



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

101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCCCTGTTGCTCAACTCTACGCTTTTGTTCGTTT

**BstEII (555)** **AgeI (552)** **NcoI (560)** **EagI (585)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATTGGAGCCCGCCGCCCCGGCCCGCCCTCGGGCC

1 M E P A G P A P G R L G P  
PstI (661)

601 GCTGCTGCTGCTGCTGCTCGCCGCTCCTGCGCTGGTCCAGGAGTGGCGGGTGGAGGAGAGCTGCAGGTGATTGAGCCTGACAAGTCCGTGTTGGTTGCA

13 L L C L L L A A S C A W S G V A G E E E L Q V I Q P D K S V L V A  
**Bsp120I (751)** **AseI (792)**

701 GCTGGAGAGACAGCCACTCTGCGCTGCACTGCGACCTCTCTGATCCCTGTGGGCCATCCAGTGGTTCAGAGGAGCTGGACCAGCCGGGAATTAATCT

47 A G E T A T L R C T A T S L I P V G P I Q W F R G A G P G R E L I

**XmaI (824)**  
801 ACAATCAAAAAGAAGGCCACTTCCCGGGTAACAACCTGTTTCAGACCTCACAAAGAGAAACAACATGGACTTTTCCATCCGCATCGGTAACATCACCCC

80 Y N Q K E G H F P R V T T V S D L T K R N N M D F S I R I G N I T P

**BspEI (932)** **BssHII (993)**  
901 AGCAGATGCCGGCACCTACTACTGTGTGAAGTTCCGGAAAGGAGCCCGATGACGTGGAGTTAAGTCTGGAGCAGGCACTGAGCTGTCTGTGCGCGCC

113 A D A G T Y Y C V K F R K G S P D D V E F K S G A G T E L S V R A

**Bsp120I (1023)** **DraIII (1051)**  
1001 AAACCCTTGCCCCGTGGTATCGGGCCCTGCGGCGAGGGCCACACCTCAGCACACAGTGAGCTTACCTGCGAGTCCACGGCTTCTCACCCAGAGACA

147 K P S A P V V S G P A A R A T P Q H T V S F T C E S H G F S P R D

**XmnI (1107)**  
1101 TCACCCTGAAATGGTTCAAAAATGGGAATGAGCTCTCAGACTTCCAGACCAACGTGGACCCCGTAGGAGAGAGCGTGTCTACAGCATCCACAGCACAGC

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

**SmaI (1271)**  
1201 CAAGTGGTGTGACCCGCGAGGACGTTCACTCTCAAGTCATCTGCGAGGTGGCCACGTCACCTTGCAGGGGGACCTCTTCGTGGGACTGCCAACTTG

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

1301 TCTGAGACCATCCGAGTTCACCCACCTTGGAGTTACTCAACAGCCCGTGGAGGAGAGAAACAGGTGAATGTCACCTGCCAGGTGAGGAAGTTCTACC

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

**Acc65I (1483)**  
1401 CCCAGAGACTACAGTGTGCTGGTGGAGAATGGAAACGTGTCCCGACAGAAACGGCCTCAACCGTTACAGAGAACAAGGATGGTACCTACAACCTGGAT

280 P Q R L Q L T W L E N G N V S R T E T A S T V T E N K D G T Y N W M

1501 GAGCTGGCTCCTGGTGAATGTATCTGCCACAGGGATGATGTGAAGCTCACCTGCCAGGTGGAGCATGACGGGCGAGCCAGCGTCCAGAAAAGCCATGAC

313 S W L L V N V S A H R D D V K L T C Q V E H D G Q P A V S K S H D

**ApaLI (1698)**  
1601 CTGAAGGTCTCAGCCACCCGAAGGAGCAGGGCTCAAATACCGCCGCTGAGAACACTGGATCTAATGAACGGAACATCTATATTGTTGGTGGGTGTGGTGT

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

1701 GCACCTTGTGTTGGCCCTACTGATGGCGGCCCTTACCTCGTCCGAATCAGACAGAAAGCCAGGGCTCCACTTCTTACAAGTTGCATGAGCC

380 C T L L V A L L M A A L Y L V R I R Q K K A Q G S T S S T R L H E P

**NdeI (1844)**  
**EcoRV (1837)**  
1801 CGAGAAGATGCCAGAGAAATAACACAGGACACAAATGATATCACATATGCAGACCTGAACCTGCCAAGGGGAAAGCCTGCTCCCGAGGTGCGGGAG

413 E K N A R E I T Q D T N D I T Y A D L N L P K G K K P A P Q A A E

**Tth111I (1981)**  
1901 CCCAACACCACACGGAGTATGCCAGCATTGACACCAGCCCGAGCCCGCTGGAGGACACCCTCACCTATGCTGACCTGGATGGTCCACCTCAACC

447 P N N H T E Y A S I Q T S P Q P A S E D T L T Y A D L D M V H L N

2001 GGACCCCCAAGCAGCCGCCCCCAAGCCTGAGCCGCTTCTCAGAGTACGCCAGCGTCCAGGTCCCGAGGAAGTGAATGGGACCGTGGTTTGTCTTAGC

480 R T P K Q P A P K P E P S F S E Y A S V Q V P R K •

**MscI (2132)** **NheI (2126)**  
2101 ACCCATCTACGCGCTTCTTGTCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATG

**HpaI (2264)** **MfeI (2275)**  
2201 CTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAATTGCATTCATTTTATGTTTCAG

**EcoRI (2360)**  
2301 GTTCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACCTTTAACCTCCAAA

2401 TCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTTGGCAATGTGCATTAGCTGTTTGAGCCTCACCTTCTTTC  
**SapI (2542)** **SspI (2599)**

2501 ATGGAGTTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTTTAAATGACTGACCTCCACATTCCCTTTTATGAAA  
**SwaI (2613)**

2601 ATATTCAGAAATAATTTAAATACATCATTGCAATGAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTT  
**BstXI (2903)**

2701 AGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGATGAGT  
141 • N R T Y K L P I L

2801 TCCTCAATGGTGGTTTTGACCAGCTTGCATTCACTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGC  
130 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 G C P S

2901 TGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGCTTCTGCCGTTGCTCACAGCAGACCAAT  
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

3001 GGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCCGACA  
64 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  
**StuI (3038)** **BbsI (3184)** **XmnI (3180)**

3101 TGGTGCTTGTTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGCTTCATGGTGGCCC  
30 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 M

3201 TCCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTA  
**AseI (3246)**

3301 AACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATT  
**SpeI (3401)**

3401 TACTAGTCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAAC  
**SnaBI (3529)**

3501 CGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGGGGCCATT  
**NdeI (3634)**

3601 TACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCA  
**PacI (3820)** **PstI (3813)** **SdaI (3812)** **BspLU11I (3830)**

3801 AGTTATGTAACGCCTGACAGTTAAATTAAGAACAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTCTGGCGTTTTTCCATAG  
3901 GCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGC  
4001 TCCCTCGTGCCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTA  
**ApaLI (4144)**

4101 GGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCCGTTCCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCT  
4201 TGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTG  
4301 AAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT  
4401 CCGGAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTC  
**EagI (4580)** **PacI (4560)** **SwaI (4569)** **NotI (4579)**

4501 TACGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTT  
4601 TATTTTCATTACATCTGTGTGGTTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGC  
4701 TGTCGCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA