI (3684)** TAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAAC

3701 TCCATTGACGTCAATGGGGTGGAGACTTGAATCCCGTGAGTCAAACCGTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAAT

3801 **SnaBI (3812)** AGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGCGGGCCATTTACCGTCATTGACGTCA

3901 **NdeI (3917)** ATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTG

4001 **PstI (4096)** **SdaI (4095)** GCGTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTCAGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGCTCG

4101 **PacI (4103)** **BspLU11I (4113)** AGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCTTGGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACG

4201 AGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCC

4301 TGTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTG

4401 **ApaLI (4427)** TAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAA

4501 GACACGACTTATCGCACTGGCAGCAGCCACTGGAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTA

4601 CGGTACTACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAACCACC

4701 GCTGGTAGCGGTGGTTTTTTTGGTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTC

4801 **EagI (4863)** **PacI (4843)** **SwaI (4852)** **NotI (4862)** AGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCCAATAAAATATCTTTATTTTTCATTACATCTG

4901 TGTGTTGGTTTTTTGTGTGAATCGTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTG

5001 CAGGTGCCAGAACATTTCTCTATCGAA