



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
MfeI (82) EcoNI (96)

101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **HindIII (245)** **Bsu36I (291)**
EcoNI (287)

201 GTGAACGTTCTTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACAGCGCCCGCCCTACCTGAGGCC

301 GCCATCCACGCGGTTGAGTCGCGTCTGCCGCTCCCGCTGTGGTGCCTCTGAACTGCGTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
NaeI (441)

401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGTCTTTGTTTCGTTT

KasI (535) **AgeI (552)** **BstEII (555)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCCACCATGGGTGATTTATTCTTATACTTTACGGTGTTC CAAT
1 M G D L F L Y F Q V F P M

601 GTGGACATTTAAGAGACAGTTTCTATCCTTTTTAACATGATCCTAATTTCTGGACTCCTTGGGCTAGATGGTTTCCAAAACCTCGCCTGTGATGTC
13 W T L K R Q F P I L F N M I L I S G L L G A R W F P K T L P C D V

701 ACTCTGGATGCCCAAATACCCATGTGATTGTGGACTGCACAGACAAGCATTTGACAGAAATCTGGAGGTATTCTGCCAATGCCACCAACCTTACCC
47 T L D A P N T H V I V D C T D K H L T E I P G G I P A N A T N L T

Clal (865)

801 TCACCATTAACACATAGCAGGCATCTCTCCAGCCTCCTTACCAGGCTGGACCATCTGGTGGAGATCGATTTTCAGATGCAACTGCGTACCTGTTTCGACT
80 L T I N H I A G I S P A S F H R L D H L V E I D F R C N C V P V R L

ApaLI (918)

901 GGGGCCAAAAGACAACGTGTGCACAAAAGGCTACAGATTAACCCAAACAGCTTTAGCAAACCTCAGTATTTAAAATCTCTTTACCTGGATGAAAACAG
113 G P K D N V C T K R L Q I K P N S F S K L T Y L K S L Y L D G N Q

XbaI (1002) **Bsu36I (1012)** **BspHI [m] (1071)**

1001 CTTCTAGAAATACCTCAGGATCTTCTCCAGCTTACAGCTGCTGAGCCTGGAGGCCAACACATCTTTTTGATCATGAAGGAGAATCTAACAGAAGTGG
147 L L E I P Q D L P P S L Q L L S L E A N N I F L I M K E N L T E L

1101 CCAACCTAGAAATACTCTACCTGGGCCAAAACCTGTTACTATCGTAACCTTGAATGTTTCATTTACTATCGAAAAAGATGCTTTCTAAATATGAGAAA
180 A N L E I L Y L G Q N C Y Y R N P C N V S F T I E K D A F L N M R N

1201 TTTAAAATTGCTCTCCCTAAAAGATAACAAATATCTCAGCTGTCCCACTGTTTTGCCATCTAGTTTGACAGAECTCTATCTTTACAATAACATCATTACA
213 L K L L S L K D N N I S A V P T V L P S S L T E L Y L Y N N I I T

DraIII (1397)

1301 AAAATCCAAGAAGATGATTTTAATAACCTCAGTCAACTACAAGTCTTGATCTGAGTGGAAATTGCCCTCGTTGTTATAATGTTCCATTTCTTGCACAC
247 K I Q E D D F N N L S Q L Q V L D L S G N C P R C Y N V P F P C T

Ppu10I (1445) **NsiI (1445)** **BbrPI (1495)**

1401 CCTGTGAAAATAATCTCCCTACAGATTGATCCAATGCTTTTATGATGACAGATTGCAAGTCTTACGTCTACACAGTAACTCTCTACAGCAGT
280 P C E N N S P L Q I D P N A F D A L T E L Q V L R L H S N S L Q H V

1501 GCCCAAAGATGTTTAAAAACATTAACAACTTAAAGAACTAGATCTTTCCAAAACCTTCTGGCCAAAAGAAATGGGGATGCCAAATTTTACATCTT
313 P Q R W F K N I N K L K E L D L S Q N F L A K E I G D A K F L H L

Ppu10I (1676) **NsiI (1676)**

1601 CTTCATAACCTTGCAACTGGATCTGTCTTTCAATTATGATCTTTCAGGCTACCATGCAAGTATAAACTATCAGATGCATTTTCTTACTGAAAAAAT
347 L H N L V N L D L S F N Y D L Q V Y H A V I N L S D A F S S L K K

PmeI (1745)

1701 TGAAGTTTTGCGAATCAAAGGCTATGTCTTAAAGAGCTGAATAGTTAAACCTTCCATTACATAATCTTCCCAATCTTGAAGTCTTGTATCTTGG
380 L K V L R I K G Y V F K E L N S L N L F P L H N L P N L E V L D L G

1801 CACTAACTTTATAAAAATGCTAACCTCAGCATTTTTAAACCAATTTAAACATTGAAATTCATAGATCTTTCAGTGAATAAAATATCACCTTCGGGAGAT
413 T N F I K I A N L S I F N Q F K T L K F I D L S V N K I S P S G D

Bsp120I (1951)

1901 TCACCTGAAGGTGGTTTCTGCTCTAACAGGAGAAGTCTGTAGAAGGCCATGGGCCAGGTCCTTGAACACTGCATTATTTTCAGATATGATGAGTATG
447 S P E G G F C S N R R T S V E G H G P Q V L E T L H Y F R Y D E Y

2001 CAAGGAGTCCCGGTCGAAGAGCAAAGAGCCTCCTTCTTTCTTACCTCTTAATGAAGATTGTTATATGTATGGACAGACCTTGGACCTTAGTAGAAAATAA
480 A R S C R S K S K E P P S F L P L N E D C Y M Y G Q T L D L S R N N

XcmI (2178)

2101 TATATTTTTATCAAGCCTTCTGATTTCCAGCATCTTTCTTCTCAAATGCCTAAACTTATCAGGAAATAGCATTAGCCAGACGCTTAATGGAAGTGAA
513 I F F I K P S D F Q H L S F L K C L N L S G N S I S Q T L N G S E

SspI (2222)

2201 TTTGAGCCTTTAGTGGAGTTGAAATATTTGGACTTCTCTAACAAATCGGCTTGATTTACTCTACTCAACTGCATTTGAGGAGCTGCACAACCTGGAAGTCC
547 F Q P L V E L K Y L D F S N N R L D L L Y S T A F E E L H N L E V

2301 TAGATATAAGCAGTAACAGCATTATTTCAATCAGAAGGAATTAACATGCTAAATTTTACCAAGAACCTCAAGGTTCTGAGGAACTGATGATGAA
580 L D I S S N S H Y F Q S E G I T H M L N F T K N L K V L R K L M M N

XcmI (2423) **XmnI (2443)**

2401 CTATAATGACATTGCTACCTCCACGACAGGACCATGGAGAGTGAATCTCTTCAAATCCTGGAGTTCAGAGGCAACCATTTGGATATTTTATGGAGAGAT
613 Y N D I A T S T S R T M E S E S L Q I L E F R G N H L D I L W R D

2501 GGTGATAACAGATACTTAAATCTTTAAGAATCTGCTAAACTTAGAAGAGCTAGACATCTCTGAAAATCTCTGAGTTTCTTACCTTTGGGAGTTTTTG
647 G D N R Y L K F F K N L L N L E E L D I S E N S L S F L P L G V F

2601 ATAGTATGCCTCCAATCTAAAGACTCTCTCCTTAGCCAAAAATGGGCTCAAGTCTTTCAGTTGGGAAAGACTACAGAGTCTGAAGAATCTAGAAACTTT
680▶ D S M P P N L K T L S L A K N G L K S F S W E R L Q S L K N L E T L
2701 GGACCTCAGCTTCAACCGACTGAAGACTGTCCCTGAGAGATTATCCAAGTGTTCGCCGAGCCTCAAGAACTCATACTTAAGAATAATCAAATCAGGTGC
713▶ D L S F N Q L K T V P E R L S N C S R S L K K L I L K N N Q I R C
2801 CTGACAAAGTATTTCTCCAAGGTGCTTCCAGTTGCGACATCTGGACCTCAGCTCAAATAAAATTCAGTTATCCAAAAGACGAGTTTCCAGAAAACG
747▶ L T K Y F L Q G A F Q L R H L D L S S N K I Q V I Q K T S F P E N

2901 TCCTCAACAATCTGAACATTTTGTCTGCATCACAATCGATTTCTGTGCAACTGTGATGCTGTGTGGTTTGTCTGGTGGTTAACCATACCGAGGTGAC
780▶ V L N N L N I L F L H H N R F L C N C D A V W F V W W V N H T E V T
3001 TATTCCTTACTTGGCCACAGATGTGACTTGCATGGACAGGAGCACACAAGGCCAGAGTGTAGTCTCTCTGGATCTATATACTGTGAGTTAGATCTG
813▶ I P Y L A T D V T C M G P G A H K G Q S V V S L D L Y T C E L D L

3101 ACTAATTCATCTGTTCTCACTTTCCATATCAGCAGTTCTCTCTGTGATGATGATCACAATAGCAAACCATCTCTATTTCTGGGATGTGTGGTATAGTT
847▶ T N F I L F S L S I S A V L S L M M I T I A N H L Y F W D V W Y S

3201 ATCATTCTGTAAAGCCAAAATAAAAGGGTATCGACGTCTGATATCACCAATTTCTGTATGATGCTTTTATTGTATATGACACTAAAGACCCAGCAGT
880▶ Y H F C K A K I K G Y R R L I S P N S C Y D A F I V Y D T K D P A V

3301 GACAGAGTGGGTTTTGGACGAGCTGGTGCCAAATGGAGACCCGAGAGAGAAGTGTTTAATTTATGTCTTGAGGAAAGGGACTGGTTACCAGGGCAG
913▶ T E W V L D E L V A K L E D P R E K C F N L C L E E R D W L P G Q
3401 CCTGTTCTGAAAATCTTCCAGAGCATAAGCTTAGCAAAAAGACAGTGTTTGTGATGACAGACAAGTACGCAAAGACTGAAAATTTAAGATAGCA
947▶ P V L E N L S Q S I Q L S K K T V F V M T D K Y A K T E N F K I A

3501 TTTACTTATCCCATCAGAGGCTCATGGATGAAAAAGTGGATGTAATCATCTTGATATTCCTTGAGAAGCCCTTCAGAAGTCCAAGTTTCTCCAATCCG
980▶ F Y L S H Q R L M D E K V D V I I L I F L E K P L Q K S K F L Q L R
3601 GAAGAGCTCTGTGGCAGTTCTGTCTTGTGAGTGGCCAAACCCAGGCTCACCCGTACTTGKGCAGTGTCTGAAAATGCCCTGCCACAGACAAT
1013▶ K R L C G S S V L E W P T N P Q A H P Y F W Q C L K N A L A T D N

3701 CACGTGACCTACAGTCAGGTGTTCAAAGAGACAGCCTAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCA
1047▶ H V T Y S Q V F K E T A

3801 GTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACACAACAATTGCATTCAT

3901 TTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTTCAAATACAGCATAGCAAACT
4001 TTAACCTCCAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCC
4101 TCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTGAACTAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCC

4201 CTTTTTAGTAAAAATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAAT
4301 ATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTG
1414 • N R T Y K

4401 AGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGTCTCTGCACA
134▶ 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 R C M

4501 TGCCACAGGGGCTGACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCAC
101▶ 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 S V

4601 AGCAGACCAATGGCAATGGCTTCCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATG
68▶ 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 R I

4701 GCCGCCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCT
34▶ 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 F T K

4801 TCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCT
1▶ M

4901 GACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAA

5001 GTCCCGTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGAT

5100 G T A C T G C C A A A C C G C A T C A T C A T G G T A A T A G C G A T G A C T A A T C G T A G A T G T A C T G C C A A G T A G G A A G T C C C A T A A G G T C A T G T A C T G G G C A T A A T G C

5200 C A G G C G G G C A T T T A C C G T C A T T G A C G T C A A T A G G G G G C T A C T T G C C A T A T G A T A C A C T T G A T G T A C T G C C A A G T G G G C A G T T T A C C G T A A A T A C T C C A

5300 C C C A T T G A C G T C A A T G G A A A G T C C C A T T G G C G T T A C T A T G G G A A C A T A C G T C A T T A T T G A C G T C A A T G G G C G G G G T C G T T G G G C G G T C A G C C A G G C G G

PstI (5425)
SdaI (5424) **PacI (5432)** **BspLU11I (5442)**

5400 GCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAA TTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCT
 5498 GCGGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCG
 5598 TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGCGCTTTCTC
ApaLI (5756)

5698 ATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTACGCCGACCGCTGCGCCTTATC
 5798 CGGTAACATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT
 5898 GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG
 5998 TTGGTAGCTCTTGATCCGCAAAACAAACCACCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGA

EagI (6192)
NotI (6191)

6098 TCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCC
PacI (6172) SwaI (6181)

6198 GCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAA
 6298 ACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA