



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

101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCGCTACCTGAGGCC **HindIII (245)**

301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
NaeI (441) 401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

NcoI (560)
BstEII (555)
AgeI (552) 501 TCTGTTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACCATTGGGAGACTGCCTTGACCTTCTTTAGGAGTCGACT **KasI (535)** 1▶ M G D C L D L L L G V V L

BbsI (624) 601 CCTGACCAGTCTGCACTTGGAAATGCTTCTGCTTCTTTGATGGTTGGAGAGCCATCTATCTTCTGCAACCTCACCAAGTGCCCGAGTCCCAAC
13▶ L T S P A L G M S S C F F D G W R A I Y L S C N L T Q V P Q V P N

701 ACCACCAAGAGCCTCCTGCTCAGCTTCAACTATATCAGGACAGTCACAACCGCATCTTCCCCTTCTGGAGCAGTTGCACTGCTGGAGCTGGGAAC
47▶ T T K S L L L S F N Y I R T V T T A S F P F L E Q L Q L L E L G T

BamHI (854) 801 AGTTTACCCCTTGACAATTTACAGAGAAGCCTCCGAAACCTGCCAATCTCAGGATCCTAGACCTGGGTGGAAGTCAGATAAACTTCTTGCATCCAGA
80▶ Q F T P L T I Y R E A F R N L P N L R I L D L G G S Q I N F L H P D

901 TGCTTTTACAGGACTGCCACCTCACTAACTTGCAGTGTCTTCTGCTGCTCCTCCGATGCTGATTAAAAGATGGTTATTTAGAAAATTTGGCATCT
113▶ A F Q G L P H L T K L R L F S C G L S D A V L K D G Y F R N L A S

1001 TTGACTCACTGGACCTATCAAAAATAAAATTCAGAGTCTTTACCTTCCCTCATTCCGGGAACGAATTCCTTGAAGTCCATGACTTTTCTTCA
147▶ L T H L D L S K N K I Q S L Y L H P S F R E L N S L K S I D F S F

BsrGI (1184) 1101 AAAAAACCTATTGTATGTGAGCAGGAGTTCAAACCCCTCAAGGGAAAACACTCTCTTCTTAAGCCTTGCTGATAACCAGCTGTACAGCAGGGTCTC
180▶ N K I P I V C E Q E F K P L Q G K T L S F L S L A D N Q L Y S R V S

BspHI (1286) 1201 AGTGGACTGGAACAAGTGTCTGAACCCATTGAAACATGGTCTGGAACCCTAGATGTTTCTGGCAATGGCTGGGGGTGGACATCATGAGAACTTC
213▶ V D W N K C L N P F R N M V L E T L D V S G N G W G V D I M R N F

BamHI (1386) **XcmI (1398)** 1301 AGCAATGCCATCAATGGGAGCCAGATTTTCTCTTTGGTCTTACTCGTCACATTATGGTTTCTCATTGGCTTCTCCAACTGAAGGATCCTGACTCC
247▶ S N A I N G S Q I F S L V L T R H I M G S S F G F S N L K D P D Y

MscI (1415) 1401 ACACCTTTGCTGGGCTGGCCAGGAGCTCGATGATACAGCTGGACATTTACATGGGTATATCTTCTCCCTGAACTTCCGAATCTTTGAGACACTCCAGGA
280▶ H T F A G L A R S S M I Q L D I S H G Y I F S L N F R I F E T L Q E

1501 GCTGAAGGTCTAAACCTCGCCTACAACAAGATAAACAGTATTTCGAGGAAATGCGTTTATGGACTCGACAACCTCCAAGTTCTCAATATATCATATAAC
313▶ L K V L N L A Y N K I N S I S R N A F Y G L D N L Q V L N I S Y N

PstI (1660) 1601 CTTCTGGGGAATATAIAATTATGATTTGATGGACTACCTAAGGTGGCCTATATTGACCTGCAGAAGAATCACATTGGGATCATTGAGACAAAACAT
347▶ L L G E L Y N Y D F D G L P K V A Y I D L L Q K N H I G I I Q D K T

1701 TCAAATCCTGGGAAATTAATACCTTGGACCTCCGGGATAATGCTCTTAAACAATTTATTTTCTTCAAGCATTCTAATATCTTCTTAAGTGGCAA
380▶ F K F L G K L N T L D L R D N A L K T I Y F L P S I P N I F L S G N

Asp718I
Acc65I 1801 TAAGCTGATGACTTTGCCAAACATCCCACCTACAGCCAATTTATCCAATTATCAGAGAATAGGCTAGAAAATCTGAATGATCTCTACTTCTCTCCAG
413▶ K L M T L P N I P L T A N F I Q L S E N R L E N L N D L Y F L L Q

HindIII (1990) 1901 GTACCTCATCTCCAGATTCTCATTTAAATCAAATCGCTTTTCTTTTGTCCACAAAACCATGCCCTTCCAGAGAATCCAGCCTAGAAAAGCTTTTCC
447▶ V P H L Q I L I L N Q N R F S F C H Q N H A P S E N S S L E K L F

2001 TTGGAGAAAATATGTTGCAGCTTGCCTGGGAAACTGGGTGATGTTGGGATATTTTAAAGGGCTTCTCATCTCCAACCTCTATCTAAATAAACTA
480▶ L G E N M L Q L A W E T G S C W D I F K G L S H L Q L L Y L N K N Y

2101 CCTGAATTTCTTCCACAGGAGATTTTCTCATCTGACTGCACTGAGGGGACTCAGCCTCAAAGACAACAGGCTGACTGTTCTTTTCTGGCGACTTA
513▶ L N F L P P G V F H H L T A L R G L S L K D N R L T V L F P G D L

DraIII (2270) 2201 CCTGTAACCTAGAGATCCTGGATATATCTGAAACCAGCTCCTCTCTCTGATCCTGATTTATTTGCATCACTGAGTGCATAGACATAACTCATAACA
547▶ P A N L E I L D I S G N Q L L S P D P D L F A S L S A I D I T H N

BspLU11I (2381) 2301 ATTTCACTGTGAGTGTGAAGTCACTGCTTTTATCCATTGGCTCAATCAAACCAATATACAATAGCTGGGTCTCCAGCAGACATGACTGCATGATACC
580▶ N F I C E C E L S A F I H W L N Q T N I T I A G S P A D M Y C M Y P

2401 AAATCTCTGGCTGGGTTTCCATTTACAGTCTTTCCACAGAAAGTTGTGAAGAAGAAGAAGTTTGTAGAGTCCCTAAAGTTTCCCTTTTCTGCTGCT
613▶ N S L A G V S I Y S L S T E S C E E E E V L E S L K F S L F I L V

2501 ACTGTCACCTGACTCTGTTCTCGTATCACCTCGCGGTTACTAAGTTGAGGGGTTTGTTCATCTGTTATAAGAAAGCCAGAGTCTGCTGTTCA
647▶ T V T L T L F L V I T L A V T K F R G F C F I C Y K K A Q S L L F

SnaBI (2628) 2601 AGGACCCATCAAGGGAAGAGAATCAGATACGTACAAATACGATGCCTATTTGTGCTTACAGTAGCAAAGACTTTGAATGGGTGCAGAACGCTTTGCTCAA
680▶ K D P I K G R E S D T Y K Y D A Y L C F S S K D F E W V Q N A L L K

2701 ACACCTGGATGTCCAGTACCACAGCCAAAACAGATTTAACCTGTGCTTTGAAGAAAGAGACTTTATGCCGGGGAAAACACATTGCCAACATCCAGGAT
713▶ H L D V Q Y H S Q N R F N L C F E E R D F M P G E N H I A N I Q D
2801 GCTGTGTGGAGCAGCAGAAAGATTGTCTGTCTCGTGAGCAGACTTCCTTAGAGACGGGTGGTGCCTCGAAGCCTTCAGTTATGCCAGAGCAGGTGCT
747▶ A V W S S R K I V C L V S R H F L R D G W C L E A F S Y A Q S R C
KasI (2914) SandI (2938)
2901 TAGCTGACCTCAATGGCGCCCTCATCATGGTGGTGGTGGGTCCTGTCCAGTTCATCTGATGAGGCATCAGTCCATCAGAGGGTTCGTCCAGAAACG
780▶ L A D L N G A L I M V V V G S L S Q F H L M R H Q S I R G F V Q K R
ScaI (3002) BglII (3021) Ppu10I (3063)
3001 GCAGTACTTGAGGTGGCCTGAAGATCTCCAGGATGTTGACTGGTTTCTCAATAAACTCTCTCAATGCATTCTAAAGAAAGAAAAGGAAAGGAAGAAAGAC
813▶ Q Y L R W P E D L Q D V D W F L N K L S Q C I L K K E K E R K K D
NsiI (3063)
MscI (3151)
3101 AGTGCTATTCAAGCTGCAAAATGTAACACCACATCTCCTAAGCGGTGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAATA
847▶ S A I Q L Q N V T T I S •
NheI (3145)
3201 GAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAACAAGTTAAACAACAACAATTG
HpaI (3283) MfeI (3294)
3301 CATTCAATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATCTAAAATACAGCATAG
3401 CAAAACCTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTT
3501 TGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGATTTTCCCAAGGTTTGAAGTACTGCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCC
SspI (3618) SmaI (3632)
3601 ACATTCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCT
3701 TCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGT
141▶ • N R T
3801 GTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCT
137▶ 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
BstXI (3922)
3901 CTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGT
103▶ 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 N
StuI (4057)
4001 TGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGT
70▶ 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 T
4101 CCTGATGGCCGCCGACATGGTCTTGTGCTCCTCATAGAGCATGGTATCTTCTCAGTGGCAGCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTG
37▶ 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
BspHI (4207)
BbsI (4203) AseI (4265)
4201 AAGGTTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAG
3▶ F T K M
4301 CTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACAT
SpeI (4420)
4400 TTTGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGC
SnaBI (4548)
4499 CCATTGATGACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCAAGTAGGAAAGTCCATAAGGTGATGACTGGG
NdeI (4653)
4599 CATAATGCCAGGCGGGCCATTTACCGTCAATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTAAGTCCCAAGTGGGCAGTTTACCGTAA
4699 ATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTCAG
PstI (4832) SdaI (4831) PacI (4839) BspLU11I (4849)
4799 CCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTG 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 G C C A G A A C C G T A A A A A G G C C
4897 GCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCAGCGTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGAT
4997 ACCAGGCGTTTTCCCTGGAAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGC
ApaLI (5163)
5097 GCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGC
5197 GCCTTATCCGGTAACTATCGTCTTGAGTCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATG
5297 TAGGGCGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG
5397 AAAAAAGATTGGTAGCTCTTGATCCGGCAAACAAACCCTGGTAGCGGTGGTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCT
PacI (5579) SmaI (5588)
5497 CAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAAT

EagI (5599)

NotI (5598)

5597 C AGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAA
5697 CAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTATCGAA