



```
      PvuI (7)
      SgfI (6)
1  GGATCTGCATCGCTCCGGTGCCTGTCAGTGGGCAGAGCGACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101  GAGAAAGTGGCGCGGGTAACTGGAAAGTGATGTCGTACTGGTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
      Psp1406I (203)                                HindIII (245)                                Bsu36I (291)
201  GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCGCGCCCTACCTGAGGCC
301  GCCATCCACGCGGGTTGAGTCGCGTTTCTGCGCCCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCGCCGTCTAGGTAAGTTTAAAGTCAAGTCAGGTCGAGACC
401  GGGCCTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCCTGCTTACTCACTTACGTCTTTGTTTCGTTT
      SphI (560)
      KasI (535)              AgeI (552)              KasI (570)
501  TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCCTACCTGAGATCACGGGTGACATGCCGAAGGCGCCCAAGCAGCAGCCGCGGAGCCCGA
113  M P K A P K Q Q P P E P E
      Psp1406I (678)SapI (688)
601  GTGGATCGGGACGGAGAGACACGAGCCATCAGACAAAGTGGTGAAGAAAAGGAAGAAGGACAGAAGAATCAAAAAACGTTCTTTGAAGAGCTGGCA
131  W I G D G E S T S P S D K V V K K G K K D K K I K K T F F E E L A
701  GTAGAAGATAAACAGGCTGGGGAAGAAGAGAAAGTCAAGGAGAAGAGCAGCAGCAGCAGCAACAGCAACAGCAGCAAAAAAAGCGAGATACCC
147  V E D K Q A G E E E K V L K E Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q
801  GAAAAGCAGGCGGAAGAAGGATGTGGATGATGATGGAGAAGAGAAAGAGCTCATGGAGCTCTTAAAGACTCTCAGTGCCACCAAGTGTGAGGAGGA
180  R K G R R K K D V D D D G E E K E L M E R L K K L S V P T S D E E D
      SacII (921)
901  TGAAGTACCCGCCAAAAACCCCGGAGGGAAGAAAACCAAGGTGGTAATGTTTTTGACGCTGATTGAGATCAGAGTGAAGGAGGAGGAGGAA
113  E V P A P K P R G G K K T K G G N V F A A L I Q D Q S E E E E E E
1001  GAAAAATCCTCTAAGCCTGCCAAGCGGAGAAGAATCGGTCAATAAGGCCGTATCTGAGGAACAGCAGCCTGCACTCAAGGGCAAAAAGGGAAAGG
147  E K H P P K P A K P E K N R I N K A V S E E Q Q P A L K G K K G K
1101  AAGAGAAGTCAAAAGGGAAGGCTAAGCCTCAAAATAAATTGCTGCTCTGACAATGAAGAGGAGGATAAAGAAGAAGAAATTAAGGAAAAGGAGCC
180  E E K S K G K A K P Q N K F A A L D N E E E E D K E E E I I K E K E P
1201  TCCCAACAAAGGGAAGGAGAAGGCAAGGCAAGGAGGCGAGGTTTACAGGGAAGAAGGAGAAGGGAAGGAGAAGGAGGAGGAGGAGGAGTCTAAGGCA
213  P K Q G K E K A K K A E Q G S E E E G E G E E E E E E G G E S K A
      BstAPI (1387)
1301  GATGATCCCTATGCTCATCTTAGCAAAAAGGAGAAGAAAAGCTGAAAAACAGATGGAGTATGAGCGCCAAGTGGCTTCATTAAGCAGCCAAATGCAG
247  D D P Y A H L S K K E K K K L K K Q M E Y E R Q V A S L K A N A
      NsiI (1460)                                BsrBI (1496)
1401  CTGAAAATGACTTCTCCGTGTCACAGGCGGAGATGTCCTCCCGCAAGCCATGTTAGAAAATGCATCTGACATCAAGCTGGAGAAGTTCAGCATCTCCGC
280  A E N D F S V S Q A E M S S R Q A M L E N A S D I K L E K F S I S A
      BsrGI (1532)              EagI (1545)
1501  TCATGGCAAGGAGCTGTTCTGATACATTGTAGCCCGCCCTGCTACGGGCTGGTAGACCCAATGGCAAGGGCAAGACCACACTCCTCCT
313  H G K E L F V N A D L Y I V A G R R Y G L V G P N G K G K T T L L
1601  AAGCACATTGCCAACGAGCCCTGAGCATCCCTCCCAACATTGATGTTGCTGTGTGAGCAGGAGGTTAGCAGATGAGACACCAGCAGTCCAGGCTG
347  K H I A N R A L S I P P N I D V L L C E Q E V V A D E T P A V Q A
      BsrBI (1741)
1701  TTCTTCGAGCTGACACCAAGCGATTGAAGCTGTGGAAGAGGAGCGGCTTCAGGGACAGCTGGAACAAGGGGATGACACAGCTGTGAGAGGCTAGA
380  V L R A D T K R L K L L E E E R R L Q G Q L E Q G D D T A A E R L E
      BamHI (1859)
1801  GAAGGTGATGAGGAATTGCGGGCCACTGGGGCGCAGCTGCAGAGGCCAAAGCAGCCGCGATCTGGCTGGCTGGCTTTGACCTGAAATGCAGAAT
413  K V Y E E L R A T G A A A A E A K A R R I L A G L G F D P E M Q N
      SphI (1931)              MscI (1946)                                XcmI (1994)
1901  CGACCCACACAGAAGTTCTCAGGGGGCTGGCGCATGCGTGTCTCCCTGCCCAGGGCAGCTGTTATGAGCCACACTGCTGATGCTGGATGAGCCACCA
447  R P T Q K F S G G W R M R V S L A R A L F M E P T L L M L D E P T
2001  ACCACCTGGACTCAACGCTGTATCTGGCTTAAATAACTACCTCCAGGCTGGCGGAAGACCTTGTGATCGTCTCCATGACCCAGGGCTTCTTGATGA
480  N H L D L N A V I W L N N Y L Q G W R K T L L I V S H D Q G F L D D
      EcoRV (2110)
2101  TGTCTGCACTGATATCATCCACCTGATGCCAGCGGCTCCACTACTATAGGGCAATTACATGACCTTCAAAAAGATGTACCAGCAGAAGCAGAAAAGAA
513  V C T D I I H L D A Q R L H Y Y R G N Y M T F K K M Y Q Q K Q K E
2201  CTGCTGAAACAGATGAGAAGCAAGAGAAAAGCTGAAGGAGCTGAAGCAGGCGGGAAGTCCACCAAGCAGGGGAAAAACAACGAAGGAAGCCCTGA
547  L L K Q Y E K Q E K K L K E L K A G G K S T K Q A E K Q T K E A L
      Bsu36I (2374)
2301  CTCGGAAGCAGCAGAAAATGCCAGCGAAACCAAGATGAGGAATCCAGGAGGCCCTGAGCTCCTGAAGCGCCCTAAGGAGTACACTGTGCGCTTCCAC
580  T R K Q Q K C R R K N Q D E E S Q E A P E L L K R P K E Y T V R F T
2401  TTTTCCAGACCCCACTCAGCCCTCAGTGTGTTGTCATGTTGTCATTGCCTACCAGGAGCAGAAACCACTTTAAGAACTTGATTTT
613  F P D P P P L S P P V L G L H G V T F G Y Q G Q K P L F K N L D F
      Bsp120I (2532)                                NcoI (2595)
2501  GGCATCGACATGGATTCAAGGATTTGCATTGTGGGCCAATGGTGTGGGGAAGAGTACGCTACTCTGTGCTGACTGGCAAGCTGACACCCAGCCATG
647  G I D M D S R I C I V G P N G V G K S T L L L L T G K L T P T H
      FspI (2663)
2601  GGGAAATGAGAAAGAACCCGCGTAAAATTGGCTTCTTCAACCAGCAGTATGCAGAGCAGCTGCGCATGGAGGAGACGCCACTGAGTACCTGCAGCG
680  G E M R K N H R L K I G F F N Q Y A E Q L R M E E T P T E Y L Q R
      SfiI (2740)                                BglIII (2777)
2701  GGGCTTCAACCTGCCCTACCAGGATGCCCGCAAGTGCCTGGCGCTTCCGCCTGGAGAGTACGCCCCACCATCCAGATCTGCAAACTCTCTGGTGGT
713  G F N L P Y Q D A R K C L G R F G L E S H A H T I Q I C K L S G G
```

2801 CAGAAGGCGCGAGTTGTGTTTGGCTGAGCTGGCCTGTCGGGAACCTGATGTCCTCATCTTGGACGAGCCAAACCAATAACCTGGACATAGAGTCTATTGATG  
747▶ Q K A R V V F A E L A C R E P D V L I L D E P T N N L D I E S I D  
2901 CTCTAGGGGAGGCATCAATGAATACAAGGGTGTGTGATCGTTGTGACGGATGATGCCCGACTCATCAGAAAACCAATTGCCAGCTGTGGTGGTGGGA  
780▶ A L G E A I N E Y K G A V I V V S H D A R L I T E T N C Q L W V V E

**ClaI (3019)**

3001 GGAGCAGAGTGTAGCCAAATCGATGGTGACTTTGAAGACTACAAGCGGGAGGTGTTGGAGGCCCTGGGTGAAGTCATGGTCAGCCGGCCCGAGAGTGA  
813▶ E Q S V S Q I D G D F E D Y K R E V L E A L G E V M V S R P R E •

**MscI (3131)**

3101 GCTTTCCTCCAGAAAGTCTCCCGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGC

**HpaI (3263)**

3201 TTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAACAAGTTAAACAACAACATTGCATTCAATTTATGTTTCAGG

**EcoRI (3359)**

3301 TTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAAATCTAAATAACAGCATAGCAAACTTTAACCTCCAAAT  
3401 CAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCA

**SapI (3541)** **SspI (3598)**

3501 TGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAAGTCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCCTTTTATAGTAAA

**SwaI (3612)**

3601 TATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTA

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

3801 CCTCAATGGTGGTTTTGACCAGCTTCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCT  
130▶ 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 G C P S

3901 GACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATG  
97▶ 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 A S G I

**StuI (4037)**

4001 GCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCCGACAT  
63▶ 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 A G V H

**XmnI (4179)**

4101 GGTGCTTGTGCTCATAGAGCATGGTATCTTCTCAGTGGCAGCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCT  
30▶ 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 M

**AseI (4245)**

4201 CCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAA  
4301 ACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTCGCTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTT

**SpeI (4400)**

4401 ACTAGTCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACC

**SnaBI (4528)**

4501 GCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGTTCATGTACTGGCATAATGCCAGGCGGGCCATTT

**NdeI (4633)**

4601 ACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCCAAGTGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAA  
4701 TGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGGCGGGGTCGTTGGCGGTGAGCCAGGCGGGCCATTTACCGTAA

**PacI (4819)**

**SdaI (4811)** **BspLU11I (4829)**

4801 GTTATGTAACGCCTGCAGGTTAAATAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAGGCCGCTTGTGGCGTTTTCCATAGG  
4901 CTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCT  
5001 CCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAAGCGTGGCGCTTCTCATAGCTCACGCTGTAG

**ApaI (5143)**

5101 GTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTT  
5201 GAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGA  
5301 AGTGGTGGCCTAACTACGGCTACACTAGAAGAAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATC  
5401 CGGCAAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCT

**EagI (5579)**

**PacI (5559)** **SwaI (5568)** **NotI (5578)**

5501 ACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGCCGCAATAAAAATATCTTT  
5601 ATTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAAATAGGCT  
5701 GTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA