



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

101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGCTGTACTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTAGGTAAGTTAAAGTCAAGTGCAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGTTGCTCAACTCTACGCTTTTGTTCGTTT

**NcoI (560)**  
**BstEII (555)**  
**AgeI (552)**  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGTACCATGGTGAACCTCACGGTAGACCAGATCCGCGCCATCAT  
1 M V N F T V D Q I R A I M

**BspLU1I (623)** **BbrPI (640)** **NcoI (648)**  
601 GGACAAGAAGGCCAACATCCGCAACATGCTGTGCATCGCCACGTGGACCATGGCAAGTCCACGCTGACAGACTCCCTGGTGTCAAGGCGGGCATCATC  
13 D K K A N I R N M S V I A H V D H G K S T L T D S L V C K A G I I

**XmaI (707)**  
**SrfI (706)**  
701 GCCTCGGCCCGGGCCGGGAGACACGCTTCACTGATACCCGGAAGGACGAGCAGGAGCGTTGCATCACCATCAAGTCAACTGCCATCTCCCTCTTCTACG  
47 A S A R A G E T R F T D T R K D E Q E R C I T I K S T A I S L F Y

**XmaI (877)** **Sall (886)**  
801 AGCTCTCGGAGAATGACTTGAACCTCATCAAGCAGAGCAAGGACGGTCCCGGCTTCTCATCAACCTCATTGACTCCCCGGGCATGTCGACTTCTCCTC  
80 E L S E N D L N F I K Q S K D G A G F L I N L I D S P G H V D F S S

901 GGAGGTGACTGCTGCCCTCCGAGTACCAGTGGCGCATTGGTGGTGGTGGACTGCGTGTGAGCGTGTGCGTGACAGCAGGAGACAGTGTGCGGCAGGCC  
113 E V T A A L R V T D G A L V V V D C V S G V C V Q T E T V L R Q A

**PstI (1060)**  
1001 ATTGCCGAGCGCATCAAGCTGTGCTGATGATGAACAAGTGGACCGCGCCCTGCTGGAGCTGCAGCTGGAGCCCGAGGAGCTTACCAGACTTCCAGC  
147 I A E R I K P V L M M N K M D R A L L E L Q L E P E E L Y Q T F Q

**NcoI (1157)** **Clal (1174)**  
**BsrBI (1149)** **BspHI [m] (1169)** **Acc65I (1189)**  
1101 GCATCGTGGAGAAGCTGAACGTCATCATCTCCACCTACGGGAGGGGAGAGCGGCCCATGGCAACATCATGATCGATCCTGTCTCGGTACCGTGGG  
180 R I V E N V N V I I S T Y G E G E S G P M G N I M I D P V L G T V G

1201 CTTTGGGTCTGGCCTCCACGGTGGGCCTTCCACCTGAAGCAGTTTGGCCGAGATGATGTGGCCAAGTTCCGCCCAAGGGGGAGGGCCAGTTGGGGCCT  
213 F G S G L H G W A F T L K Q F A E M Y V A K F A A K G E G Q L G P

**BsrBI (1303)**  
1301 GCCGAGCGGGCAAGAAAGTAGAGGACATGATGAAGAAGCTGTGGGTGACAGGTACTTTGACCCAGCCAACGGCAAGTTCAGCAAGTCAAGCCACGACC  
247 A E R A K K V E D M M K K L W G D R Y F D P A N G K F S K S A T S

**BspHI [m] (1470)**  
1401 CCGAAGGGAAGAAGCTGCCACGCACCTTCTGCCAGCTGATCCTGGACCCCATCTCAAGGTGTTTGTGCGATCATGAATTTCAAGAAAGAGGAGACAGC  
280 P E G K K L P R T F C Q L I L D P I F K V F D A I M N F K K E E T A

1501 AAAACTGATAGAGAACTGGACATCAAAGTGGACAGCAGGACAAGGACAAGGCAAAACCCCTGCTGAAGGTGTGATGCGCCGCTGGCTGCCTGCC  
313 K L I E K L D B I K L D S E D K D K E G K P L L K A V M R R W L P A

**BsaBI (1620)**  
1601 GGAGACGCTTGTGAGATGATCACCATCCACCTGCCCTCCCCTGTGACGGCCAGAAGTACCCTGCGAGCTCCTGTACGAGGGGCCCGGACGACG  
347 G D A L L Q M I T I H L P S P V T A Q K Y R C E L L Y E G P P D D

**NcoI (1706)** **PshAI (1777)**  
1701 AGGCTGCCATGGGCATTAAGCTGTGACCCCAAAGCCCTTATGATGTATATTTCCAAAATGGTGCCAACTCCGACAAAGGTCGGTTCTACGCCTT  
380 E A M G I K S C D P K G P L M M Y I S K M V P T S D K G R F Y A F

1801 TGGACGAGTCTTCTCGGGCTGTCTCCACTGGCCTGAAGTCAAGGATCATGGGGCCAACTATACCCCTGGGAAGAGGAGGACCTTCACTGAAGCCA  
413 G R V F S G L V S T G L K V R I M G P N Y T P G K K E D L Y L K P

1901 ATCCAGAGAAATCTTGTGATGATGGCCGCTACGTGGAGCCCATCGAGGATGTGCTTGTGGGAACATTTGGGCGCTGTTGGGCGTGGACAGTTCCTGG  
447 I Q R T I L M M G R Y V E P I E D V P C G N I V G L V G V D Q F L

**SfiI (2082)**  
2001 TGAAGACGGGCACCATCACCACCTTCGAGCAGCGCACAAATGCGGGTGTGAAGTTCAGCGTCAAGCTGTTGTCAGAGTGGCCGTGGAGGCCAAGAA  
480 V K T G T I T T F E H A H N M R V M K F S V S P V V R V A V E A K N

**NcoI (2156)**  
2101 CCCGGCTGACCTGCCAAGCTGGTGGAGGGGCTGAAGCGGCTGGCCAAGTCCGACCCCATGGTGCAGTGCATCATCGAGGAGTGGGAGAGCACATCATC  
513 P A D L P K L V E G L K R L A K S D P M V Q C I I E E S G E H I I

**SgrAI (2205)** **BglII (2225)** **Tth111I (2281)**  
2201 GCGGGCGCCGCGAGCTGCACCTGGAGATCTGCCTGAAGGACCTGGAGGAGGACCAGCCTGCATCCCATCAAGAAATCTGACCCGGTCTGCTCGTACC  
547 A G A G E L H L E I C L K D L E E D H A C I P I K K S D P V V S Y

**BsrGI (2366)**  
2301 GCGAGACGGTCAAGAGTGAAGAGTGAACGCTGCTCTGCCTCTCCAAGTCCCAACAAAGCACAACCGGCTGTACATGAAGGCGCGGCCCTTCCCCGACGGCCT  
580 R E T V S E E S N V L C L S K S P N K H N R L Y M K A R P F P D G L

**Clal (2410)** **BssHII (2455)**  
2401 GGCCGAGGACATGATAAAGGCGAGGTGTCCGCCGTCAGGAGCTCAAGCAGCGGGCGGCTACCTGGCCGAGAAGTACGAGTGGGACGTGGCTGAGGCC  
613 A E D I D K G E V S A R Q E L K Q R A R Y L A E K Y E W D V A E A

**BglII (2504)**  
2501 CGCAAGATCTGGTCTTTGGGCCGACGGCACCAGCCCAACATCCTACCGACATCAAGGGTGTGACGTACCTCAACGAGATCAAGGACAGTGTGG  
647 R K I W C F G P D G T G P N I L T D I T K G V Q Y L N E I K D S V

2601 TGGCCGGCTTCAGTGGGCCACCAAGGAGGGCGCACTGTGTGAGGAGAACATGCGGGGTGTGCGCTTCGACGTCCACGACGTACCCTGCACGCCGACGC  
680▶ V A G F Q W A T K E G A L C E E N M R G V R F D V H D V T L H A D A

2701 CATCCACCGGGAGGGGCCAGATCATCCCCACAGCACGGCGTGCCTCTACGCCAGTGTGCTGACCGCCAGCCAGCCTCATGGAGCCATCTACCTT  
713▶ I H R G G G Q I I P T A R R C L Y A S V L T A Q P R L M E P I Y L

2801 GTGGAGATCCAGTGTCCAGAGCAGTGGTGGTGCATCTACGGGTTTTGAACAGGAAGCGGGCCACGTGTTTCGAGGAGTCCAGGTGGCCGACCC  
747▶ V E I Q C P E Q V V G G I Y G V L N R K R G H V F E E S Q V A G T

2901 CCATGTTTGGTCAAGGCTATCTGCCCGTCAACGAGTCTTTGGCTTACCCTGACCTGAGGTCCAAACACGGGCGCCAGGCGTTCCCCAGTGTGT  
780▶ P M F V V K A Y L P V N E S F G F T A D L R S N T G G Q A F P Q C V

3001 GTTTGACCACTGGCAGATCTGCCCGGAGACCCCTTCGACAACAGCAGCCGCCACAGCCAGGTGGTGGCGGAGACCCGCAAGCGCAAGGCCTGAAAGAA  
813▶ F D H W Q I L P G D P F D N S S R P S Q V V A E T R K R K G L K E

3101 GGCATCCCTGCCCTGGACAATTCTGGACAAATTGTAGGCGCCCTCCGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCAC  
847▶ G I P A L D N F L D K L •

3201 AACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAAACAAC  
HpaI (3288) MfeI (3299)

3301 AATTGCATTCAATTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTTGGTATGGAATTTCTAAAAATACAG  
EcoRI (3384)

3401 CATAGCAAACTTTAACCTCAAATCAAGCCTCTACTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTCATTAG  
SapI (3566)

3501 CTGTTTGACGCTCACCTCTTTTCATGGAGTTTAAAGATAGTGTATTTCCCAAGGTTTGAAGTCTTTCTTCTTTATGTTTTAAATGCACTGAC

3601 CCCCCACATTCCTTTTTAGTAAATATTACAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAG  
SspI (3623) SnaI (3637)

3701 GCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTC  
141 • N

3801 CTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTTCATCTCAATGAGCACAAGCAGTCCAGGAGCATAGTCAGAGATGA  
138▶ R T Y 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

3901 GCTCTGCACATGCCACAGGGCTGACCACCTGTGCTCCACTCAGATAGGAGTGGCTGACAGCCACAATGTGTCAAAGTCTTCTG  
105▶ E R 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

4001 CCCGTTGCTCACAGCACCCCAATGGCAATGGCTTACGACAGCAGTACCCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTC  
72▶ G N 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

4101 TTGGTCTGATGGCGCCGACATGGTCTTTGTCTCATAGAGCATGGTGTATCTTCTAGTGGCAGCTCCAGCAGCTCCAGTCTGCTGAGAGA  
38▶ K T R 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

4201 TGTTGAAGGTCTTCAATGGTGGCCCTCTATAGTGGTCTTACTATGCCGATATACTATGCCGATGATTAATTTGTCAAAACAGCGTGGATGGCGTC  
5▶ N F T K M

4301 TCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACG  
AseI (4270)

4401 ACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAACCTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCA  
SpeI (4425)

4501 CGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGACT  
SnaBI (4553)

4601 GGGCATAATGCCAGGCGGCCATTTACCCTCATTGACGTCAATAGGGGCGTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCGAGTTTACC  
NdeI (4658)

4701 TAAATACTCCACCATTGACGTCAATGGAAAGTCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGGGGGGTCTTGGGGCGT

4801 CAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGACAGGTTAAATAAGAACTGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGC  
PacI (4844) PstI (4837) SdaI (4836) BspLU11I (4854)

4901 CGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGA  
BspLU11I (4854)

5001 TACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCGCTTACCGGATACCTGTCCGCCTTCTCCTTCGGAAGCGTGG  
ApaLI (5168)

5101 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTTGCCTCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCCCAGCGCTG  
ApaLI (5168)

5201 CGCCTTATCCGGTAACTATCGTCTTGTAGTCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTAT

5301 GTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG

5401 GAAAAAGAGTTGGTAGCTCTTGTATCCGGCAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATC

5501 TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAA

---

EagI (5604)

NotI (5603)

5601 TCAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAA  
5701 ACAAACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA