



PvuI (7)  
SgfI (6) MfeI (82)  
1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGAAAAGTATGTCGTGTACTGGCTCCGCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) HindIII (245)  
201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCGCGTTTTCGCGCCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

BstEII (555)  
AgeI (552) NcoI (560)  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGGCCTCTGCTGACAAGAATGGCGGGAGCGTGCCTC  
1 M A S A D K N G G S V S S  
SphI (698)  
601 TGTGTCCAGCAGCCGCTGCAGAGCCGGAAGCCACCCAACCTCTCCATCACCATCCCACCACCGAGAAAGAGACCCAGGCCCTGGCAGCAGGACAGC  
13 V S S S R L Q S R K P P N L S I T I P P P E K E T Q A P G E Q D S  
701 ATGCTGCCTGAGGTTTTCAGAAATAGGAGGCTAAAGAAAAGCCAGCCAGGACCTGGGCTTCACACACCACCGCTGCCCTCCCTCCTCCCAAGA  
47 M L P E G F Q N R R L K K S Q P R T W A S H T T A C P P S F L P K  
XcmI (841)  
801 GGAAGAACCAGCCTACTTGAAGAGCGTCAGCCTCCAGGAGCCAGCCGATGGCAGGAGATTTCAGAGAAGCCCTGGCTTCCGCCGCCAGGCCTC  
80 R K N P A Y L K S V S L Q E P R S R W Q E S S E K R P G F R R Q A S  
901 ACTGTCCAGAGCATCCGCAAGGGCGCAGCCAGTGGTTTGGAGTCAGCGGCGACTGGGAGGGGCGAGCGCAGCAGTGGCAGCGCCGAGCCTGCACCAC  
113 L S Q S I R K G A A Q W F G V S G D W E G Q R Q Q W Q R R S L H H  
SphI (1004)  
1001 TGCAAGATGCGCTACGGCCGCTGAAGGCCTCGTGCCAGCGTGACCTGGAGCTCCCCAGCCAGGAGGCACCGTCTTCCAGGGCACTGAGTCCCCAAAGC  
147 C S M R Y G R L K A S C Q R D L E L P S Q E A P S F Q G T E S P K  
BamHI (1122)  
1101 CCTGCAAGATGCCAAGATTGTGGATCCGCTGGCCCGGGGCCGGCCTTCCGCCACCCGAGGAGATGGACAGGCCCCACGCCCCGACCCACCGCTGAC  
180 P C K M P K I V D P L A R G R A F R H P E E M D R P H A P H P P L T  
1201 CCCCAGGATCCTGTCCCTCACCTCCTTACCAGTGTCCGTTTGGCTACTCCACCTGCCACGCCGAAGAGAATGTCTGTGGCCACATGAGCTTGCAA  
213 P G V L S L T S F T S V R S G Y S H L P R R K R M S V A H M S L Q  
1301 GCTGCCGCTGCCCTCCTCAAGGGGCGCTCGGTGCTGGATGCCACCGGACAGCGGTGCCGGGTGGTCAAGCGCAGCTTTGCCTTCCGAGCTTCTGGAGG  
247 A A A A L L K G R S V L D A T G Q R C R V V K R S F A F P S F L E  
1401 AGGATGTGGTCGATGGGCGAGACGTTTACTCCTCTTTTTAGTAAGGAAGAAATGAGCTCCATGCTGATGATGCTTTGAGTCCCCCACTCTC  
280 E D V V D G A D T F D S S F F S K E E M S S M P D D V F E S P P L S  
BamHI (1517)  
1501 TGCCAGCTACTTCCGAGGGATCCACACTCAGCCTCCCCTGTCTCCCCGATGGGGTGCAAAATCCCTCTGAAGGAGTATGGCCGAGCCCCAGTCCCCGGG  
313 A S Y F R G I P H S A S P V S P D G V Q I P L K E Y G R A P V P G  
1601 CCCCAGGCGCGCAAGCGCATCGCTCCAAGGTGAAGCACTTTGCTTTGATCGGAAGAAGCGGCACTACGGCCTCGCGGTGGTGGCAACTGGCTGAACC  
347 P R R G K R I A S K V K H F A F D R K K R H Y G L G V V G N W L N  
PshAI (1785)  
1701 GCAGCTACCGCCGAGCATCAGCAGCACTGTGCAGCGGAGCTGGAGAGCTTGCAGACCCACCGCCCTACTTCACTACTGGCTGACCTTGCCTCATGT  
380 R S Y R R S I S S T V Q R Q L E S F D S H R P Y F T Y W L T F V H V  
1801 CATCATCAGCTGCTGGTATTTGCAGTATGGCATGCACCCGTGGGCTTTGCCAGCAGTACCACCCAGCTGGTGTGCGGAACAAAGTGTGTAC  
413 I I T L L V I C T Y G I A P V G F A Q H V T T Q L V L R N K G V Y  
1901 GAGAGCGTGAAGTACATCCAGCAGGAGAATCTTGGTTGGCCCCAGCTCGATTGACCTGATCCACCTGGGGGCAAGTTCTACCCTGCATCCGGAAGG  
447 E S V K Y I Q Q E N F W V G P S S I D L I H L G A K F S P C I R K  
BstAPI (2013) BsrBI (2041) NsiI (2082)  
2001 ACGGGCAGATCGAGCAGCTGGTGTGCGGAGCGAGACCTGGAGCGGGACTCAGGCTGCTGTGTCAGAATGACCACTCCGGATGCATCCAGACCCAGCG  
480 D G Q I E Q L V L R E R D L E R D S G C C V Q N D H S G C I Q T Q R  
MscI (2120) NcoI (2159)  
2101 GAAGGACTGCTCGGAGACTTTGGCCACTTTTGTCAAGTGGCAGGATGACACTGGGCCCCCATGGACAAGTCTGATCTGGCCAGAAGCGGACTTCCGGG  
513 K D C S E T L A T F V K W Q D D T G P P M D K S D L G Q K R T S G  
DraIII (2274)  
2201 GCTGTCTGCCACAGGACCCAGGACCTGCGAGGAGCCAGCCTCCAGCGGTGCCACATCTGGCCCGATGACATCAAGTGGCCGATCTGCACAGAGC  
547 A V C H Q D P R T C E E P A S S G A H I W P D D I T K W P I C T E  
2301 AGGCCAGGAGCAACCACAGGCTTCTGCACATGGACTGCGAGATCAAGGGCCGCCCTGCTGCATCGGCACCAAGGGCAGCTGTGAGATCACCACCG  
580 Q A R S N H T G F L H M D C E I K G R P C C I G T K G S C E I T T R  
2401 GGAATACTGTGAGTTCATGCACGGCTATTTCCATGAGGAAGCAACTCTGCTCCAGGTGCACTGCTGGACAAGGTGTGTGGGCTGCTGCCCTTCTC  
613 E Y C E F M H G Y F H E E A T L C S Q V H C L D K V C G L L P F L  
2501 AACCTGAGGTCCAGATCAGTTCTACAGGCTCTGGCTGTCTCTTCTACATGCTGGCGTGGTGCCTGCTGCTGTGGTCTTTCAAATGACCA  
647 N P E V P D Q F Y R L W L S L F L H A G V V H C L V S V V F Q M T  
2601 TCCTGAGGACCTGGAGAAGCTGGCGGCTGGCACCCTATCGCCATCATCTTCTCCTCAGTGGCATCACAGGCAACCTCGCCAGTGCATCTTTCTCC  
680 I L R D L E K L A G W H R I A I I F I L S G I T G N L A S A I F L P

2701 ATACCGGGCAGAGGTGGGCCCGCCGGCTCACAGTTCCGGCCTCCTCGCTGCCTCTTTCGTGGAGCTCTCCAGAGCTGGCCGCTGCTGGAGAGGCCCTGG  
713▶ Y R A E V G P A G S Q F G L L A C L F V E L F Q S W P L L E R P W  
2801 AAGGCTTCTCAACCTCTCGGCATCGTGCTTCTCTGTTTCATCTGTGGCCTCCTGCCCTGGATCGACAACATCGCCACATCTTCGGCTTCTCAGTG  
747▶ K A F L N L S A I V L F L F I C G L L P W I D N I A H I F G F L S  
2901 GCCTGCTGCTGGCCTTTCGCTTCTGCTTACATCACCTTCGGCACCAGCGACAAGTACCGCAAGCGGGCACTCATCTGGTGTACTGCTGGCCTTTGC  
780▶ G L L L A F A F L P Y I T F G T S D K Y R K R A L I L V S L L A F A

**BsrGI (3029)**

3001 CGGCTTCTCGCCGCCCTCGTGTGTGGCTGTACATCTACCCATAACTGGCCTGGATCGAGCACCTCACCTGCTTCCCTTACCAGCCGCTTCTGC  
813▶ G L F A A L V L W L Y I Y P I N W P W I E H L T C F P F T S R F C

**NheI (3147)**

XcmI (3141) MseI (3153)

3101 GAGAAGTATGAGCTGGACCAGGTGCTGCACTGACCGCTGGCCACACGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC  
847▶ E K Y E L D Q V L H •

**HpaI (3285) MfeI (3296)**

3201 TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAACAAT

**EcoRI (3381)**

3301 TGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTTAAAATACAGCAT  
3401 AGCAAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGTGC  
3501 TTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAACACTGCTTTCATTTCTTTATGTTTTAAATGCACTGACCTC

**SspI (3620)**

**SwaI (3634)**

3601 CCACATTCCTTTTTAGTAAATATTCCAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCC  
3701 CTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTTAGTTCCTG

141 ◀ • N R

3801 GTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCT  
137▶ 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 E  
3901 CTCTGCACATGCCACAGGGGTGACCACCTGATGGATCTGTCCACCTCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCC  
104▶ 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 G  
4001 GTTGCTCAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAAGTCTTG  
71▶ 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 K  
4101 GTCCTGATGGCCGCCGACATGGTGTGTTGTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGT  
37▶ 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 N

**BbsI (4205)**

XmnI (4201)

**AseI (4267)**

4201 TGAAGGTCTTCATGGTGGCCCTCTATAGTGAGTTCGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCC  
4▶ F T K M  
4301 AGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTCGCTCAATGGGGCGGAGTTGTACGACA

**SpeI (4422)**

4401 TTTTGAAAGTCCCGTTGATTTACTAGTCAAACAACTCCATTGACGTCAATGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGC

**SnaBI (4550)**

4501 CCATTGATGACTGCCAAACCGCATCATCATGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGTTCATGACTGGG

**NdeI (4655)**

4601 CATAATGCCAGGCGGGCATTACCCTGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCGAGTTACCCTGAA  
4701 ATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGAAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTGAG

**PacI (4841)**

**SdaI (4833)**

**BspLU11I (4851)**

4801 CCAGGCGGGCATTACCCTGTAAGTTATGTAACGCTGCAGGTTAAATAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGC  
4901 GTTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAC  
5001 CAGGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGC  
5101 TTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGC  
5201 CTTATCCGGTAACTATCGTCTTGTAGTCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTA  
5301 GGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA  
5401 AAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCA

PacI (5581) SmaI (5590)

5501 AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACACAGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCA

---

**NotI (5600)**

5601 GCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACA

5701 AAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA