



**PvuI (7)**  
**SgfI (6)** 1 GGATCTGGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
**MfeI (82)**  
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGCTGTACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

**Psp1406I (203)** **HindIII (245)** **Bsu36I (291)**  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTCACTTACGTCTTTGTTTCGTTT

**NcoI (560)**  
**BstEII (555)**  
**KasI (535)** **AgeI (552)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACCATGGAGTCCATGTTCCGAGGACGACATCAGCATCCTGAC  
1 M E S M F E D D I S I L T

**PciI (665)**  
**BspLU111 (665)**  
601 GCAGGAGGCCTGGGGCTAGTGAGGTGTGGCTGGACAGTCCCGGAGACCCCTCGCTTGGGGGAGACATGTGCTCCGCTCCCACTTTGCCCTCATCACT  
13 Q E A L G P S E V W L D S P G D P S L G G D M C S A S H F A L I T  
701 GCTTACGGAGACATCAAGGAACGGCTGGGGGCTGGAGAGGAGAAGCCACCCTCCGACGCGCCTCAAAGTCTACGAGATCAAGTACCCACTGATCA  
47 A Y G D I K E R L G G L E R E N A T L R R R L K V Y E I K Y P L I  
801 GTGACTTTGGAGAGGAGCATGGCTTTTCTGTATGAAATCAAGGATGGCTCCCTGCTGGAGTGGAGAAGGTGAGCCTGCAGCAACGGCTCAACCAGTT  
80 S D F G E E H G F S L Y E I K D G S L L E V E K V S L Q R R L N Q F  
901 CCAGCATGAGTTACAGAAGAACAAGGAGCAGGAAGAACAGCTTGAGAGATGATCCAGGCCTACGAGAAACTTGCCTGGAGAAGAGTACTTGGAGACA  
113 Q H E L Q K N K E Q E E Q L G E M I Q A Y E K L C V E K S D L E T

**Bsp120I (1016)** **BglII (1043)**  
1001 GAGCTGAGGGAGATGAGGGCCCTGGTGGAGACGACCTGCGTCAGATCTGTGGTTTGGAGCAGCAGTGGCGCAACAGCAAGGCCTCCAGGATGCAGCCT  
147 E L R E M R A L V E T H L R Q I C G L E Q Q L R Q Q Q G L Q D A A  
1101 TCTCAACCTGAGCCACCGCCAGCCCCGCCCCCTCCCTGCACTGATTTAGACCTGCACTACCTGGCACTGAGAGGGGATCTGGCTTGGATCATGCAGG  
180 F S N L S P P P A P A P P C T D L D L H Y L A L R G G S G L S H A G

**DraIII (1214)**  
1201 CTGGCCGGGCTCCACACCCAGTGTGAGTGACCTGGAGCGCGCGCTAGAAAGGCTTTGGAGCGCGCAGGGAGAGGCCCGGGGGCTCAGTCCGG  
213 W P G S T P S V S D L E R R R L E E A L E A A Q G E A R G A Q L R

**Bsp120I (1364)**  
1301 GAGGAGCAGCTCCAGGCCGAGTGCAGCGGCTGCAGGGGAGCTGAAGCAGCTGCAGGAGACCCGGGCCAGGATCTGGCCTCCAACAGTCCGGAGCGAG  
247 E E Q L Q A A E C E R L Q G E L K Q L Q E T R A Q D L A S N Q S E R  
1401 ACATGGCGTGGGTGAAAAGAGTTGGGGATGATCAGTGAATTTGGCKCTGGCCTACACGAGTGCAGCGAGGAGTGGCGCGCTTGGGAGTTGAGTTTC  
280 D M A W V K R V G D D Q V N L A L A Y T E L T E E L G R L R E L S S  
1501 CCTACAGGGGAGAATCTTGAAGACTCTGTTGCAGGAACAGGCCGGAGTGGCGCCAGAGGCACTCGCCGCTGTACAACGCCACTCCCCGGCCCCCAG  
313 L Q G R I L R T L L Q E Q A R S G G Q R H S P L S Q R H S P A P Q  
1601 TGCCCCCTCCCCCTCCCCGCTGCCGAGCGGCTCCCCGTGCCCCCGTGCAGTCCCCGTCCCCAGCGCCGCTCTCCCGTCCCCCGTCCCCCTCGC  
347 C P S P S P P A R A A P P C P P C Q S P V P Q R R S P V P P C P S

**SgrAI (1760)**  
1701 CGCAGCAGCGCGCTCTCCGGCCTCACCTCCTGCCCGTCCCGCTCCCGCAGCGCCGCTCGCCGGTCCCGCCTCGTGCAGTCCCCAGCCCGCAGCG  
380 P Q Q R R S P A S P S C P S P V P Q R R S P V P P S C Q S P S P Q R

**BssHII (1884)**  
1801 CCGTCCCCGGTCCCCCAGCTGCCCGCCCCGAGCCCCGCCACCGCCGCCCCCCGCCGGCGAGAGGACGCTGGCCGAGCGGCCTACGCCAAG  
413 R S P V P P S C P A P Q P R P P P P P P G E R T L A E R A Y A K

**Eco72I (1912)**  
**BbrPI (1912)**  
**DraIII (1909)** **KasI (1978)**  
1901 CCGCCAGCCACACGTTGAAGGCCGGCTTCCAGGGCCCGCAGCTACTCTGAGCTGGCGGAGGGCGCGCCTACGCGGGCGCTCCCCGCCCTGGCTGC  
447 P P S H H V K A G F Q G R R S Y S E L A E G A A Y A G A S P P W L

**SfiI (2001)** **BssHII (2079)**  
**AscI (2078)** **SacII (2074)**  
2001 AGGCCGAAGCGGCCACTCTCCCCAAGCCCCGGCCTACGGCAGCGAGCTCTACGGCCCTGGCAGGCCCTCAGCCCGCGCGCCTTCGAGGGCATCCG  
480 Q A E A A T L P K P R A Y G S E L Y G P G R P L S P R R A F E G I R

**XcmI (2159)**  
2101 GCTGCGCTTCCAGAAGCAGCGCTCGGAGGAGGACGAGTGGGCTGTGCCACCAGCCCCGAGCCCGGAGGTGGGCACCATCCGCTGCGCCTCCTTCTGC  
513 L R F E K Q P S E E D E W A V P T S P P S P E V G T I R C A S F C  
2201 GCGGGCTTCCCCATCCCCGAGTCCCGCCGCCACCGCTACGCCACGCCGAGCAGCGCAGTCTGCGCGTCCATCAACTGCTGATGGAGACGGTGG  
547 A G F P I P E S P A A T A Y A H A E H A Q S W P S I N L L M E T V

**Asp718I (2348)**  
**Acc65I (2348)**  
2301 GCTCCGACATCCGAGCTGCCCCCTCTGCCAGCTGGGTTTCCCTGTCCGGTACCCGGATGATGCCTCATCAACACATTGACTCCCACCTGGAGAACAG  
580 G S D I R S C P L C Q L G F P V G Y P D D A L I K H I D S H L E N S

**BglIII (2402)** **MscI (2449)**  
**NheI (2443)**  
2401 CAAGATCTAGGGCACCAGCCCCACCACTGGCTGTTTCTCCGTGCTAGTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAACCAACTAGA  
613 K I •

2501 **HpaI (2581)** MfeI (2592)  
 ATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACCAAGTTAACCAACAACAAATTGCA

2601 **EcoRI (2677)**  
 TTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAAATCTAAAATACAGCATAGCA

2701 AAACCTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTG

2801 CAGCCTCACCTTCTTTCATGAGGTTAAGATATAGTGATTTTTCCCAAGGTTTGAAGTACTCTTCTATTCTTTATGTTTTAAATGACTGACCTCCCAC

2901 **SspI (2916)** SwaI (2930)  
 ATTCCCTTTTTAGTAAAAATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTC

3001 ATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTAAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGT  
 141 • N R T Y

3101 ACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTCCATTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCT  
 136 K L 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

3201 **BstXI (3220)**  
 GCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCCTTCTGCCGTTG  
 103 C M 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

3301 CTCACAGCAGACCAATGGCAATGGCTTCTGACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCC  
 69 S V 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

3401 **XmnI (3497)**  
 TGATGGCCGCCCCGACATGGTGCTTGTGCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAA  
 36 I A 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

3501 **BspHI (3505)** **VspI (3563)**  
**AseI (3563)**  
 GGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGC T  
 3 T K M

3601 TATCTGACGGTTCACTAAACGAGCTCTGTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTT

3701 **SpeI (3718)**  
 GGAAAGTCCCGTTGATTTACTAGTCAAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCCA

3800 **SnaBI (3846)**  
**Eco105I (3846)**  
 TTGATGTAAGTCCAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCCATAGGTCATGTAAGTGGGCAT

3900 **NdeI (3951)**  
 AATGCCAGGGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTAAGTCCCAAGTGGGCAGTTTACCGTAAATA

4000 CTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCAATGGGGGGGGTCTTGGGGGGTCCAGCCA

4100 **SdaI (4129)** **PacI (4137)** **PciI (4147)**  
**BspLU11I (4147)**  
 GCGGGCCATTTACCGTAAGTTATGTAACCGCTG C A G G T T A A T T A A G A A C A T G T G A G C A A A A G G C C A G A A A A G G C C A G A A C C G T A A A A G G C C G C G

4198 TTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAGAAAATCGACGCTCAAGTCAAGGAGGTGGCGAAACCCGACAGGACTATAAAGATACC

4298 AAGCGTTTTCCCTGGAAGTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGGAAAGCGTGGCGCT

4398 **ApaLI (4461)**  
 TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCCAGCCGCTGCGCC

4498 TTATCCGGTAACTATCGTCTTGAAGTGGTGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGTAAGCCAGTTACCTTCGGAAA

4598 GCGGTGCTACAGAGTTCTTGAAGTGGTGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGTAAGCCAGTTACCTTCGGAAA

4698 AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAA

4798 **PacI (4877)** **SwaI (4886)** **NotI (4896)**  
 GAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AG

4898 **EagI (4897)**  
 CGGCCGAATAAAATATCTTTATTTTTCATTACATCTGTGTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAA

4998 AACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA