



PvuI (7) SgfI (6) MfeI (82)  
 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
 101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

---

Psp1406I (203) HindIII (245) Bsu36I (291)  
 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCGCTACCTGAGGCC  
 301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTAGGTAAGTTAAAGCTCAGGTCGAGACC  
 401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

---

KasI (535) AgeI (552)  
 501 TCTGTTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACCATTGGAGCTCCCGAGCCGACGCCGACCCCGATGGCGG  
 1 M D V P E P Q P D P D G G

---

XmaI (619) PvuI (698)  
 601 GGACGGCCCCGGCCACGAGCCCCGGGAGTCCCAAGACGAGCTGGACTTTTCCATCCTCTTTCGATTATGACTATCTGAACCTATCGAAGAAGAACC  
 13 D G P G H E P G G S P Q D E L D F S I L F D Y D Y L N P I E E E P

BspEI (727)  
 701 ATCGCACATAAGGCCATCAGCTCACCTCCGACTCGCATACCCGGATGATGTCCTGGACTATGGCTCAAGCCATGCAACCCCTTGCAGTCTCTCTG  
 47 I A H K A I S S P S G L A Y P D D V L D Y G L K P C N P L A S L S

BsaBI (897)  
 801 GCGAGCCCTGGCCGGTTCGGAGAGCCGATAGTATAGGGTCCAGAAGTCTGAGCCCGTCAAGCCAGCAGGGGCTTCGGGCCGAGCCCTCGGAT  
 80 G E P P G R F G E P D S I G F Q N F L S P V K P A G A S G P S P R I

SfiI (993)  
 901 CGAGTACTCCATCCCACGAAGTATGAGGACAGGGGGGCCCTCCGTGGGAGAGACGCCGCTGTCCCCGAGCAGCCGGCCCTGGCCCTGGCCGGC  
 113 E I T P S H E L M Q A G G A L R G R D A G L S P E Q P A L A L A G

NheI (1075)  
 1001 GTGGCCGACGCCGAGGTTACACTGCCCGTCCCGGCTACGAGGGCTACCGGAGCCGCTTTGCTTGGCCCGCTAGCAGCGGCTCCTTGCAGCT  
 147 V A A S P R F T L P V P G Y E G Y R E P L C L S P A S S G S S A S

PshAI (1165)  
 1101 TCATTTCTGACACCTTCTCCCCACACTCGCCCTGCGTCTACCCAAATAAGCCGGGCCGACGACCTGTGTCCCCAGTTTCAAACATCCCTGCTCA  
 180 F I S D T F S P Y T S P C V S P N N A G P D D L C P Q F Q N I P A H  
 1201 TTATCCCCAGAACCTCCAATAATGCACCTCGAACAGCCTCGCCGAGGACAGCTGCCTGGCCGACACTGCCCCGTCGCCGATCCCCG  
 213 Y S P R T S P I M S P R T S L A E D S C L G R H S P V R P A S R

DraIII (1307) FspI (1333) Eco4VII (1376)  
 1301 TCCTCCTACCCGGTGCCAAAGCGGAGGCACTTGTGCGCAGAGGCTTTGGTTGCTCCTCTGCCCGCAGCCCTACCCAGCGCTCCCGGAGCCCTCGCCAC  
 247 S S S P G A K R R H S C A E A L V A P L P A A S P Q R S R S P S P

BbrPI (1411) Asp718I (1445) EagI (1457)  
 1401 AGCCCTCGCTCAGTGGCACCGCAGGACGACAGCATCCCCGCTGGGTACCCCCACGGCCGGCTCTGTGTTCTCATGGATGCCCTCAACACCCTGGC  
 280 Q P S P H V A P Q D D S I P A G Y P P T A G S A V L M D A L N T L A

BamHI (1517) XcmI (1577)  
 1501 CACCGACTGCCCTGGGGATCCCTCCAAGATATGGAAGACCAGTCTGACCCGACGCTGTGTCCACCCTCCGTCGAAGGCTGGCCCGCCAC  
 313 T D S P C G I P S K I W K T S P D P T P V S T A P S K A G L A R H

EcoRI (1650) Asp718I (1677)  
 1601 ATCTACCTACTGTGAGTTCCTGGGCGATGTGAGCAGGAGGAGGAATTCCGCTCCAGAGTCCATCCTGCTGGTACCACCTACTTGGCCCAAGC  
 347 I Y P T V E F L G P C E Q E E R R N S A P E S I L L V P P T W P K

