



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

101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCGGGTTGAGTCGGCTTTCGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCTCTAGGTAAGTTAAAGTCAAGTGCAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTGTTTCGTTT

NcoI (560) **BstEII (555)** **Tth111I (583)**
501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCGCTACCTGAGATCACCGGTACCATTGGGAGAGAATGCTGATGGTGACCAGGTCTATGGAGAA **BstEII (580)**
1► M G E N A D G D Q V M E N

601 TCTGCTTACAGTGCATGTCACCTTACATGGAAGCTGCTATTTGAAAATAATGACATACCTGATTTGGAAGTGAAGTCTCAGAGCAGGTCCAGTTCTT
13► L L Q L R C H F T W K L L F E N N D I P D L E V R I S E Q V Q F L

PstI (775)
701 GACATCAAGAACCATTGGGGATGCACAACCTCCTGGCCTACGTGAGGCACCTGAAAGGCCAGCAGGACGAAGCCCTGCAGAGCTTGAAGAAGCTGAAG
47► D I K N P L G M H N L L A Y V R H L K G Q Q D E A L Q S L K E A E

801 CCTTGATCCAGAGCAGCAGCTGAGCAAGAGAAGCCTGGCGACCTGGGCACTGTGCCTGGTGCATTACCACAGGGCAGCTTGGCAGAAGCCAGGT
80► A L I Q S E Q L S K R S L A T W G N C A W L H Y H R G S L A E A Q V

901 CTACCTGGACAAGGTGGAGAAGGTGTGCAAGGAATTTCAAGTCCCTTCCGCTACAGGCTGGAGTGTGCTGAGATGACTGTGAGGAAGGCTGGGCTTG
113► Y L D K V E K V C K E F S S P F R Y R L E C A E M D C E E G W A L

NcoI (1031)
1001 CTGAAGTGTGGAGGAGAAATTATAACAAGCCATGGCCTGCTTTCGGAAGGCTCTGAAAGTGGAGCCAGAAAACCCTGAGTACAACACTGGCTATGCAG
147► L K C G G G N Y K Q A M A C F A K A L K V E P E N P E Y N T G Y A

XbaI (1140)
1101 TCGTAGCCTATCGCAAGATTTAGATGACAACCTTATTTCTCTAGAACCTTTGAGGAAGGCTGTCCGGTTAAATCCAGAAGATCCATACCTTAAAGTTCT
180► V V A Y R Q D L D D N F I S L E P L R K A V R L N P E D P Y L K V L

HindIII (1210) **BspLU11I (1230)** **BstAPI (1271)**
1201 CCTTGCCCTAAAGCTTCAGGATTTAGGAGAACATGTTGAAGCAGAAGCACACATTGAAGAAGCCCTCAGCAGCACATCTTGCCAAAGCTATGTCATTGCG
213► L A L K L Q D L G E H V E A E A H I E E A L S S T S C Q S Y V I R

SspI (1310) **SapI (1342)** **EcoO109I (1360)**
1301 TATGCAGCAAATATTTCCGTAGGAAACATCGCGTAGACAAGCTCTTCACTTAAACAGGGCCTTGCAAGGCATCACCTTCTCTGGCTACTTACATT
247► Y A A K Y F R R K H R V D K A L H L L N R A L Q A S P S S G Y L H

1401 ATCAAAAAGGGCTCTGCTACAAGCAACAATCTCCAAGTGGAGACATCCCGAAACAGGCAGCCAGAAGGCAGGACAATGTGCAAGAATTGGCACAACA
280► Y Q K G L C Y K Q Q I S Q L R T S R N R Q P R R Q D N V Q E L A Q Q

BspHI (1506) **XcmI (1538)**
1501 GGCATTCATGAATTTCAAGAGACTTTGAAACTGAGGCCACATTTGAGATGGCCTATGTTTGCATGGCTGAAGTGCAGGCAGAAATTCACCAGTATGAA
313► A I H E F Q E T L K L R P T F E M A Y V C M A E V Q A E I H Q Y E

1601 GAAGCAGAGAGAATTTCCAGAAGGCACTGAACAACAAGACCCTCGTGGCTCACATAGAGCAGGATATCACCTCCGCTATGGCCGTTTCTACAGTTTC
347► E A E R N F Q K A L N N K T L V A H I E Q D I H L R Y G R F L Q F

BbsI (1711)
1701 ATAAGCAGTCAGAAGACAAGCAATCACCTCTACTTAAAAGGCTAAAAGTGAAGAGAAGTCTTTGCTTGGAGAACTACTGACTGCTTTGGAGAA
380► H K Q S E D K A I T L Y L K G L K V E E K S F A W R K L L T A L E K

1801 AGTGGCTGAAAGAGCTGTTTGTGAGAATGTTTCATCTGTAGAGAGTACCAGCCTTCTTGGGCTAGTCTACAACTGAAAGGGCAAGAGAAAAATGCTCTG
413► V A E R R V C Q N V H L V E S T S L L G L V Y K L K G Q E K N A L

MscI (1985) **NheI (1979)**
1901 TTTTACTATGAGAAGGCACTGAGGCTCACTGGGAAATGAACCCTGCATTCTGAATGCAGCTCACCTCTGTGACGTTAAGCTAGCTGGCCAGACATGATA
447► F Y Y E K A L R L T G E M N P A F •

2001 AGATACATTGATGAGTTTGGACAAACCAACTAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTA

HpaI (2117) **MfeI (2128)**
2101 TAAGTGCATAAACAAGTTAACAACAACAATTGCATTATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTA

EcoRI (2213)
2201 CAAATGTGGTATGGAATTTAAATAACAGCATAGCAAACTTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGG

SapI (2395)
2301 CATCAGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGATTTTCCCAAGGTTTGAACACTGCTCT

SspI (2452) **SwaI (2466)**
2401 TCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAATAATTTAAATACATCATTGCAATGAAAATAAATGTT

EcoO109I (2527)
2501 TTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCAATAATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTAAATAGAAATTTGGACA

2601 GCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGTGTTGACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCATCTCAATGAGCAC
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

2701 AAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTG
115 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 P H R

2801 ACAGCCACAATGGTGTCAAAGTCCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAA
81 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 Y A E I

2901 TGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGCTTGTGTCTCATAGAGCATGGTGATCTTCTCAGTGGCGAC
48 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 T A V

3001 CTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGA
15 E V L E L D Q Q S I N F T K M

3101 TTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCC

3201 ATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTG

3301 GAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCA

3401 AGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACCT

3501 TGATGTACTGCCAAGTGGGACGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTG

3601 ACGTCAATGGGCGGGGTCGTTGGGCGGTACGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGC
3701 CAGCAAAAGGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGA
3801 GGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAAGTCCCTCGTGGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCT

3901 GTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTG

4001 CACGAACCCCCGTTTCCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCA

4101 CTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTAT

4201 CTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTTGTGCAAG

4301 CAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGCTGTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTT

4401 TGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAACTA

4501 ACATACGCTCTCCATCAAACAAACGAAACAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA