



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

MfeI (82)  
101 GAGAAGGTGGCGCGGGTAAACTGGAAAAGTATGTCGTGTACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

**NgoMIV (441)**  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

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

1 M D P G K P R K N V L V V

**BstEII (642)**  
601 GGCTCTCCTTGTCAATTTCCAGGTGTGCTTCTGCCAAGATGAGGTCACCGATGACTACATCGGCGAGAATACCACGGTGGACTACACCCTGTACGAGTCG

13 A L L V I F Q V C F C Q D E V T D D Y I G E N T T V D Y T L Y E S

701 GTGTGCTCAAGAAGGATGTGCGGAACCTTAAGGCCTGGTCTCGCTCATGTATTCTGTCATCTGCTCGTGGGCTGCTCGGCAACGGGCTGGTGA

47 V C F K K D V R N F K A W F L P L M Y S V I C F V G L L G N G L V

801 TACTGACGTACATCTATTTCAAGAGGCTCAAGACCATGACGGATACCTACCTGCTCAACCTGGCCGTGGCAGACATCTTTTCTCCTGATTCTTCCCTT

80 I L T Y I Y F K R L K T M T D T Y L L N L A V A D I L F L L I L P F

**HindIII (975)**  
901 CTGGCCCTACAGCGAAGCAAGTCTGGATCTTTGGCTCTACCTGTGTAAGGCATCTTTGGCATCTATAAGTTAAGCTTCTCAGCGGGATGCTGCTG

113 W A Y S E A K S W I F G V Y L C K G I F G I Y K L S F F S G M L L

**NsiI (1005)** **SnaBI (1024)**  
1001 CTCCTATGCATCAGCATTGACCGCTACGTAGCCATCGTCCAGGCCGTGTCGGCTCATCGCCACCGCGCCCGGTGCTTCTCATCAGCAAGCTGTCCTGTG

147 L L C I S I D R Y V A I V Q A V S A H R H R A R V L L I S K L S C

1101 TGGGCATCTGGATGCTGGCCCTTCTCCTCCATCCCGGAGCTGCTCTACAGCGCCTCCAGAAGAACAGCGCGAGGACACGCTGAGATGCTCACTGGT

180 V G I W M L A L F L S I P E L L Y S G L Q K N S G E D T L R C S L V

**BstXI (1226)**  
1201 CAGTGCCCAAGTGGAGGCCTTGATCACCATCCAAGTGGCCAGATGGTTTTGGGTTCTAGTGCCTATGCTGGCTATGAGTTTCTGCTACCTCATTATC

213 S A Q V E A L I T I Q V A Q M V F G F L V P M L A M S F C Y L I I

**BsrBI (1330)**  
1301 ATCCGTACCTTGTCCAGGCACGCAACTTTGAGCGGAACAAGGCCATCAAGGTGATCATTGCCGTGGTGGTAGTCTTCATAGTCTTCCAGCTGCCCTACA

247 I R T L L Q A R N F E R N K A I K V I I A V V V V F I V F Q L P Y

**MscI (1424)**  
1401 ATGGGGTGGTCTGGCTCAGACGGTGGCCAACCTTCAACATCACCAATAGCAGCTGCGAAACCAGCAAGCAGCTCAACATTGCCTATGACGTACCTACAG

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

**SapI (1578)**  
1501 CCTGGCCTCCGCTCCGCTGCTGCGTCAACCCCTTCTGTATGCCTTCATCGGCGTCAAGTTCGCGACGACCTTCAAGCTTCAAGGACTTGGGCTGT

313 L A S V R C C V N P F L Y A F I G V K F R S D L F K L F K D L G C

**NgoMIV (1634)** **MluI (1649)**  
1601 CTCAGCCAGGAACGGCTCCGGCACTGGTCTTCTGCCGGCATGTACGGAACCGTCCGGTGAGCATGGAGGCGGAGACCACCACAACCTTCTCCCGTAGG

347 L S Q E R L R H W S S C R H V R N A S V S M E A E T T T T F S P •

**MscI (1790)**  
1701 GGGCTCCCCTGCCGGACTACAAGGACCTCTCCAGGAGCCTAATGTGGTGCACACATGCACAGACTCTACATCCACCGAATTGCTAGCTGGCCAGACA

**EcoO109I (1722)** **ApaLI (1749)** **NheI (1784)**  
1801 TGATAAGATACATTGATGAGTTTGGACAACCACAAGTGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAAAC

**HpaI (1922)** **MfeI (1933)**  
1901 CATTATAAGTGAATAAACAAGTTAAACAACAACAATTGCATTCTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAAAC

**EcoRI (2018)**  
2001 CTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGC

2101 ATAGGCATCAGGGGCTGTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGATTTTCCCAAGGTTTGAACTA

**SapI (2200)** **SspI (2257)** **SwaI (2271)**  
2201 GCTCTTCAATTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATA

**EcoO109I (2332)**  
2301 ATGTTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATT

2401 GGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATG  
141 • N R T Y K L P I L E E I T T K V L K G N M E I  
SacI (2532) BstXI (2561)  
2501 AGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGT  
116 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 V E D S Y P H  
2601 GCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGC  
83 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 V R G I Y A  
2701 CTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTCTTGTTCCTCATAGAGCATGGTATCTTCTCAGTG  
50 E I 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  
XmnI (2838)  
2801 GGCACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCC  
16 A V E V L E L D Q Q S I N F T K M  
AseI (2904) SacI (2961)  
2901 GATGATTAATTGTCAAACACGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTAC  
SpeI (3059)  
3001 CGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAG  
SnaBI (3187)  
3101 ACTTGAAATCCCCTGAGTCAAACCGTATCCACGCCATTGATGTACTGCCAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAC  
NdeI (3292)  
3201 TGCCAAGTAGGAAAGTCCATAAGGTCATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGAT  
3301 ACACCTTGATGACTGCCAAGTGGGCGAGTTTACCCTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCA  
PacI (3478)  
PstI (3471)  
SdaI (3470) BspLU11I (3488)  
3401 TATTGACGTCAATGGGCGGGGTGCTTGGGCGGTGAGCCAGGCGGGCCATTTACCCTAAGTTATGTAACCGCTGCAGGTTAAATAAGAACATGTGAGCAA  
3501 AAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCGCGGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG  
3601 TCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGA  
3701 TACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGCT  
ApaLI (3802)  
3801 GTGTGCACGAACCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGCTAACTATCGTCTTGTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC  
3901 AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTT  
4001 GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTT  
4101 GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGG  
EagI (4238)  
PacI (4218) SwaI (4227) NotI (4237)  
4201 GATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGT  
4301 AACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCG  
4401 AA