XhoI (1756)  
 1701 AGTTGGTCCGGCCATTCCCATCTGCAGCATCCCTGTGACTGCATCCCTCCACCACTCGAGTGGCCACTCTCCAATCAGTCGGGCTCCTATGAGCTACG  
 380 Q L V P A I P I C S I P V T A S L P P L E W P L S N Q S G S Y E L R

DraIII (1885)  
 1801 GATTGAGTCCAACCCAAGCCCCATCACGGGCCACTATGAGACGGAGGCGAGCCGTGGCGTGTCAAAGCCCCAACAGGAGGACACCTGTGGTGACG  
 413 I E V Q P K P H H R A H Y E T E G S R G A V K A P T G G H P V V Q

BglII (1937) BamHI (1964)  
 1901 CTCCACGGCTACATGGAGAACAAGCCTCTGGGGCTTACAGTCTTATTGGGACAGCAGATGAGAGGATCCTTAAGCCGACGCGCTTCTACCAAGTACACA  
 447 L H G Y M E N K P L G L Q I F I G T A D E R I L K P H A F Y Q V H  
 2001 GGATCACTGGGAAAACGTCACCAACGAGCTATGAGAAGATCGTAGGCAACCAAGGCTCTGGAGATCCCCCTGGAGCAAAGAACAACATGAGAGC  
 480 R I T G K T V T T T S Y E K I V G N T K V L E I P L E P K N N M R A

MluI (2185)  
 2101 CACCATCGACTGTGAGGATCCTGAAGCTCCGAAACGCTGACATCGAGCTGCGGAAGGGCGAGACGGACATCGGCAGGAAGAACAACGCTGTGCGCCTG  
 513 T I D C A G I L K L R N A D I E L R K G E T D I G R K N T R V R L

ApaLI (2209) SdaI (2244) Eco4VII (2279)  
 2201 GTGTTCCGCGTGCAGTCCAGAGCCAGTGGGCGCATCGTCTCCCTGAGGCTGCGTCCAACCCATCGAGTCTCTCAGCGCTCTGCCACGAGCTGC  
 547 V F R V H V P E P S G R I V S L Q A A S N P I E C S Q R S A H E L  
 2301 CCATGGTGGAGACAAGACATGGACAGCTGCCTGGTCTACGGGGCCAGCAGATGATCCTCAGGCCAGAACTTACAGCGGAGTCCAAGTTGTGTT  
 580 P M V E R Q D M D S C L V Y G G Q Q M I L T G Q N F T A E S K V V F  
 2401 CATGGAGAAGACTACAGATGGGCAGCAGATTTGGGAGATGGAAGTACGGTGGATAAAGACAAGACCAGCCTAACATGCTTTTTGTTGAGATCCCCGAG  
 613 M E K T T D G Q Q I W E M E A T V D K D K S Q P N M L F V E I P E  
 2501 TATCGAAACAAGCACATCCGCGTGCCTGAAAGTCAACTTCTACGTCATCAACGGAAGAGGAAACGAAGTCAAGCAGCAGCTTTTACCTACCCACCCAG  
 647 Y R N K H I R V P V K V N F Y I N G A G R K R R S Q P Q H T Y H P  
 2601 TCCCTGCCATCAAGACAGACCCAGCGATGAGTATGAACCATTTTGTATGTCAGCCCGCCATGGAGCCTGGGGAGCCAGCCATATTACCCACAGCA  
 680 V P A I K T E P S D E Y E P S L I C S P A H G G L G S Q P Y Y P Q H

2701 CCAATGCTGGCCGAGTCCCCCTCTGCCTTGTGGCTACCATGGCCCCGTGCCAACAGTTCGGCTCGGGGCTCTCATCCCCGATGCTCGCTACCAACAG  
713▶ P M L A E S P S C L V A T M A P C Q Q F R S G L S S P D A R Y Q Q  
2801 CAGAGCCCCGAGCTGCCCTCTACCAGAGAAGCAAGAGCTGAGTCCCGGCTGCTGGGTACACGACGCGTCCCTCTGGCAGCACCTTGGGTCTGG  
747▶ Q S P A A A L Y Q R S K S L S P G L L G Y Q Q P S L L A A P L G L

**SphI (2925)**

2901 CTGATGCCACCGCTCTGTGCTGGTGCATGCTGGTTCTCAGGGGCGAGGGCTCCACCTGCCACACACATCTCGGCCAGCCAGGAGGCTCACC  
780▶ A D A H R S V L V H A G S Q G Q G Q G S T L P H T S S A S Q Q A S P  
3001 CGTGATCCACTACTACCCACCAACCAGCAGCTTCGCGGTGGGGTCCACAGGAGTCCAGCATATCATGTACTGTGAAACTTCGGCCCCAGCTCTGCC  
813▶ V I H Y S P T N Q Q L R G G G H Q E F Q H I M Y C E N F G P S S A

KasI (3145)  
XmaI (3141)  
**SrfI (3140)**

3101 AGGCCTGGCCGCTCCCATCAACCAAGGTCAGAGGCTGAGCCCCGGCGCTACCCACAGTCATCCAACAACAGACTGCCCGAGCCAAAGAGCTGCCA  
847▶ R P G P P P I N Q G Q R L S P G A Y P T V I Q Q Q T A P S Q R A A  
3201 AAAACGGACCCAGTGACCAGAAGGAAGCTCTGCCACGGGAGTGACCGTCAAACAGGAACAGAACCTGGACCAGACCTACCTGGATGACGAGTTGATAGA  
880▶ K N G P S D Q K E A L P T G V T V K Q E Q N L D Q T Y L D D E L I D  
3301 CACACACCTTAGCTGGATACAAAACATATTATGAGACCCAGTGACTGGTCTTTGATCCGAGAAATCAAAGTGAAAGTTAATGAAATCATCAGGAAGGAGT  
913▶ T H L S W I Q N I L •

**AvrII (3454)**  
**NruI (3446)**

3401 TTTCAGGACCTCCCTCCCGAAATCAGACCTAGAAAACAAACATAATCTCGCGACGCTAGGACTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTG

**HpaI (3598)**

3501 GACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGT

MfeI (3609) EcoRI (3694)

3601 TAACAACAACAATTGCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTC

3701 TAAAAACAGCATAGCAAACTTAACTCCAATCAAGCCTCTACTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAA

3801 TGTGCATTAGCTGTTTGCAGCCTCACCTCTTTTCATGGAGTTAAGATATAGTGATTTTTCCCAAGGTTGAACTAGCTCTTCATTTCTTTATGTTTTAA

**SspI (3933)** SwaI (3947)

3901 ATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTTATTAGGCAGAATCCA

4001 GATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTA

4101 GCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAG  
141▶ • N 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

SacI (4208)

4201 TCAGAGATGAGCTCTCGCATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAA  
108▶ 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 R V A V I T D F  
4301 AGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGAT  
75▶ 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 E I H V A S I I  
4401 CTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTGTTGTCTCATAGAGCATGGTGTCTCTCAGTGGCGACCTCCACCAGCTCCAGATCC  
42▶ 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 A V E V L E L D

**BspHI (4522)**  
**XmnI (4514)** **AseI (4580)**

4501 TGCTGAGAGATGTTGAAGGCTTTCATGATGGCCCTCTATAGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGT  
8▶ Q Q S I N F T K M

SacI (4637)

4601 GGATGGCGTCTCCAGCTTATCTGACGGTCTACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCG

**SpeI (4735)**

4700 GAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGAGACTTGAAATCCCCGTGAGTCA

**SnaBI (4863)**

4799 AACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGGAAAGTCCATA

**NdeI (4968)**

4899 AGGTCATGTAAGTGGCATAATGCCAGGCGGGCATTACCCTGATGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTG

4999 GGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGG

SdaI (5146) PaeI (5154) **BspLU11I (5164)**

5099 TCGTTGGGCGGTACGCCAGGCGGGCCATTTACCCTAAGTTATGTAACGCTG 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 C C A G A A A A G C C A G G

5197 AACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAGAAAAATCAGCGCTCAAGTCAGAGGTGGCGAAACCCGA

5297 CAGGACTATAAGATAACCAGGCGTTTTCCCTGGAAAGTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCC

ApaLI (5478)

5397 TTCCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCCGTT

5497 CAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGTAGTCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA

5597 GCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAA

5697 GCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGGCAAGCAGCAGATTACGCGC

5797 AGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTT PacI (5894)

EagI (5914)

Swal (5903) **NotI (5913)**

5897 AATTAACATTTAAATC AGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGGTGGTTTTTTGTGTAATCGTAACTAACATACGCTCTCCAT

5997 CAAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA