



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

HindIII (245)
Psp1406I (203) **PvuII (239)** **EcoNI (287)**
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACAGCGCCCGCCGCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTGCAGTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTAACTCTACGCTTTGTTTCGTTT

NcoI (560) **BstEII (555)**
KasI (535) **AgeI (552)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCCACATGGAAAAATGTTTCCTTCAGTCGTCATGCTGACCTG
601 CATTTCCTGCTAATATCTGTTCTGTGAGTTATGCGCCGAAGAAAAATTTTCTAGAAGCTATCCTTGTGATGAGAAAAAGCAAAATGACTCATGTTAT
13> I F L L I S G S C E L C A E E N F S R S Y P C D E K K Q N D S V I
1> M E N M F L Q S S M L T C

SaII (716)
701 GCAGAGTGCAGCAATCGTGCAGTACAGGAAGTTCCCAAAACGGTGGGCAAATATGTGACAGAACTAGACCTGTCTGATAATTTTCATCACACACATAACGA
47> A E C S N R R L Q E V P Q T V G K Y V T E L D L S D N F I T H I T

BsrGI (858) **Bst1107I (884)**
801 ATGAATCATTTCAGGGGTGCAAAATCTCACTAAAATAAATCTAAACCACAACCCCAATGTACAGCACCAGAACGAAATCCCGGTATACAATCAAATGG
80> N E S F Q G L Q N L T K I N L N H N P N V Q H Q N G N P G I Q S N G

BbsI (958)
901 CTTGAATATCACAGACGGGCGATTCTCTCAACCTAAAAACCTAAGGGAGTTACTGCTTGAAGACAACAGTTACCCCAATACCTCTGGTTTGCAGAG
113> L N I T D G A F L N L K N L R E L L L E D N Q L P Q I P S G L P E
1001 TCTTTGACAGAATAGTCTAATTCAAAACAATATATAACAATAACTAAAGAGGGCATTTCAGACTTATAAATTTGAAAAATCTCTATTTGGCCTGGA
147> S L T E L S L I Q N N I Y N I T K E G I S R L I N L K N L Y L A W
1101 ACTGCTATTTTAAACAAAGTTTGCAGAAAACTAACATAGAAGATGGAGTATTTGAAACGCTGACAAATTTGGAGTTGCTACTATCTTTCAATTTCTCT
180> N C Y F N K V C E K T N I E D G V F E T L T N L E L S L S F N S L
1201 TTCACATGTGCCACCCAACTGCCAAGCTCCCTACGAAAATTTTTCTGAGCAACCCAGATCAAATACATTAGTGAAGAAGATTTCAAGGGATTGATA
213> S H V P P K L P S S L R K L F L S N T Q I K Y I S E E D F K G L I
1301 AATTTAACATTACTAGATTTAAGCGGAACTGTCCGAGGTGCTCAATGCCCATTTCCATGCGTGCCTTGTGATGGTGGTGCCTCAATTAATATAGATC
247> N L T L L D L S G N C P R C F N A P F P C V P C D G G A S I N I D
1401 GTTTTGCTTTTCAAACTTGACCAACTCGATCACTAAACCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCCTGGTTAAAAATATGCCTCATCT
280> R F A F Q N L T Q L R Y L N L S G S T S L R K I N A A W F K P H L

EcoRI (1516)
1501 GAAGGTGCTGGATCTTGAATTCACATTTAGTGGGAGAAATAGCCTCTGGGGCATTTTTAAACGATGCTGCCCGCTTAGAAATACTTGACTTGTCTTTT
313> K V L D L E F F N Y L V G E I A S G A F T T M L P R L E I L D L S F
1601 AACTATATAAAGGGAGTTATCCACAGCATATTAATATTTCCAGAAACTTCTCAAACCTTTTGTCTCTACGGGCAATTGCAATTAAGAGGTTATGTGTCC
347> N Y I K G S Y P Q H I N I S R N F S K L L S L R A L H L R G Y V F

ClaI (1786)
1701 AGGAACTCAGAGAAGATGATTTCCAGCCCTGATGCGACTTCCAAACTTATCGACTATCAACTTGGGTATTAATTTTATTAAGCAAATCGATTTCAAAC
380> Q E L R E D D F Q P L M Q L P N L S T I N L G I N F I K Q I D F K L
1801 TTTCAAATTTTCTCAATCTGGAATTTACTTGTGAGAAAAAGAAATATCACCGTTGGTAAAAGATACCCGGCAGAGTTATGCAAAATGTTCTCT
413> F Q N F S N L E I I Y L S E N R I S P L V K D T R Q S Y A N S S S

BspEI (1913) **BstBI (1953)** **DraIII (1991)**
1901 TTTCAACGTCATATCCGGAAACGACGCTCAACAGATTTTGTGTTGACCCACATTCGAATTTTATCATTTCACCCGCTCTTTAATAAAGCCACAATGTG
447> F Q R H I R R R S T D F E F D P H S N F Y H F T R P L I K P Q C
2001 CTGCTTATGAAAAAGCCTTAGATTTAAGCCTCAACAGTATTTTCTTATTGGGCCAAACCAATTTGAAAAATCTTCTGACATTGCCTGTTTAAATCTGTC
480> A A Y G K A L D L S L N S I F F I G P N Q F E N L P D I A C L N L S
2101 TGCAAATAGCAATGCTCAAGTGTAAAGTGAAGTGAATTTTCCAGCATTCTCATGTCAAATATTTGGATTTGACAAAACAAATAGACTAGACTTTGATAAT
513> A N S N A Q V L S G T E F S A I P H V K Y L D L T N N R L D F D N

BglII (2238)
2201 GCTAGTGTCTTACTGAATTTGCCGACTTGAAGTTCTAGATCTCAGCTATAATTCACACTATTTTCAAGATAGCAGGGCTAACACATCATCTAGAATTTA
547> A S A L T E L S D L E V L D L S Y N S H Y F R I A G V T H H L E F
2301 TTTCAAATTTTCAAAATCTAAAAGTTTAAACTTGAGCCACAACAACATTTTACTTTTAAACAGATAAGTAACTGGAAAGCAAGTCCCTGGTGAAT
580> I Q N F T N L K V L N L S H N N I Y T L T D K Y N L E S K S L V E L
2401 AGTTTTCAGTGGCAATCGCCTTGACATTTTGTGGAATGATGACAAACAGTATATCTCCATTTTCAAAGTGTCAAGAATCTGACACGCTGCTGGATTTA
613> V F S G N R L D I L W N D D N R Y I S I F K G L K N L T R L D L
2501 TCCCTTAATAGGCTCAAGCATCCAAATGAAGCATTCTTAATTTGCCAGCGAGTCTCACTGAAGTACATATAAATGATAATATGTTAAAGTTTTTTA
647> S L N R L K H I P N E A F L N L P A S L T E L H I N D N M L K F F

XhoI (2629)
2601 ACTGGACATTACTCCAGCAGTTTCTCGTCTCGAGTTGCTTACTTACGTGGAACAAACTACTCTTTTAACTGATAGCCTATCTGACTTTACATCTTC
680> N W T L L Q Q F P R L E L L D L R G N K L L F L T D S L S D F T S S
2701 CCTTCGGACACTGCTGCTGAGTCAACAGGATTTCCACCTACCCTCTGGCTTTCTTTCTGAAGTCAAGTCTGAAGCACCTCGATTTAAGTTCCAAT
713> L R T L L L S H N R I S H L P S G F L S E V S S L K H L D L S S N
2801 CTGCTAAAAACAATAACAAATCCGCACTTGAACCTAAGACCACCAAAATATCTATGTTGGAAGTACACGGAAACCCCTTTGAATGCACCTGTGACA
747> L L K T I N K S A L E T K T T T K L S M L E L H G N P F E C T C D

BsaBI (2923)

2901 TTGGAGATTTCCGAAGATGGATGGATGAACATCTGAATGTCAAATCCAGACTGGTAGATGCATTTGTGCCAGTCCTGGGGATCAAAGAGGGGAAGAG
780▶ I G D F R R W M D E H L N V K I P R L V D V I C A S P G D Q R G K S

PstI (3041) NcoI (3080)

3001 TATTGTGAGTCTGGAGCTAAACAACCTGTGTTTCAGATGCTACTGCAGTGATATTATTTTTCTTACGTTCTTTATCACCACCATTGTTTATGTTGGCTGCC
813▶ I V S L E L T T C V S D V T A V I L F F F T F F I T T M V M L A A

3101 CTGGCTCACCATTGTTTACTGGGATGTTTGGTTTATATAATGTGTGTTAGCTAAGGTAAGGCTACAGGTCTCTTCCACATCCCAAACCTTTCT
847▶ L A H H L F Y W D V W F I Y N V C L A K V K G Y R S L S T S Q T F

Psp1406I (3295)

3201 ATGATGCTTACATTTCTTATGACACCAAAGATGCCTCTGTTACTGACTGGGTGATAAATGAGCTGCGCTACCACCTTGAAGAGAGCCGAGACAAAACGT
880▶ Y D A Y I S Y D T K D A S V T D W V I N E L R Y H L E E S R D K N V

XmaI (3330)

SmaI (3326)

EcoO109I (3326) MscI (3338)

3301 TCTCCTTTGTCTAGAGGAGAGGGATTGGGACCCGGGATTGGCCATCATCGACAACCTCATGCAGAGCATCAACCAAAGCAAGAAAACAGTATTTGTTTTA
913▶ L L C L E E R D W D P G L A I I D N L M Q S I N Q S K K T V F V L

3401 ACCAAAAATATGCAAAAAGCTGGAACCTTAAACAGCTTTTACTTGGCTTTGCAGAGGCTAATGGATGAGAACATGGATGATTATTTATCTCGC
947▶ T K K Y A K S W N F K T A F Y L A L Q R L M D E N M D V I I F I L

3501 TGGAGCCAGTGTTACAGCATTCTCAGTATTTGAGGCTACGGCAGCGGATCTGTAAGAGCTCCATCCTCCAGTGGCCTGACAACCCGAAGGCAGAAGGCTT
980▶ L E P V L Q H S Q Y L R L R Q R I C K S S I L Q W P D N P K A E G L

BamHI (3685)

3601 GTTTTGGCAAACCTGAGAAATGTGGTCTTGACTGAAAATGATTACGGTATAACAATATGTATGTCGATTCCATTAAGCAATACGGATCCTATCCCTAT
1013▶ F W Q T L R N V V L T E N D S R Y N N M Y V D S I K Q Y G S Y P Y

MscI (3788)

NdeI (3725) **NheI (3782)**

3701 GATGTGCCAGACTATGCTGGCTATCCATATGATGTTCTGATTATGCTGGATACCCTTATGATGTGCCAGACTATGCCTAAAGCTAGCTGGCCAGACATG
1047▶ D V P D Y A G Y P Y D V P D Y A G Y P Y D V P D Y A •

3801 ATAAGATACATTGATGAGTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCA

HpaI (3920) MfeI (3931)

3901 TTATAAGCTGCAATAAACAGTTAAACAACAACCAATTGCATTCTTTTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCT

EcoRI (4016)

4001 CTACAAATGTGGTATGGAATCTAAATAACAGCATAGCAAAACTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCAT

4101 AGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTCCAAGGTTTGAAGTACG

SwaI (4269)

4201 TCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAT

EcoO109I (4330)

4301 GTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGG

4401 ACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAG
1414 • N R T Y K L P I L E E I T T K V L K G N M E I L

BstXI (4559)

4501 CACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGC
1164 V F C D P A Y 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

StuI (4694)

4601 CTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCTGCCAATGTAGGCT
824 R V A V I T D F 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

4701 CAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGCTTGTGCTCCTCATAGACATGGTGATCTTCTCAGTGGC
494 I H V A S I I 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

BspHI (4844)

BbsI (4840)

4801 GACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGATGGCCCTCTATAGTGAGTCTGATTATACTATGCCGATATACTATGCCGA
164 V E V L E L D Q Q S I N F T K M

4901 TGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCG

SpeI (5057)

5001 CCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCCTGTTGATTTACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGA

SnaBI (5185)

5100 CTTGGAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACT

NdeI (5290)

5200 GCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATA

5300 CACTTGATGACTGCCAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATT

PstI (5469)

SdaI (5468) PacI (5476)

5400 ATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAA TTAAGAACATGTGAGCA

5498 AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAA

5598 GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCCTGCCGCTTACCGG

5698 ATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGGTGTAGGTCGTTTCGCTCCAAGCTGGGC

ApaI (5800)

5798 TGTGTGCACGAACCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAG

5898 CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATT

5998 TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT

6098 TGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTAAAG

EagI (6236)

PacI (6216) SwaI (6225) **NotI (6235)**

6198 GGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC **AGCGGCCG**CAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCG

6298 TAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATC

6398 GAA