



PvuI (7)  
SgfI (6) MfeI (82)  
1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCCGCAATTGAACGGGTGCCTA  
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTATGTCGTGTACTGGCTCCGCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

---

Psp1406I (203) HindIII (245)  
201 GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACCGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

---

BstEII (555)  
AgeI (552) NcoI (560)  
501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCACCATGGTCTGACGCTGCTTCTCCGCTACAACTGTG  
1 M V L T L L F S A Y K L C  
PvuI (636) RsrII (654) NcoI (682)  
601 CCGCTTTCACCATGTCAGGCCACGCGCGGGCGCGGATCGGCTGACAGTGCCTGGACCGGATCGGAGTGGTGGCGCCAGCCATGGTGGGCTGCGGGC  
13 R F F T M S G P R P G A D R L T V P G P D R S G G A S P W W A A G  
NruI (711)  
701 GGTGCGGGTCTCGCAAGTGTGCGCCGAGTGGGACTGAGGIGCAAGGCGCCCTGGAGCGTTCGCTGCCTGAGCTGCAGCAGGCGCTGTCCGAGCTGA  
47 G R G S R E V S P G V G T E V Q G A L E R S L P E L Q Q A L S E L  
801 AACAGGCAAGCGCGCGGGCTGTGGCGCGGGTCTCGCCGAGGCTTCCAGCTGGTAGAGGAAAGCTGGTGTGCCGCGCTGGCGCGGAGGTGGC  
80 K Q A S A A R A V G A G L A E V F Q L V E E A W L L P A V G R E V A  
901 CCAAGGTCTATGCGATGTATACGTCTGGACGGTGGCCTCGACTTGTGTTGCGGCTGCTCAGGCACCGGAGCTAGAGACCCGTGTGACGGCCGCGCGC  
113 Q G L C D A I R L D G G L D L L R L L Q A P E L E T R V Q A A R  
1001 TTGCTGGAGCAGATCCTGGTGGCTGAGAACCGGACCGGCTGGCGCGCATCGTCTAGGCGTGATCTTGAACCTGGCAAGGAGCGCGAGCCTGTGGAAC  
147 L L E Q I L V A E N R D R V A R I G L G V I L N L A K E R E P V E  
BspLU11I (1130) StuI (1179)  
1101 TGGCAGCAAGCGTGGCGGCATCTTGGAGCACATGTTCAAGCACTCGGAGGAGACGTGCCAGCGGCTGGTGGCGCGCCGAGGCTCGACCGGTGCTGTA  
180 L A R S V A G I L E H M F K H S E E T C Q R L V A A G G L V D A V L Y  
1201 CTGGTGCCCGCACAGACCCGCGCTGCTGCGCCACTGCGCTTTCGCTGGCGAACTGCGCGTGCACGGGGCCAGACGGTGCAGCGGTGCATGGTG  
213 W C R R T D P A L L R H C A L A L A N C A L H G G Q T V Q R C M V  
1301 GAGAAGCGCGCCCGAGTGGCTTCCGCTCGCTTCTCCAAGGAGGACGAGCTGTGCGGCTGCACGCGCTGCCTGGCGGTGGCGGTGTTGGCTACCA  
247 E K R A A E W L F P L A F S K E D E L L R L H A C L A V A V L A T  
1401 ACAAGGAGGTGGAACGCGAGGTCGAGCATTCTGGCACATTGGCGCTTGTGAGCCGCTCGTGGCATCGCTGGACCCCGCGCTTCCGCGCTGCCTGGT  
280 N K E V E R E V E H S G T L A L V E P L V A S L D P G R F A R C L V  
BstXI (1518)  
1501 GGATGCCAGTGACACAAGCCAGGGTCTGGACCAGACGACCTGCAGAGCCTGGTGTGCTGCTCGATTGCTGCGTGGAGGCTCAGTGCATAGGAGCA  
313 D A S D T S Q G R G P D D L Q S L V L L L D S S R L E A Q C I G A  
BstAPI (1612)  
FspI (1609)  
1601 TTCTACGTGCGCAGAGGCTGCCATCAAGAGCCTACAGGAAAGACCAAGGTGTTACAGCAGATCGGCGCTATCCAGAGCCTGAAACGCTGGTTTCTT  
347 F Y L C A E A A I K S L Q G K T K V F S D I G A I Q S L K R L V S  
1701 ACTCTACGAATGGCACCAGCTCGGCGTGGCCAAGCGCGCTGCGCCTATTGGGCGAGGAGTGCCAAGGCGCATCCTGCCCTGCGTGGCCAGCTGGAA  
380 Y S T N G T T S A L A K R A L R L L G E E V P R R I L P C V A S W K  
ScaI (1844) BstEII (1879)  
1801 GGAAGCTGAGGTCCAGACCTGGCTACAGCAGATCGGCTTCTCCAGTACTGCGAGAATTTCCGGAGCAGCAGGTAGATGGTACCTGCTTAAAGACTC  
413 E A E V Q T W L Q Q I G F S Q Y C E N F R E Q Q V D G D L L L R L  
AvrII (1923)  
1901 ACAGATGAAGAACTCCAGACAGACCTAGGCATGAAATCAAGCATCACCCGCAAGAGGTTCTTTAGGGAGCTCACAGAGCTCAAGACCTTCGCCAGCTACG  
447 T D E E L Q T D L G M K S S I T R K R F F R E L T E L K T F A S Y  
BamHI (2043)  
2001 CTACTTGCACCCGAGCAACCTAGCGGACTGGCTGGGACGCTGGATCCTCGCTTCCGCCAGTACACCTATGGCTGGTACGCTGCGGTCTGGACCGCTC  
480 A T C D R S N L A D W L G S L D P R F R Q Y T Y G L V S C G L D R S  
DraIII (2107)  
2101 CCTGCTGACCCGCTGTGAGCAGCAGCTCCTGGAGGACTGTGGCATCCGCTGGGAGTGCACCGCACGCGCATCCTCTGACGCCAGAGAAATGCTA  
513 L L H R V S E Q Q L L E D C G I R L G V H R T R I L S A A R E M L  
SandI (2278)  
EcoO109I (2278)  
2201 CATTCCCCGCTGCCCTGTACTGGAGGCAAGCTCAGTGGGGACACCCAGATGTCTTTATCAGTTACCGGAGGAACTCAGGGTCCCAGCTGGCCAGCCTCC  
547 H S P L P C T G G K L S G D T P D V F I S Y R R N S G S Q L A S L  
HindIII (2377)  
2301 TGAAGGTGCACCTGCAGCTTACGGCTTACGCTTTCATCGAGTGGAGAAGCTGGAAGCCGCAAATTCGAGGACAAGCTTATCCAAAGGCTCATAGC  
580 L K V H L Q L H G F S V F I D V E K L E A G K F E D K L I Q S V I A  
SphI (2447)  
2401 GGCTCGCAATTTTGTCTGCTGTGCTGCTGGGGCGCTGGATAAGTGCATGCAGGACCATGACTGCAAGGACTGGGTGCACAAGGAGATTGTGACTGCT  
613 A R N F V L V L S A G A L D K C M Q D H D C K D W V H K E I V T A

2501 TTAAGCTGTGGCAAGAACATTGTGCCCATCATTGATGGCTTTGAGTGGCTGAGCCTCAGGCGCTGCCTGAGGATATGCAGGCTGTACTCACCTTCAACG  
 647▶ L S C G K N I V P I I D G F E W P E P Q A L P E D M Q A V L T F N  
 2601 GCATCAAATGGTCCCATGAGTACCAGGAGGCCACCATCGAGAAGATCATCCGCTTCCCTACAGGGCCGCCCTCTCAGGACTCCTCTGCCGGATCGGATAC  
 680▶ G I K W S H E Y Q E A T I E K I I R F L Q G R P S Q D S S A G S D T  
NheI (2755)

2701 CAGTTTGGAGGGAGCTACGCCAATGGGTCTGCCTAACCTGTCCCCAGTCCCGTGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTGGACAA  
 713▶ S L E G A T P M G L P •

HpaI (2893)

2801 ACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACA

---

MfeI (2904)
EcoRI (2989)

2901 ACAACAATTGCATTTCATTTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAA  
 3001 TACAGCATAGCAAACCTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGC  
 3101 ATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTAGCTCTTCATTTCTTTATGTTTTAAATGCA

---

SspI (3228)
SwaI (3242)

3201 CTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTATTAGGCAGAATCCAGATGC

---

EcoO109I (3303)
141 ◀ •

3301 TCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTT

3401 AGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGA  
 140◀ 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 D S

BstXI (3532)

3501 GATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCC  
 107◀ 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 D

StuI (3667)

3601 TTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTACGACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCC  
 73◀ 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 E G  
 3701 CAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTG  
 40◀ 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 Q Q

XmnI (3809)
AseI (3875)

3801 AGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGTGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATG  
 7◀ S I N F T K M ▶

3901 GCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTG

---

SpeI (4030)

4001 TTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCT

---

SnaBI (4158)

4101 ATCCACGCCCATTTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCAT

---

NdeI (4263)

4201 GTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCGATT  
 4301 TACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGG

---

PacI (4449)
BspLU11I (4459)

4401 GCGGTGAGCCAGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAATCATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAA  
 4501 AAGGCCGCTTGTGCGGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT  
 4601 AAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGAAG  
 4701 CGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTCGGAAG  
 4801 CGTGGCGCTTATCCGGTAACTATCGTCTTGTGTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA  
 4901 GGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC  
 5001 CTTGGAAGAAAGAGTTGGTAGCTCTTGTATCCGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAA

---

PacI (5189) SwaI (5198)

5101 GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACAT

---

NotI (5208)

5201 TTAAATCAGCGGCGCAATAAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAA  
 5301 ACGAAACAAAACAACTAGCAAATAGGCTGTCCCGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA