



**PvuI (7)**  
**SgfI (6)**  
1 GGATCTGCGATCGCTCCGGTGCCCGTCACTGGGCGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGTCCGCAATTGAACGGGTGCCTA

101 GAGAAGGTGGCGGGGTAACCTGGGAAAGTGTCTGTACTGGCTCCGCCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCCGGTTGAGTCGCGTCTGCGCCCTCCCGCTGTGGTGCCCTCTGAAGTCCGCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

**NgoMI (441)**  
**NaeI (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTGGAGGCTACCTAGACTCAGCCGGCTCTCCAGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTTCGTTT

**NcoI (560)**  
**BstEII (555)**  
**AgeI (552)**  
501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGGTGAATGAATACAAGAGAATTGTTCTGCTGAGAGG

1 M V N E Y K R I V L L R G

**BsaBI (655)**  
**MscI (647)**  
601 ACTTGAATGATCAATAAGCATTATTTAGCTTATTTAAGTCATTGCTGGCCAGAGATTTAAATCTGAAAAGAGACAACCAAGAGCAATACACCAGCATT

13 L E C I N K H Y F S L F K S L L A R D L N L E R D N Q E Q Y T T I

701 CAGATTGCTAACATGATGGAAGAGAAATTTCCAGCTGATTCTGGATTGGGCAAACTGATTGATTTTGTGAAGAAGTACCAGCTCTTAGAAAAAGAGCTG

47 Q I A N M M E E K F P A D S G L G K L I E F C E E V P A L R K R A

**EcoO109I (867)** **PciI (899)** **BspLU11I (899)**  
801 AAATTCCTAAAAAGAGAGATCAGAAGTAACAGGAGAAACATCACTGGAAAAAATGGTCAAGAAGCAGGTCTGCAACACCTACATCAACTACAAGCCA

80 E I L K K E R S E V T G E T S L E K N G Q E A G P A T P T S T T S H

901 CATGTTAGCATCTGAAAAGAGGCGAGACTTCTGCAACCCAGGAAGAGACTTCCACAGCTCAGCGGGGACTTCCACAGCTCAGCGGAGGACTTCCACAGCT

113 M L A S E R G E T S A T Q E E T S T A Q A G T S T A Q A R T S T A

1001 CAGGCGAGGACTTCCACAGCTCAGGAGGGGACTTCCACAGCCAGAAAAGAAAAAGTATGAGAGAAGAAGAGACTGGAGTGAAAAGAGCAAGGCGGCTA

147 Q A R T S T A Q E G T S T A Q K R K S M R E E E T G V K K S K A A

1101 AGGAACCAGATCAGCCTCCCTGTTGTGAAGAACCACAGCCAIGTGCCAGTACCAATCTCCACAGCTCATCTTCCGCTTCATCTAACATTCCTTCGGC

180 K E P D Q P P C C E E P T A M C Q S P I L H S S S S A S S N I P S A

**SspI (1231)**  
1201 TAAGAACCAAAAAATCACAACCCAGAAATATTTCCAGAGGTGCTGTTCTCCACTCAGAGCCCTGACAGTGTGGTCTCACTGCAACAGACCCCA

213 K N Q K S Q P Q N I P R G A V L H S E P L T V M V L T A T D P

1301 TTTGAATATGAATCACCAGAACATGAAGTAAAGAACATGCTTCATGCTACAGTGGCTACAGTACAGTATTTCCATGTGAAAGTTTCAACATCAACT

247 F E Y E S P E H E V K N M L H A T V A T V S Q Y F H V K V F N I N

1401 TGAAAAGAAAGTTTCAAAAAAGAAATTTTATCATCATATCCAACTACTTTGAGAGCAAAGGCATCTGGAGATCAATGAGACTTCTCTGTGTGAGAGGC

280 L K E K F T K K N F I I I S N Y F E S K G I L E I N E T S S V L E A

**BglIII (1562)**  
1501 TGCTCTGACCAAAATGATTGAAGTGCCCAACAGTATTATCAGAAATGCAATGCAAGTCCCTAAGATCTGTGATTTCAAAGGGTACTTCTGGAGCAGTG

313 A P D Q M I E V P N S I I R N A N A S P K I C D I Q K G T S G A V

1601 TTCTATGGAGTGTTCATATTACACAAGAAAACAGTGAACCGAAAAGAACACAATCTATGAAATAAAAGATGGTTCCAGGAAGCATAGAAGTGGTGGGGAGTG

347 F Y G V F T L H K K T V N R K N T I Y E I K D G S G S I E V V G S

1701 GAAAATGGCACAAACATCAACTGCAAGGAAGGAGATAAACTCCACTCTTCTGCTTTACCTGAAAACAATTGACAGGCAACCAAGTTAGTGTGTGGAGA

380 G K W H N I N C K E G D K L H L F C F H L K T I D R Q P K L V C G E

**EcoRV (1815)** **BsaBI m (1861)** **XcmI (1877)**  
1801 ACACAGTTTCATCAAGATATCAAAGAGAGGAAATGTACCAAAGGAGCCTGTAAGGAAGAAGATCACCATCATGGTCCCAAACAAGTGGTGTCTGAAA

413 H S F I K I S K R G N V P K E P A K E E D H H H G P K Q V M V L K

1901 GTAACAGAACCATTACATATGACCTGAAAGAGGATAAAAGAATGTTTCATGCTACCGTGGCTACTGAAACTGAGTTCCTCAGAGTGAAGTGTGTTTGGACA

447 V T E P F T Y D L K E D K R M F H A T V A T E T E F F R V K V F D

2001 CAGCTCTAAAGAGCAAGTTCATCCCAAGAAATATCATTGCCATATCAGATATTTTGGGTGCAATGGGTTTCTGGAGATATACAGAGCTTCTGTGTCTC

480 T A L K S K F I P R N I A I S D Y F G C N G F L E I Y R A S C V S

**Psp1406I (2107)**  
2101 TGATGTGAACGTTAATCCAACAATGGTATCTCAAATCACTGAGACAAGAGCTAATGCAACTCCTAAAATTTCTTATCTTTTCTACAAGCAAGGGGG

513 D V N V N P T M V I S N T L R Q R A N A T P K I S Y L F S Q A R G

**ScaI (2216)**  
2201 ACATTTGTGAGTGGAGACTTAGTAAATAAGAAAACGGAGAGGAATAAATTCATTACTATGGAATTGGAGATGATACAGGGAAAATGGAAGTGGTGG

547 T F V S G E Y L V N K K T E R N K F I Y Y G I G D D T G K M E V V

**BpuAI (2306)** **BbsI (2306)**  
2301 TTTATGGAAGACTCAACATGTCAAGTGTGAACAGGCGAGTAAACTAAGACTTGTCTGCTTTGAAATGACTTCCACTGAAGATGGGTGGCAGCTGAGGTC

580 V Y G R L T N V R C E P G S K L R L V C F E L T S T E D G W Q L R S

**BamHI (2440)**  
2401 TGTAAGGCACAGTTACATGCAGGTCATCAATGTAGAAAAGGATCCTATCCCTATGATGTGCCAGACTATGCTGGCTATCCATATGATGTTCTGATTAT

613 V R H S Y M Q V I N A R K G S Y P Y D V P D Y A G Y P Y D V P D Y

**MscI (2543)**  
**NheI (2537)**  
2501 GCTGGATACCCCTTATGATGTGCCAGACTATGCCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTGGACAAAACCAACTAGAATGCAG

647 A G Y P Y D V P D Y A •

**HpaI (2675)**  
2601 TGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGTGAATAAACAAGTTAAACAACAATTCATTCAAT

**EcoRI (2771)**  
2701 TTATGTTTCAGGTTCCAGGGGAGGTGTGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTTAAAATACAGCATAGCAAACCTT

2801 TAACCTCCAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCT

2901 CACCTTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGTTTGAAGCTAGCTCTTCATTCTTTATGTTTTAAATGCACTGACCTCCACATTCCC

3001 TTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATA

3101 TCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGA  
 141 ◀ • N R T Y K L

3201 GGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACAT  
 134 ◀ P I L E E I T T 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

3301 GCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTCTGCCCCGTTGCTCACA  
 101 ◀ 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 F D K Q G N S V

3401 GCAGACCAATGGCAATGGCTTCAGCAGACAGTAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGG  
 67 ◀ 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 I E G T K T R I A

3501 CCGCCCCGACATGGTCTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCGACTCCACCAGTCCAGATCTGCTGAGAGATGTTGAAGGCTTT  
 34 ◀ A G V H H K N D 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

3601 CATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTG  
 1 ◀ M

3701 ACGGTTCACTAAACGAGCTCTGTTATATAGACCTCCACCCTACACGCCTACCGCCATTTGCGTCAATGGGGGGAGTTGTTACGACATTTTGGAAAAG

3801 TCCCGTTGATTACTAGTCAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGT

3901 ACTGCCAAAACCGCATCATGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTAAGTGGCATAATGCCA

4001 GCGGGCCATTACCGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGCAGTTTACCGTAAATACTCCACC

4101 CATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGC

4201 CATTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCG  
 4224 ◀

4301 TTTTCCATAGGCTCCGCCCCCTGACGAGCATCAGAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACCAGCGTTTC

4401 CCCCTGGAAGCTCCCTGTCGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCTTTTCCCTTCGGGAAGCGTGGCGTTTCTCATAG

4501 CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCAGAACCCCGTTCAGCCCAGCCGCTGCGCTTATCCGGT  
 4555 ◀

4601 AACTATCGTCTTGTAGTCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTA

4701 CAGAGTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGG

4801 TAGCTCTTGATCCGGCAAACAAACCCGCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCT

4901 TTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGCCGCGAA  
 4991 ◀

5001 TAAAAATCTTTATTTTATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTA

5101 GCAAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA