



PvuI (7)
SgfI (6)
MfeI (82)

1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA

101 GAGAAGGTGGCGCGGGTAAACTGGAAAGTATGTCGTGACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203)
HindIII (245)
Bsu36I (291)

201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC

301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

NcoI (566)
BspEI (558)

501 TCTGTTTGTCCGGCTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGCTCCGGACACCATGGCCGAAGCGCCAGGTGGTGGAGACCGA

1▶ M A E A P Q V V E T D

Eco4VII (626)

601 CCCGACTTCGAGCCGCTGCCCCGGCAGCGCTCTGTACTGGCCGCTGCCAGGCCGAGTTTAAACAGTCAAACCTGACCACCTCCAGTCCGGCGCCG

11▶ P D F E P L P R Q R S C T W P L P R P E F N Q S N S T T S S P A P

NotI (709)

701 TCGGGCGCGCGGCCCAACCCGACGCCGCGGAGCCTGGCTCGGCTCCGCTGTCAGCACCGACTTTATGAGCAACCTGAGCCTGCTGGAGGAGA

45▶ S G G A A A N P D A A A S L A S A S A V S T D F M S N L S L L E E

Bsp120I (882)

801 GTGAGGACTTCGCGCGGGCGCCAGGCTGCGTGGCCGTGGCGCGCGGCTGCGGCCAGCAGGGGCTGTGCGGGGACTTCCAGGGCCCGAGGCGGGCTG

78▶ S E D F A R A P G C V A V A A A A A S R G L C G D F Q G P E A G C

ApaI (901)

901 CGTGCACCCAGCGCCGACAGCCCCACCGCCGCGCTGTGCGAGCCCCACCCGTGCTCCTCCGCTGCCGCGCCGCGGGGCCACTCGCGGGA

111▶ V H P A P P Q P P P T G P L S Q P P P V P P S A A A A A G P L A G

MluI (1031)
BsiWI (1049)

1001 CAGCCGCGCAAGACCAGCTCGTCGCGCCGCAACGCGTGGGCAACCTGTCGTACGCCGACCTCATACCAAGGCCATCGAGAGCTCAGCCGAGAAGAGGC

145▶ Q P R K T S S S R R N A W G N L S Y A D L I T K A I E S S A E K R

BglII (1112)
EcoRI (1188)

1101 TCACCCTGTCGAGATCTACGAGTGGATGGTGAAGAGCGTGCCTACTTCAAGGATAAGGGCGACAGCAACAGCTCGGCGGCTGGAAGAATTCAATTCG

178▶ L T L S Q I Y E W M V K S V P Y F K D K G D S N S S A G W K N S I R

XmnI (1253)
BsrBI (1293)

1201 CCACAATCTGTCCCTTACAGCAAGTTTATTGAGTGCAGAATGAAGGAACCTGAAAGAGTTCTTGGTGGATGCTCAATCCAGAGGGAGGCAAGAGCGGA

211▶ H N L S L H S K F I R V Q N E G T G K S S W W M L N P E G G K S G

NcoI (1325)

1301 AAATCACCCCGAGAAGAGCTGCGTCCATGGACAACAACAGTAAATTTGCTAAGAGCCGAGGGCGGGCTGCTAAGAAAAAGCATCTCTCCAGTCTGGGC

245▶ K S P R R R A A S M D N N S K F A K S R G R A A K K K A S L Q S G

SmaI (1456)

1401 AAGAGGGTCTGGAGACAGCCCTGGGTCTCAGTTTTCTAAGTGGCCTGCGAGTCTGGGTCCACAGCAACGATGACTTTGATAACTGGAGTACATTTCCG

278▶ Q E G P G D S P G S Q F S K W P A S P G S H S N D D F D N W S T F R

BspHI (1550)

1501 TCCTGAACCAGCTCAAATGCTAGTACCATCAGTGGGAGACTTTCTCCCATCATGACAGAGCAGGATGACCTGGGAGATGGGGACGTGCATTCCCTGGTG

311▶ P R T S S N A S T I S G R L S P I M T E Q D D L G D G D V H S L V

1601 TATCCACCCTCTGCTGCAAGATGGCGTCTACGCTGCCAGTGTCTGAAATCAGCAATCCAGAAAACATGGAGAACCTTCTGGATAATCTCAACCTTC

345▶ Y P P S A A K M A S T L P S L S E I S N P E N M E N L L D N L N L

XcmI (1731)

1701 TCTCGTCCCCAACATCTTAACTGTGTCACCCAGTCTCGCTGGCAGCATGATGCAGCAGACCATGCTATTGTTTTGCACCGCCAAACACCAGTCT

378▶ L S S P T S L T V S T Q S S P G S M M Q Q T P C Y S F A P P N T S L

1801 AAATTCACCCAGTCCAACTACTCAAAGTACACATACGGCAAATCCAGCATGAGCCCTTTGCCCCAGATGCCTATGCAGACACTTCAGGACAGCAAATCA

411▶ N S P S P N Y S K Y T Y G Q S S M S P L P Q M P M Q T L Q D S K S

AgeI (1992)

1901 AGTTACGGAGGATTGAACCAAGTATAACTGTGCCCCAGGACTCTTGAAGAGTTGTTGACTTCTGACTCTCTCCCAATGACATTATGTCACCGGTTG

445▶ S Y G G L N Q Y N C A P G L L K E L L T S D S P P H N D I M S P V

Bsp120I (2052)

2001 ATCCCGGAGTGGCCCAACCAACAGTCCGGTCTGGGCCAAAATGTAATGATGGGCCAATTCGGTTCATGCCAGCGTATGGCAGCCAGGCATCTCATAA

478▶ D P G V A Q P N S R V L G Q N V M M G P N S V M P A Y G S Q A S H N

XcmI (2121)

2101 CAAAATGATGAACCCAGCTCCACACCCACCTGGACATGCACAGCAAACGCTTCGGTCAACGGCCGTACCTGCCCATGTGGTGAACACCATGCCT

511▶ K M M N P S S H T H P G H A Q Q T A S V N G R T L P H V V N T M P

BbsI (2232)

2201 CACACATCTGCCATGAACCGCTTGACCCCGTGAAGACACCTTACAAGTGCCTCTGTCCACCCCATGCAGATGAGTGCCTGGGACGACTACTCCTCGG

545▶ H T S A M N R L T P V K T P L Q V P L S H P M Q M S A L G S Y S S

2301 TGAGCAGCTGCAATGGCTATGGTAGGATGGGTGCCTCCACCAGGAGAAGCTCCCAAGTGACTTGGATGGCATGTTTATTGAGCGCTTGGACTGTGACAT
578▶ V S S C N G Y G R M G V L H Q E K L P S D L D G M F I E R L D C D M
Eco47III (2381)

2401 GGAGTCCATCATTGGAATGACCTCATGGATGGAGATACCTTGGATTTAACTTTGATAATGTGTTGCCAACCAAAGCTTCCACACAGTGTCAAGACT
611▶ E S I I R N D L M D G D T L D F N F D N V L P N Q S F P H S V K T
DraIII (2483)

2501 ACAACACACAGCTGGGTGTCAGGCTAAGAGTTAGTGAGCAGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAAT
645▶ T T H S W V S G •
MscI (2547)

2601 GCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAATTGCATT
HpaI (2679) MfeI (2690)

2701 CATTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCATAAATACAGCATAGCAAA
EcoRI (2775)

2801 ACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGA
HpaI (2679) MfeI (2690)

2901 GCCTCACCTTCTTCATGGAGTTAAGATATAGTGATTTTCCCAAGTTTGAAGTACTCTTCATTCTTTATGTTTAAATGCACTGACCTCCACAT
EcoRI (2775)

3001 TCCCTTTTGTAGAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCAT
SspI (3014) SwaI (3028)

3101 AATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTAC
141◀ • N R T Y

3201 TTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGC
135◀ 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 R C

3301 ACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCT
102◀ 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 S
BstXI (3318)

3401 CACAGCAGACCAATGGCAATGGCTTCAACAGACAGAGTACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGT
69◀ 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 R
StuI (3453)

3501 ATGGCCGCCCCGACATGGTGCTTGTGCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACAGCTCCAGATCCTGCTGAGAGATGTTGAAGG
35◀ 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 F T
BbsI (3599) XmnI (3595)

3601 TCTTCATGGTGGCCCTCTATAGTGAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGATGGCGTCTCCAGCTTA
2◀ K M
AseI (3661)

3701 TCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGG
SpeI (3816)

3801 AAAGTCCCCTGTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTG
SnaBI (3944)

3901 ATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCCATAAAGTTCATGTAAGTGGCATAAT
NdeI (4049)

4001 GCCAGGCGGGCCATTTACCGTCAATGAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGCAGTTTACCGTAAATACTC
NdeI (4049)

4101 CACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTACGCCAGGC
PacI (4235)

4201 GGGCATTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTCT
PstI (4228) SdaI (4227) BspLU11I (4245)

4301 GGGCTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAAGGTTGGCGAAACCCGACAGGACTATAAAGATACAGCGG
BspLU11I (4245)

4401 TTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTC
ApaLI (4559)

4501 ATAGTCAAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCTTATC
ApaLI (4559)

4601 CGGTAAGTATCGTCTTGAAGTCCAAACCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGT
ApaLI (4559)

4701 GCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAG
ApaLI (4559)

4801 TTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGA

PacI (4975) SwaI (4984) NotI (4994)

4901 TCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCC

5001 GCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAA

5101 ACTAGCAAAATAGGCTGTCCCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA