



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

101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCCGGTTGAGTCGCGTTTCTGCCGCCTCCCGCTGTGGTGCCTCCTGAAGTCGCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCTCAACTCTACGCTTTTGTTCGTTT

**NcoI (560)**  
**BstEII (555)**  
**AgeI (552)** 501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATTGGCGCAGTTTGCAGTGAATACCAGCGCCTAGAGGC  
1 M A Q F D T E Y Q R L E A

601 CTCCTATAGTGATTCACCCAGGGGAGGAGCCTGTTGGTGACGTCGCCGAGGGGAGCAAGTCACCTTGGCACCATTGAAACCTTGACCTCTTC  
13 S Y S D S P P G E E D L L V H V A E G S K S P W H H I E N L D L F

**XhoI (704)** **BglII (758)**  
701 TTCTCTGAGTTTATAATCTGCACCAGAAGAATGGCTTACATGTATGCTCATCGGGGAGATCTTTGAGCTCATGCAGTTCCTCTTTGTGGTTGCCTTCA  
47 F S R V Y N L H Q K N G F T C M L I G E I F E L M Q F L F V V A F

801 CTACCTTCTGGTCACTGCGTGGACTATGACATCCTATTTGCCAACAAGATGGTGAACCACAGTCTTACCCTACTGAACCCGTCAAGGTCACCTTGC  
80 T T F L V S C V D Y D I L F A N K M V N H S L H P T E P V K V T L P

**BamHI (989)**  
901 AGACGCCTTTTGTCTGCTCAAGTCTGTAGTGCAGGATTGAGAAAATGGCTCCTTATCACCATCCTGGTCATTGCTGGTGTCTTGGATCCACCGG  
113 D A F L P A Q V C S A R I Q E N G S L I T I L V I A G V F W I H R

**FspI (1064)** **BbrPI (1095)**  
1001 CTTATCAAGTTCATCTATAACATTTGCTGCTACTGGGAGATCCACTCCTTCTACCTGCAGCCTCTGCGCATCCCTATGCTGCCCTTCCGTATTGACGT  
147 L I K F I Y N I C C Y W E I H S F Y L H A L R I P M S A L P Y C T

**BsaBI (1146)** **BglII (1145)**  
1101 GGCAAGAAGTGCAGGCCCGGATCGTGACAGCGCAGAAGGAGCACCATGATCTGCATCCACAACCTGAGCTGACAGAAGTGCAGTCTACACCGCATCCT  
180 W Q E V Q A R I V Q T Q K E H Q I C I H K R E L T E L D I Y H R I L

**HpaI (1228)**  
1201 CCGTTTCCAGAACTACATGGTGGCACTGGTTAAACAAATCCCTCCTGCCTCTGCGCTTCCGCTGCCTGGCCTCGGGGAAGCTGTCTTCTTACCCGTTG  
213 R F Q N Y M V A L V N K S L L P L R F R L P G L G E A V F F T R G

1301 CTCAAGTACAACCTTTGAGCTGATCCTTCTTGGGGACCTGGCTCTCTGTTTCTCAATGAATGGAGCCTCAAGGCCGAGTACAACCTGGGGGGCAACGGC  
247 L K Y N F E L I L F W G P G S L F L N E W S L K A E Y K R G G Q R

**BsaBI (1437)** **NdeI (1476)**  
1401 TAGAGCTGGCCAGCGCCTCAGCAACCGCATCCTGTGGATTGGCATCGCTAATCTCTGCTGTGCCCCCTCATCTCATATGGCAAATCCTCTATGCCTT  
280 L E L A Q R L S N R I L W I G I A N F L L C P L I L I W Q I L Y A F

1501 CTTGAGCTATGCTGAGGTGCTGAAGCGGAGCCGGGGCCCTGGGAGCAGCTGCTGGTCACTCTATGGCCGCTGCTACCTCCGCCACTTCAACGAGCTG  
313 F S Y A E V L K R E P G A L G A R C W S L Y G R C Y L R H F N E L

**MscI (1688)**  
1601 GAGCAGAGCTGCAGTCCCGCTCAACCGTGGCTACAAGCCCGCTCCAAGTACATGAATTGCTTCTTGTACCTCTTTGACACTGCTGGCAAGAATG  
347 E H E L Q S R L N R G Y K P A S K Y M N C F L S P L L T L L A K N

**PshAI (1785)**  
1701 GAGCCTTCTTCGCTGGCTCCATCCTGGCTGTGCTTATTGCCCTCACCATTATGACGAAGATGTGTTGGCTGTGGAACATGTGCTGACCACCGTCACACT  
380 G A F F A G S I L A V L I A L T I Y D E D V L A V E H V L T T V T L

**BstEII (1806)** **DraIII (1846)**  
1801 CCTGGGGTACCCTGACCGTGTGCAGTCTTTATCCCGGACCGACATGGTGTCTGCCCTGAGCAGCTGCTCCGCGTGATCCTCGCTCACATCCAC  
413 L G V T V T V C R S F I P D Q H M V F C P E Q L L R V I L A H I H

**BsrBI (1932)** **SapI (1962)**  
1901 TACATGCCTGACCACTGGCAGGTAATGCCACCGCTCGCAGACCCGGGACGAGTTTGGCCAGCTTCCAGTACAAGGCAGTGTTCATTTTGGAGAGT  
447 Y M P D H W Q G N A H R S Q T R D E F A Q L F Q Y K A V F I L E E

2001 TGCTGAGCCCATTTGTCACACCCCTCATCTCTTCTGCTGCGCCACGGCCCTGGAGATTATAGACTTCTTCCGAAACTTACCGTGGAGGTCGT  
480 L L S P I V T P L I L I F C L R P R A L E I I D F F R N F T V E V V

2101 TGGTGTGGGAGATACCTGCTCTTTGCTCAGATGGATGTTGCGCAGCATGGTCACTCCCAAGTGGCTATCTGCTGGGAGAGAGGCTCAGTGTACCA  
513 G V G D T C S F A Q M D V R Q H G H P Q W L S A G Q T E A S V Y Q

**BbrPI (2270)** **DraIII (2267)** **AvrII (2289)**  
2201 CAAGCTGAGGATGAAAGACAGAGTTTGTCACTCATGCACTTTGCCATACCAACCCTGGCTGGCAGCCACCAGTGGAGGACAGCCTTCTAGGCTTCC  
547 Q A E D G K T E L S L M H F A I T N P G W Q P P R E S T A F L G F

**NheI (2332)**  
2301 TCAAGGAGCAGGTTACGCGGGATGGAGCAGCTGCTAGCCTCGCCCAAGGGGTCTGCTCCCTGAAAATGCCCTTTTACGCTATCCAGTCTTACAATC  
580 L K E Q V Q R D G A A A S L A Q G G L L P E N A L F T S I Q S L Q S

**SdaI (2475)**  
2401 TGAGTCTGAGCCCTGAGCCTTATCGAAATGTGGTAGCTGGCTCATCTGCCGGGGCCCTCCACTGCCAGAGACCTGACGGCTCCAGGCACAGGGCT  
613 E S E P L S L I A N V V A G S S C R G P P L P R D L Q G S R H R A

2501 GAAGTCGCTCTGCCCTGCGCTCCTTCTCCCGCTGCAACCCGGGAGGGCCACAGCCGGGCTCACAGCACCATGACAGGCTCTGGGGTGGATGCCA  
647 E V A S A L R S F S P L Q P G Q A P T G R A H S T M T G S G V D A

2601 GGACAGCCAGCTCCGGGAGCAGCGTGTGGGAAGGACAGCTGCAGAGCCTGGTGTGTGAGAATATGCATCCACAGAGATGAGCCTGCATGCCCTCTATAT  
680▶ R T A S S G S S V W E G Q L Q S L V L S E Y A S T E M S L H A L Y M

2701 GCACCAGCTCCACAAGCAGCAGGCCAGGCTGAACCTGAGCGGCATGTATGGCACCGCCGGGAGAGTGATGAGAGTGGAGAAAGCCCCCTGATGAAGGG  
713▶ H Q L H K Q Q A Q A E P E R H V W H R R E S D E S G E S A P D E G

2801 GGAGAGGGCGCCCGGCCCCAGTCTATCCCTCGCTCTGCTAGCTATCCCTGTGCAGCACCCCGGCCTGGAGCTCCTGAGACCACCGCCCTGCATGGGG  
747▶ G E G A R A P Q S I P R S A S Y P C A A P R P G A P E T T A L H G

2901 GCTTCAGAGCGCTACGGTGGCATCACAGATCCTGGCACAGTGCACAGGTTCCCTCTCATTTCTCTCGGCTCCTGGAGGGTGGGCAGAAAGATGG  
780▶ G F Q R R Y G G I T D P G T V P R V P S H F S R L P L G G W A E D G

3001 GCAGTCGGCATCAAGGCACCTGAGCCCGTGCCTGAGAGGGCTCGGAGGATGAGCTACCCCTCAGGTGCACAAGGTATAGACAAGGCTGAGCAGGGTCT  
813▶ Q S A S R H P E P V P E E G S E D E L P P Q V H K V •

3101 TAGACTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAACCACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATG

3201 CTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAACATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGT

3301 TTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCAAAAACACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTT

3401 TCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTAT

3501 TTTCCAAGGTTTGAAGTACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTATAGTAAAATATTCAGAAAATAATTTAAATAC

3601 ATCATTGCAATGAAAATAAATGTTTTTTATTAGCAGAATCCAGATGCTCAAGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAA

3701 GGAACCTTAATAGAAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAG  
141▶ • N R T Y K L P I L E E I T T K V L

3801 CTTGCCATTCATCTCAATGAGCACAAGCAGTCAGGAGCATAGTGCAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCC  
123▶ 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 R D

3901 ACCTCATCAGAGTAGGGGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCAGACAG  
89▶ 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 V T

4001 TGACCCTGCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCAGTCTTGGTCTGATGGCCGCCCCGACATGGTGTCTGTTGCTCATAGAG  
56▶ V R G I Y A E I H V A S I E G T K T R I A A G V H H K N D E Y L

4101 CATGGTATCTTCTCAGTGGCGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCTGATTATAC  
23▶ M T I K E T A V E V L E L D Q Q S I N F T K M

4201 TATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCATAACGAGCTCTGCTTATATAGACC

4301 TCCCACGTACACGCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCAT

4401 TGACGTCAATGGGTTGGAGACTTGAAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGAT

4501 GACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGTTCATGTACTGGCATAATGCCAGGCGGGCCATTTACCGTATTGACGTCAATAGGG

4601 GGGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCGATTTACCGTAAATACTCCACCCATTGACGTCAATGAAAAGTCCCTATTGGCGTTA

4701 CTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGACAGTTA

4801 ATTAAGAACATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATC

4901 ACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAACCCGACAGGACTATAAAGATACGAGCGTTTTCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCC

5001 GACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGAGGTATCTCAGTTCGGTGTAGGTC

5101 GTTCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGTGCCTTATCCGGTAACTATCGTCTTGAGTCAACCCGGTAAGACAG

5201 ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTA

5301 CACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGT

5401 AGCGGTGGTTTTTTTGGTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGGA

5501 ACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGCGCGCAATAAAAATATCTTTATTTTATTACATCTGTGTGTT

5601 GGTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAAAACAAACGAAACAAACAAACTAGCAAAAATAGGCTGTCCCAGTGAAGTGCAGGTG

5701 CCAGAACATTTCTCTATCGAA