



**PvuI (7)**  
**SgfI (6)**  
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA  
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

**HindIII (245)**  
**Psp1406I (203)** **PvuII (239)** **Bsu36I (291)**  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTACCGCGCCCGCCGCTACCTGAGGCC  
301 GCCATCCACGCGGTTGAGTGCAGTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

**KasI (535)** **BspHI (560)**  
501 TCTGTTTGTCCGGCTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGGACATCATGACTTTTAAACAGTTTTGAAGGAAGTGAAGTCTTTGT  
601 ACTTGACACACCAATAAGGATGAAGAATTTGTAGAAGAGTTTAAATAGATTAACATTTGCTAACTTCCCAAGTAGTAGTCTGTTTCAGCATCAACA  
13> L A D T N K D E E F V E E F N R L K T F A N F P S S S P V S A S T  
701 TTGGCGGAGCTGGGTTTCTTTATACCGGTGAAGGAGACACCGTGAATGTTTTCAGTTGTCATGCGGCAATAGATAGATGGCAGTATGGAGACTCAGCTG  
47> L A R A G F L Y T G E G D T V Q C F S C H A A I D R W Q Y G D S A  
801 TTGGAAGACACAGGAGAATATCCCAAATTCAGATTTATCAATGGTTTTTATTTTGAAGGCTGCACAGTCTACAAATCCTGGTATCCAAAATGG  
80> V G R H R R I S P N C R F I N G F Y F E N G A A Q S T N P G I Q N G  
901 CCAGTACAAATCTGAAAAGTGTGGGAAATAGAAATCCTTTTGGCCCTGACAGGCCACCTGAGACTCATGCTGATTATCTTTGAGAACTGGACAGGTT  
113> Q Y K S E N C V G N R N P F A P D R P P E T H A D Y L L R T G Q V  
1001 GTAGATATTTAGACACCATATACCCGAGGAACCTGCCATGTGTAGTGAAGAAGCCAGATTGAAGTCATTTCAAGAACTGGCCGGACTATGCTCATTTAA  
147> V D I S D T I Y P R N P A M C S E E A R L K S F Q N W P D Y A H L  
1101 CCCCAGAGAGTTAGCTAGTGTGGCCTCTACTACACAGGGGCTGATGATCAAGTGAATGCTTTTGTGGGGGAAAACTGAAAAATGGGAACCTG  
180> T P R E L A S A G L Y Y T G A D D Q V Q C F C C G G K L K N W E P C

**AgeI (724)** **PvuII (794)**  
701 TTGGCGGAGCTGGGTTTCTTTATACCGGTGAAGGAGACACCGTGAATGTTTTCAGTTGTCATGCGGCAATAGATAGATGGCAGTATGGAGACTCAGCTG  
47> L A R A G F L Y T G E G D T V Q C F S C H A A I D R W Q Y G D S A  
801 TTGGAAGACACAGGAGAATATCCCAAATTCAGATTTATCAATGGTTTTTATTTTGAAGGCTGCACAGTCTACAAATCCTGGTATCCAAAATGG  
80> V G R H R R I S P N C R F I N G F Y F E N G A A Q S T N P G I Q N G  
901 CCAGTACAAATCTGAAAAGTGTGGGAAATAGAAATCCTTTTGGCCCTGACAGGCCACCTGAGACTCATGCTGATTATCTTTGAGAACTGGACAGGTT  
113> Q Y K S E N C V G N R N P F A P D R P P E T H A D Y L L R T G Q V  
1001 GTAGATATTTAGACACCATATACCCGAGGAACCTGCCATGTGTAGTGAAGAAGCCAGATTGAAGTCATTTCAAGAACTGGCCGGACTATGCTCATTTAA  
147> V D I S D T I Y P R N P A M C S E E A R L K S F Q N W P D Y A H L  
1101 CCCCAGAGAGTTAGCTAGTGTGGCCTCTACTACACAGGGGCTGATGATCAAGTGAATGCTTTTGTGGGGGAAAACTGAAAAATGGGAACCTG  
180> T P R E L A S A G L Y Y T G A D D Q V Q C F C C G G K L K N W E P C

**MscI (897)**  
801 TTGGAAGACACAGGAGAATATCCCAAATTCAGATTTATCAATGGTTTTTATTTTGAAGGCTGCACAGTCTACAAATCCTGGTATCCAAAATGG  
80> V G R H R R I S P N C R F I N G F Y F E N G A A Q S T N P G I Q N G  
901 CCAGTACAAATCTGAAAAGTGTGGGAAATAGAAATCCTTTTGGCCCTGACAGGCCACCTGAGACTCATGCTGATTATCTTTGAGAACTGGACAGGTT  
113> Q Y K S E N C V G N R N P F A P D R P P E T H A D Y L L R T G Q V  
1001 GTAGATATTTAGACACCATATACCCGAGGAACCTGCCATGTGTAGTGAAGAAGCCAGATTGAAGTCATTTCAAGAACTGGCCGGACTATGCTCATTTAA  
147> V D I S D T I Y P R N P A M C S E E A R L K S F Q N W P D Y A H L  
1101 CCCCAGAGAGTTAGCTAGTGTGGCCTCTACTACACAGGGGCTGATGATCAAGTGAATGCTTTTGTGGGGGAAAACTGAAAAATGGGAACCTG  
180> T P R E L A S A G L Y Y T G A D D Q V Q C F C C G G K L K N W E P C

**BstBI (1271)**  
1201 TGATCGTGCCTGGTCCAGAACACAGGAGACACTTTCCCAATTGCTTTTTTGTGGCCGGAACGTTAATGTTTGAAGTGAATCTGGTGTGAGTCTGAT  
213> D R A W S E H R R H F P N C F F V L G R N V N V R S E S G V S S D

**Psp1406I (1261)**  
1301 AGGAATTTCCCAAATTCACAAACTCTCCAAGAAATCCAGCCATGGCAGAATATGAAGCACGGATCGTACTTTTGAACATGGACATCCTCAGTTAAACA  
247> R N F P N S T N S P R N P A M A E Y E A R I V T F G G T W T S S V N  
1401 AGGAGCAGCTTGAAGAGCTGGATTTTATGCTTTAGGTGAAGCGATAAAGTGAAGTCTTACTGTGGAGGAGGCTCACGGATTGGAAGCAAGTGA  
280> K E Q L A R A G F Y A L G E G D K V K C F H C G G G L T D W K P S E

**NeoI (1340)** **HpaI (1393)**  
1301 AGGAATTTCCCAAATTCACAAACTCTCCAAGAAATCCAGCCATGGCAGAATATGAAGCACGGATCGTACTTTTGAACATGGACATCCTCAGTTAAACA  
247> R N F P N S T N S P R N P A M A E Y E A R I V T F G G T W T S S V N  
1401 AGGAGCAGCTTGAAGAGCTGGATTTTATGCTTTAGGTGAAGCGATAAAGTGAAGTCTTACTGTGGAGGAGGCTCACGGATTGGAAGCAAGTGA  
280> K E Q L A R A G F Y A L G E G D K V K C F H C G G G L T D W K P S E

**SphI (1515)** **Acc65I (1526)**  
1501 AGACCTTTGGGAACAGCATGCGAAGTGGTACCCAGGGTCAAATACCTATTGGATGAGAAGGGGCAAGAATATATAAATAATATTCATTTAACCCATTCA  
313> D P W E Q H A K W Y P G C K Y L L D E K G Q E Y I N N I H L T H S

**ClaI (1654)**  
1601 CTTGAGGAATCTTTGGGAAGAACTGCTGAAAAACACCATCGCTAACTAAAAAATCGATGATACCATCTCCAGAATCCTATGGTGAAGAAGCTATAC  
347> L E E S L G R T A E K T P S L T K K I D D T I F Q N P M V Q E A I

**EcoO109I (1781)** **BglII (1794)**  
1701 GAATGGGATTTAGCTTCAAGACATTAAGAAAACAATGGAAGAAAAAATCAAACATCCGGGAGCAGTATCTATCACTTGAAGTCCGATTCAGATCT  
380> R M G F S F K D I K K T M E E K I Q T S G S S Y L S L E V L I A D L

**ScaI (1864)** **KasI (1883)**  
1801 TGTGAGTGCTCAGAAAGATAATACGGAGGATGAGTCAAGTCAAATTCATTGCAGAAAGACATTAGTACTGAAGAGCAGCTAAGGCGCCTACAAGAGGAG  
413> V S A Q K D N T E D E S S Q T S L Q K D I S T E E Q L R R L Q E E

**HindIII (1900)** **PshAI (1993)**  
1901 AAGCTTTGCAAAATCTGTATGGATAGAATATTGCTATCGTTTTTGTCTTGTGGACATCTGGTCACTTGTAAACAGTGTGCAGAAGCAGTTGACAAAT  
447> K L C K I C M D R N I A I V F V P C G H L V T C K Q C A E A V D K

**NheI (2063)**  
2001 GTCCCATGTGTACACCGTCATTACGTTCAAGCAAAAAATTTTATGTCTTAGTGGGGCACCAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGT  
480> C P M C Y T V I T F K Q K I F M S •

**BstXI (2060)** **MscI (2069)**  
2101 TTGGACAAACCACAACCTAGAATGCAAGTGAAGAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACA

**HpaI (2201)** **EcoRI (2297)**  
2201 AGTTAAACAACAACAAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAA  
2301 TTCTAAATACAGCATAGCAAACTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGC  
2401 CAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAAGTCTTCTTCTTTATGTTT

**Swal (2550)**  
2501 TAAATGCACTGACCTCCACATTCCCTTTTGTAAATATTAGAAATATTGAAATATTGAAATATTGAAATATTGAAATATTGAAATATTGAAATATTGAAAT

**EcoO109I (2611)**  
2601 CCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTTAATAGAAATTTGACAGCAAGAAGCGAGCTT

2701 CTAGCTTTAGTTCCTGGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCA  
—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

2801 TAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGT  
109 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 V A V I T D

2901 CAAAGTCCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGAT  
76 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 H V A S I

3001 GATCTCCCAAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTGTTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGA  
43 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 E V L E L

3101 TCCTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAG  
9 D Q Q S I N F T K M

3201 CGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGG

3300 GCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAG

3399 TCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCC

3499 ATAAGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGCCATATGATACACTTGATGTAAGTGGCAA

3599 GTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGG

3699 GGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCCTGCAGGTTAA TTAAGAACATGTGAGCAAAGGCCAGCAAAGGCC

3797 AGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC

3897 CGACAGGACTATAAAGATACCAGGCTTTCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTTCT

3997 CCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC

4097 GTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA

4197 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCT

4297 GAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAACAACCCAGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACG

4397 CGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTGGTCAATGGCTA

4497 GTTAATTAACATTTAAATC AGCGGCCGCAATAAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACATAACGCTCTC

4597 CATCAAAAACAAACGAAACAAACAAACTAGCAAAATAGGCTGTCCCAAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA