



PvuI (7) SgfI (6) MfeI (82)  
1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGTGCTGTACTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) Bsu36I (291)  
201 GTGAACGTTCTTTTTTCGAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCGGTTCGCGCCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

AgeI (552) SphI (560)  
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGGTGAGCATGCGTTGCTGGCTCCACGCCCTGCTGGGTCTACCT  
1 M R C L A P R P A G S Y L

BstXI (632) DraIII (620) XcmI (629)  
601 GTCAGAGCCCCAAGGCAGCTCACAGTGTGCCACCATGGAGTTGGGGCCCTAGAAGTGGCTACCTGGAGTCTTAAACAGCGATGCTGACCCCGTGTGC  
13 S E P Q G S S Q C A T M E L G P L E G G Y L E L L N S D A D P V C  
701 CTCTACCATTCTATGACCAGATGGACCTGGCTGGAGAAGAAGAGATTGAGCTCTACTCAGAACCCGACACAGACACCATCAACTGCGACCAGTTCAGCA  
47 L Y H F Y D Q M D L A G E E E I E L Y S E P D T D T I N C D Q F S  
801 GGCTGTTGTGTGACATGGAAGGTGATGAAGAGACCAGGGAGGCTTATGCCAATATCGCGGAACTGGACCAGTATGTCTTCCAGGACTCCCAGCTGGAGGG  
80 R L L C D M E G D E E T R E A Y A N I A E L D Q Y V F Q D S S Q L E G  
901 CCTGAGCAAGGACATTTTCATAGACACATAGACCAGATGAAGTATCGGTGAGAGTATGGAGATGCCAGCAGAAGTTGGCGAGAAAAGTCAGAAAAGA  
113 L S K D I F I E H I G P D E V I G E S M E M P A E V G Q K S Q K R  
1001 CCCTCCAGAGGAGCTTCCGGCAGACCTGAAGCACTGGAAGCCAGCTGAGCCCCACTGTGGTACTGGCAGTCTCCTAGTGGGACCAGTGGAGGACT  
147 P F P E E L P A D L K H W K P A E P P T V V T G S L L V G P V S D  
1101 GCTCCACCTGCCCTGCCTGCCACTGCCTGCGCTGTTCAACCAGGAGCCAGCTCCGGCCAGATGCGCTGGAGAAAACCGACCAGATTCCCATGCCCTT  
180 C S T L P C L P L P A L F N Q E P A S G Q M R L E K T D Q I P M P F

SandI (1236) Bsu36I (1231)  
1201 CTCCAGTTCCTCGTTGAGCTGCCTGAATCTCCCTGAGGGACCCATCCAGTTTGTCCCCACCATCTCCACTCTGCCCATGGGCTCTGGCAAATCTCTGAG  
213 S S S S L S C L N L P E G P I Q F V P T I S T L P H G L W Q I S E  
1301 GCTGGAACAGGGTCTCCAGTATATTCACTACCATGGTGGAGTGCCCGAGCCAGCAAGTACCCCTCCAGTGGATTCACTGTCCACGGCTCCCAA  
247 A G T G V S S I F I Y H G E V P Q A S Q V P P P S G F T V H G L P

SphI (1460)  
1401 CATCTCCAGACCGCCAGGCTCCACCAGCCCTTCGCTCCATCAGCCACTGACCTGCCAGCATGCCTGAACCTGCCCTGACCTCCCAGCAAACATGAC  
280 T S P D R P G S T S P F A P S A T D L P S M P E P A L T S R A N M T  
1501 AGAGCACAAGACGTCCCCACCAATGCCCGCAGCTGGAGAGGTCTCCAACAAGTTCAAAATGGCCTGAGCCGGTGGAGCAGTCTACCCTCACTG  
313 E H K T S P T Q C P A A G E V S N K L P K W P E P V E Q F Y R S L  
1601 CAGGACACGTATGGTGCCGAGCCCGAGCCCGGATGGCATCTAGTGAGGTGGATCTGGTGACGGCCAGGCTGGAGAGGAGCAGCAGCAAGAGCCTGG  
347 Q D T Y G A E P A G P D G I L V E V D L V Q A R L E R S S S K S L

NotI (1788)  
1701 AGCGGAACTGGCCACCCGGACTGGGCGAAGCGCAGCTGGCCAAAGGAGGCTGGCTGAGGTGCTGTTGGCTGCCAAGGAGCACCGGCCGCGCTGA  
380 E R E L A T P D W A E R Q L A Q G G L A E V L L A A K E H R P R E  
1801 GACACGAGTGATTGCTGTGCTGGCAAAGCTGGTCAGGGCAAGAGCTATTGGCTGGGCGAGTGGCCGGCTGGGCTTGTGGCGGCTTCCCAGTAC  
413 T R V I A V L G K A G Q G K S Y W A G A V S R A W A C G R L P Q Y

Tth111I (1900) SdaI (1956) SapI (1968)  
1901 GACTTTGTCTTCTGTCCCCTGCCATTGCTTGAACCGTCCGGGGGATGCCTATGGCTGCGAGATCTGCTTCTCCTGGGCCACAGCCACTCGTGG  
447 D F V F S V P C H C L N R P G D A Y G L Q D L L F S L G P Q P L V  
2001 CGCCGATGAGGTTTTAGCCACATCTTGAAGAGACCTGACCGGTTCTGCTCATCTAGACGCCTTCGAGGAGCTGGAAGCGCAAGATGGCTTCTGCA  
480 A A D E V F S H I L K R P D R V L L I L D A F E E L E A Q D G F L H

RsrII (2109) BbrPI (2103)  
2101 CAGCACGTGCGGACCGGACCGCGGAGCCCTGCTCCCTCCGGGGGCTGCTGGCCGGCCTTTCCAGAAGAAGCTGCTCCGAGGTTGCACCCTCCTCCTC  
513 S T C G P A P A E P C S L R G L L A G L F Q K K L L R G C T L L L  
2201 ACAGCCGGGCCCGGGCCGCTGGTCCAGAGCTGAGCAAGCCGACGCCCTATTTGAGCTGTCCGGCTTCTCCATGGAGCAGGCCAGGCATACGTGA  
547 T A R P R G R L V Q S L S K A D A L F E L S G F S M E Q A Q A Y V  
2301 TGCGCTACTTTGAGAGCTCAGGGATGACAGAGCACCAAGACAGAGCCCTGACGCTCCTCCGGGACCGCCACTTCTTCTCAGTACAGCCACAGCCCTAC  
580 M R Y F E S S G M T E H Q D R A L T L L R D R P L L L S H S H S P T  
2401 TTTGTGCCGGCAGTGTGCCAGCTCTCAGAGCCCTGCTGGAGCTTGGGGAGGACGCAAGCTGCCCTCCACGCTCACGGACTCTATGTCGGCTGCTG  
613 L C R A V C Q L S E A L L E L G E D A K L P S T L T G L Y V G L L  
2501 GGCCGTGACGCCCTGACAGCCCCCGGGGCCCTGGCAGAGCTGGCAAGCTGGCCTGGGAGCTGGGCCGACAGATCAAAGTACCTACAGGAGGACC  
647 G R A A L D S P P G A L A E L A K L A W E L G R R H Q S T L Q E D

SacII (2659)  
2601 AGTTCCATCCGAGACGTGAGGACCTGGGCGATGGCCAAAGGCTTAGTCCAACACCCAGCCGGCCGAGAGTCCGAGCTGGCCTTCCCAGCTTCTCCT  
680 Q F P S A D V R T W A M A K G L V Q H P P R A A E S E L A F P S F L

2701 CCTGCAATGCTTCTGGGGCCCTGTGGCTGGCTCTGAGTGGCGAAATCAAGGACAAGGAGCTCCCGCAGTACCTAGCATTGACCCCAAGGAAGAAGAGG  
713▶ L Q C F L G A L W L A L S G E I K D K E L P Q Y L A L T P R K K R  
2801 CCCTATGACAACTGGCTGGAGGGCGTGCACGCTTTCTGGCTGGGCTGATCTTCCAGCCTCCCGCCGCTGCCTGGGAGCCCTACTCGGGCCATCGGGCGG  
747▶ P Y D N W L E G V P R F L A G L I F Q P P A R C L G A L L G P S A

Acc65I (2936)

2901 CTGCCTCGGTGGACAGGAAGCAGAAGGTGCTTGCAGGTACCTGAAGCGGCTGCAGCCGGGGACAATTGCGGGCGCGGAGCTGCTTGAAGTGTGCACTG  
780▶ A A S V D R K Q K V L A R Y L K R L Q P G T L R A R Q L L E L L H C

BbrPI (3034)

BspLU11I (3099)

3001 CGCCCACGAGGCCGAGGAGGCTGGAATTTGGCAGCAGTGGTACAGGAGCTCCCCGGCCGCTCTCTTTTCTGGGCACCCGCCTCACGCCTCCTGATGCA  
813▶ A H E A E E A G I W Q H V V Q E L P G R L S F L G T R L T P P D A  
3101 CATGTAAGTGGCAAGGCTGGAGGCGGCGCAAGACTTCCCTGGACCTCCGAGCAGTGGCATTGGCCCTCTGGATTGGGGAGCCTCGTGGGAC  
847▶ H V L G K A L E A A G Q D F S L D L R S T G I C P S G L G S L V G  
3201 TCAGCTGTGTACCCGTTTCAGGGCTGCCTTGAAGCAGACGGTGGCGCTGTGGGAGTCCCTGCGGAGCATGGGGAGACCAAGCTACTTCAGGAGCAGAG  
880▶ L S C V T R F R A A L S D T V A L W E S L R Q H G E T K L L Q A A E  
3301 GGAGAAGTTCACCATCGAGCCTTTCAAAGCCAAGTCCCTGAAGGATGTGGAAGACCTGGGAAAGCTTGTGCAGACTCAGAGGACGAGAAGTTCCTCGGAA  
913▶ E K F T I E P F K A K S L K D V E D L G K L V Q T Q R T R S S S E

BamHI (3497)

3401 GACACAGCTGGGAGCTCCCTGCTGTTCCGGACCTAAAGAACTGGAGTTTGCCTGGGCCCTGTCTCAGGCCCCAGGCTTTCCCCAACTGGTGGCGGA  
947▶ D T A G E L P A V R D L K K L E F A L G P V S G P Q A F P K L V R  
3501 TCCTCACGGCCTTTCTCCCTGCAGCATCTGGACCTGGATGCCTGAGTGAACAAGATCGGGGACGAGGGTGTCTCGAGCTCTAGCCACCTTCCC  
980▶ I L T A F S S L Q H L D L D A L S E N K I G D E G V S Q L S A T F P  
3601 CCAGCTGAAGTCTTGGAAACCTCAATCTGTCCAGAACAACATCACTGACCTGGGTGCCTACAACTCGCCGAGGCCCTGCCTCGCTCGTGCATCC  
1013▶ Q L K S L E T L N L S Q N N I T D L G A Y K L A E A L P S L A A S

BsrGI (3716)

BspEI (3768)

3701 CTGCTCAGGCTAAGCTTGTACAATACTGCATCTGCGACGTGGGAGCCGAGAGCTTGGCTCGTGTCTCCGGACATGGTGTCCCTCCGGGTGATGGAGC  
1047▶ L L R L S L Y N N C I C D V G A E S L A R V L P D M V S L R V M D  
3801 TCCAGTACAACAAGTTTACGGCTGCCGGGCCAGCAGCTCGCTGCCAGCCTTCCGAGGTGTCTCATGTGGAGACGCTGGCGATGTGGACGCCACCAT  
1080▶ V Q Y N K F T A A G A Q Q L A A S L R R C P H V E T L A M W T P T I

NheI (3960)

3901 CCCATTAGTGTCCAGGAACACCTGCAACAACAGGATTCACGGATCAGCCTGAGATGATCGCTAGTGCCAGACATGATAAGATACATTGATGAGTTTG  
1113▶ P F S V Q E H L Q Q Q D S R I S L R •

HpaI (4098)

4001 GACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGT

MfeI (4109)

EcoRI (4194)

4101 TAAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTC

4201 TAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAA

SapI (4376)

4301 TGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCAAGGTTTGAAGTACTGCTCTTCATTTCTTTATGTTTTAA

SspI (4433)

SwaI (4447)

4401 ATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATAGGCAGAATCCA

4501 GATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTA

4601 GCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAAGCAGTCAAGGAGCATAG  
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

BstXI (4737)

4701 TCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAA  
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

4801 AGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTTCAGCACAGACAGTACCCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAGATGAT  
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

4901 CTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCC  
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

XmnI (5014)

AseI (5080)

5001 TGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGAAGTCTGATTTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGT  
8▶ Q Q S I N F T K M

5101 GGATGGCGTCTCCAGCTTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTCGCTCAATGGGGCGG

SpeI (5235)

5201 AGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCCATGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAA

SnaBI (5363)

5301 CCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGGAAAGTCCATAAG

NdeI (5468)

5401 GTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCTTACGCTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGG

5501 CAGTTTACGTAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGGTC

SdaI (5646) PacI (5654) BspLU11I (5664)

5601 GTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACC

5701 GTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG

5801 ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCG

ApaLI (5978)

5901 GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGC

6001 CCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG

6101 AGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA

6201 GTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAA

PacI (6394)

6301 AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATT

SwaI (6403) NotI (6413)

6401 AACATTTAAATCAGCGGCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAA

6501 ACAAACGAAACAAACAAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA