



175

PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA

MfeI (82)
101 GAGAAAGTGGCGGGGTAAGTGGAAAGTGTGCTGTACTGGCTCCGCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

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

301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTCGCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

NcoI (560)
BstEII (555)
501 TCTGTTCTGCGCGGTTACAGATCCAAGCTGTGACCGCGGCTACCTGAGATCACCGGTACCATTGGGCGGTGCTTGCAGCTCCTTTACCACCTCTCCAG

1 M G G A C S S F T T S S S
601 CCCTACCATTATTCTACCTCAGTCACCGACAGCAAGGCTATGCAAGTGGAGAGCTGCTCCTCAGCCGTGGGGTAAGTAACAGAGGGTAAGTGAAGG

13 P T I Y S T S V T D S K A M Q V E S C S S A V G V S N R G V S E K
BsrGI (758)
701 CAGTTAACAGTAACACAGTTCAGCAGCATCCATCAACACCGAAGAGGACACAGTCTGTACATCTCACCACCCTGAGGACTTGTGGATAACAGTC

47 Q L T S N T V Q Q H P S T P K R H T V L Y I S P P P E D L L D N S
BsaBI (827)
801 GGATGTCCTGCCAGGATGAGGGTGTGGATTGGAATCTGAGCAGAGCTGCAAGTGTGGATGGAGGATCCCCCTCCAACCTCAGTAACATGAGCACCAG

80 R M S C Q D E G C G L E S E Q S C S M W M E D S P S N F S N M S T S
901 TTCCTACAATGATAACACTGAGGTACCTCGTAAATCACGAAAACGAAATCAAAGCAGAGGCCGGGGTCAAACGACGAGATTGTGAAGAATCTAATATG

113 S Y N D N T E V P R K S R K R N P K Q R P G V K R R D C E E S N M
DraIII (1033)
1001 GATATATTGATGCCGACAGTGCCAAAGCACCTCACTATGTGCTTCTCAGCTTACCACGGACAACAAAGGCAACTCAAAGCGGAAATGGAACATTGG

147 D I F D A D S A K A P H Y V L S Q L T T D N K G N S K A G N G T L
XcmI (1104) **BsrGI (1188)**
1101 AAAACCAAAAAGGAACGGAGTAAAGAAGAGCCCTATGTTGTGTGGACAATCTCTGTTAAAAGTGGGGAAAGGAGCTGAAGATAGTTGACAACTGA

180 E N Q K G T G V K K S P M L C G Q Y P V K S E G K E L K I V V Q P E
1201 GACACAGCACCGAGCTCGGTACCTGACTGAGGGCAGCCGTGGCTCAGTGAAAGATAGAACACAGCAAGGCTTCTACAGTAAAGCTGGAAGGCCATAAT

213 T Q H R A R Y L T E G S R G S V K D R T Q Q G F P T V K L E G H N
StuI (1367)
1301 GAACCTGTAGTGTGCAAGTGTGTTGGGCAACGACTCTGGACGAGTGAACACATGGATTTATCAGGCTGCAGAGTAACTGGACGAAATACAACCTC

247 E P V V L Q V F V G N D S G R V K P H G F Y Q A C R V T G R N T T
1401 CTTGCAAAGAAGTGGACATTGAAGGCACACTGTTATAGAAGTGGCCTTGATCCTAGCAACAACATGACACTGGCGTGGACTGCGTAGGGATATTGAA

280 P C K E V D I E G T T V I E V G L D P S N N M T L A V D C V G I L K
1501 ATTGAGGAATGCTGATGTCGAAGCCAGAATAGGAATTGCTGGTCCAAGAAGAAAAGCACTCGTGCCAGATTGGTTTTTCGAGTAAATATCATGAGGAAA

313 L R N A D V E A R I G I A G S K K K S T R A R L V F R V N I M R K
HindIII (1683)
1601 GATGGCTCCACTTTGACACTGCAACACCCCTCTTCTCCAATTTGTGTACTCAGCCAGCAGGAGTGCCAGAAATCTTAAAGAAAAGCTTGCATAGCTGTT

347 D G S T L T L Q T P S S P I L C T Q P A G V P E I L K K S L H S C
1701 CAGTGAAGGAGAAGAAGTGTGTTTTAATCGCAAGAACTTCTGAAAGGAACTAAAGTATTTTTCCAAGAAAATGTTTCTGATGAAAACCTTTGGAA

380 S V K G E E E V F L I G K N F L K G T K V I F Q E N V S D E N S W K
1801 GTCAGAAGCTGAAATGATGGAACCTATTTATCAGAACTCATCTTATTGTGAAGTTCTCCCTATCATGACCAACATATAACTTTGCTGTGTCAGTG

413 S E A E I D M E L F H Q N H L I V K V P P Y H D Q H I T L P V S V
BglII (1927)
1901 GGAATATATGATGACAAATGCTGGAAGTCTCATGATGTTCAACCATTCACTTACACTCCAGACCCAGCAGCTGGTGCTTTGAATGTAATGTGA

447 G I Y V V T N A G R S H D V Q P F T Y T P D P A A A G A L N V N V
2001 AGAAGGAAATATCTAGTCCAGCAAGACCTTGTCTTTTGAAGAGCCATGAAAGCAATGAAAACACTGGATGTAATTTAGATAAGGTAATATTATCCC

480 K K E I S S P A R P C S F E E A M K A M K T T G C N L D K V N I I P
BglII (2182)
2101 TAATGCCCTGATGACTCCACTCATACCAAGCAGTATGATTAAGAGTGAAGATGTTACTCCAATGGAAGTAAACAGCAGAAAAAGATCTTCCACTATTTTT

513 N A L M T P L I P S S M I K S E D V T P M E V T A E K R S S T I F
2201 AAGACTACAAAGTCTGTTGGATCAACTCAGCAACATTAGAAAACATCTCAAACATAGCAGGAAATGGCTCTTTTTTATCACCATCATCTTCCACCTAC

547 K T T K S V G S T Q Q T L E N I S N I A G N G S F S S P S S S H L
2301 CTTCTGAAAATGAAAACAGCAGCAGATTTCAGCCCAAGGCATACAACCCAGAGACCCTGACAACCTATTCAAACCCAGGACATCTCACAGCCTGGTACTTT

580 P S E N E K Q Q Q I Q P K A Y N P E T L T T I Q T Q D I S Q P G T F
XbaI (2485)
2401 TCCAGCAGTTTCTGCTTCTAGTCAGCTGCCAACAGCGATGCACTATTGCAGCAGGCTACACAGTTTCAGACAAGAGAACTCAGTCTAGAGAGATATTA

613 P A V S A S S Q L P N S D A L L Q Q A T Q F Q T R E T Q S R E I L
2501 CAGTCAGATGGTACAGTGGTAAATTTGTCAAACTGACTGAGGCATCAACAACAGCAGCAGTACCACCTACAAGAACAAGCACAGACTTTACAGCAGC

647 Q S D G T V V N L S Q L T E A S Q Q Q Q Q S P L Q E Q A Q T L Q Q
ScaI (2689)
2601 AGATTTTCATCAAATATTTTTCCATACCAAATAGTGTGAGTCAGCTTTCAGAATACTATTTCAGCAGCTGCAAGCAGGGAGTTTCACAGGCACTACTGCTAG

680 Q I S S N I F P S P N S V S Q L Q N T I Q Q L Q A G S F T G S T A S
Tth111I (2719)
2701 TGGCAGCAGTGAAGTGTGACTTGGTCCAAAGTTTTAGAGGCACAGCAGGTTATCTTTCAGTTTTATTTTCTGCTCCAGATGGTAATGAGAATGTT

713 G S S G S V D L V Q Q V L E A Q Q Q L S S V L F S A P D G N E N V

2801 CAAGAGCAGCTTAGTGCAGATATTTTTCAACAAGTCAGTCAAATTCAGAGTGGTGAAGCCCTGGAATGTTTTCTCAACAGAGCCAACAGTCCATACCA
747▶ Q E Q L S A D I F Q Q V S Q I Q S G V S P G M F S S T E P T V H T
2901 GACCAGATAATTTATTACCTGGAAGAGCTGAAAGTGTTCATCCACAGTCTGAAACACGTTATCTAATCAACAGCAGCAGCAGCAGCAGCAACAGCAAGT
780▶ R P D N L L P G R A E S V H P Q S E N T L S N Q Q Q Q Q Q Q Q V
3001 GATGGAATCTTCAGCCGCAATGGTGATGGAGATGCAACAGAGTATCTGCCAGGCGAGTGGCCAGATTGAGTCAAGATTTCCCTTCAACTGCTTCAGCA
813▶ M E S S A A M V M E M Q Q S I C Q A A A Q I Q S E L F P S T A S A
3101 AATGGAACCTTCAGCAATCGCCAGTTTACCAGCAGACTTCTCACATGATGAGTGCATTGTCTACCAATGAGGATATGCAAAATGCAAGTGTGAATTTGTTTT
847▶ N G N L Q Q Q S P V Y Q Q T S H M M S A L S T N E D M Q M Q C E L F
3201 CTTCTCCTCTGCAGTTTCTGGAATGAACTTCTACAACACACAGAGTGGCAACCCCTGGCACTACCATGTTTCAGACATCAAGTTTCAGGAGA
880▶ S S P P A V S G N E T S T T T T Q Q V A T P G T T M F Q T S S G D

NcoI (3353)

PshAI (3351)

3301 TGGAGAAGAACTGGAACACAAGCAAAACAGATTGAGAACAGTGTCTTTGAGCATTGGTCCAATGCAACATAGTGGGGACAATCAACCTCAAGTTAAC
913▶ G E E T G T Q A K Q I Q N S V F Q Q C T M V Q M Q H S G D N Q P Q V N
3401 CTTTTTTCATCCAAAAAGTATGATGAGTGTTCAGAATAGTGGTACCAACAACAAGGTAATGGTTTATTCCAGCAAGGGAATGAGATGATGCTCACTTC
947▶ L F S S T K S M M S V Q N S G T Q Q Q G N G L F Q Q G N E M M S L

BbsI (3522)

3501 AATCTGGAATTTTTGTCAGCAGTCTTCTCATTCCAGGCCCCAATTTTTTCATCCTCAAAATCCTATTGCCGATGCTCAGAACCTTTCCAGGAAACTCA
980▶ Q S G N F L Q Q S S H S Q A Q L F H P Q N P I A D A Q N L S Q E T Q
3601 AGTTTCTCTCTTTCATAGTCCAAATCCTATTGTCCACAGTCAGACTTCTACAACCTCCTGAAACAAATGCAGCCTCCAATGTTTCACTCTCAAAGTACC
1013▶ G S L F H S P N P I V H S Q T S T T S S E Q M Q P P M F H S Q S T
3701 ATTGCTGTGTACAGGGCTCTTCAGTTCCTCAAGCAGCAGTCAACCAATATTTCTTTCCAGAGTCCCATGAATAATCTTCAGACTAACACAGTAG
1047▶ I A V L Q G S S V P Q D Q Q S T N I F L S Q S P M N N L Q T N T V
3801 CCCAAGAAGCATTTTTTGTCAGCACCAGCACTCAATTTCTCCACTTCAGTCAACATCAAACAGTGAACAACAAGCTGCTTCCAACAGCAAGCTCCAATATC
1080▶ A Q E A F F A A P N S I S P L Q S T S N S E Q Q A A F Q Q Q A P I S
3901 ACACATCCAGACTCCTATGCTTTCCCAAGAACAGGCACAAACCCCGCAGCAGGTTTATTTGAGCCTCAGGTGGCCCTGGGCTCCCTCCACCTAATCCA
1113▶ H I Q T P M L S Q E Q A Q P P Q Q G L F Q P Q V A L G S L P P N P
4001 ATGCCTCAAAGCCAACAAGGAACCATGTTCCAGTCACAGCACTCAATAGTTGCCATGCAGAGTAACTCTCCATCCAGGAACAGCAGCAGCAGCAAC
1147▶ M P Q S Q Q G T M F Q S Q H S I V A M Q S N S P S Q E Q Q Q Q Q

XcmI (4157)

NcoI (4157)

4101 AGCAGCAGCAACAGCAGCAGCAACAACAACAGAGCATTATTCAGTAATCAGAATACCATGGCTACAATGGCGTCTCCAAGCAACCACCACCAACAT
1180▶ Q Q Q Q Q Q Q Q Q Q S I L F S N Q N T M A T M A S P K Q P P P N M
4201 GATATTCAACCCAAATCAAAATCCAATGGCTAATCAGGAGCAACAGAACCAGTCAATTTTTACCAACAAGTAACATGGCCCAATGAATCAAGAGCAA
1213▶ I F N P N Q N P M A N Q E Q Q N Q S I F H Q Q S N M A P M N Q E Q

EcoO109I (4350)

4301 CAGCCCATGCAATTTTCAGAGTCAGTCCACAGTTTCTCCTCACTTCAGAACCAGGTCCTACCCAGTCGGAATCATCACAGACCCCTTGTTCATAGCTCTC
1247▶ Q P M Q F Q S Q S T V S S L Q N P Q G P T Q S E S S Q T P L F H S S

BstEII (4421)

4401 CTCAGATTGAGTTGGTACAAGGGTACCTAGTTCTCAAGAGCAGCAAGTAACTCTCTTCTTATCTCCAGCATCCATGTCTGCCTTGCAGACCAGTATAAA
1280▶ P Q I Q L V Q G T S P S S Q E Q Q V T L F L S P A S M S A L Q T S I N
4501 TCAACAAGATATGCAACAGTCTCTCTTTTATTCCCTCAGAACAACATGCCTGGAATTCAGGAGCCACATCTTGCCTCAACCACAGGCTACTTTATTT
1313▶ Q Q D M Q Q S P L Y S P Q N N M P G I Q G A T S S P Q P Q A T L F
4601 CACAACACAGCAGGAGGCACAATGAACCAACTGCAGAATCTCCTGGCTCATCTCAGCAGACATCAGGAATGTTCTTATTTGGCATTCAAATAACTGTA
1347▶ H N T A G G T M N Q L Q N S P G S S Q Q T S G M F L F G I Q N N C

MseI (4745)

4701 GTCAGCTTTAACTCTGGACCAGCTACATTGCCTGATCAGTTGATGGCCATAAGTCAGCAGGCCCAACCACAAAACGAGGGCCAGCCACCTGTGACAAC
1380▶ S Q L L T S G P A T L P D Q L M A I S Q P G Q P Q N E G Q P P V T T
4801 ACTTCTTCTCAGCAAATGCCAGAGAATCTCCACTGGCATCTCTATAAACCAACAGAAATCGAAAAGATTGATTTGCTTGTTCATTGCAAAAC
1413▶ L L S Q Q M P E N S P L A S S I N T N Q N I E K I D L L V S L Q N

MseI (4942)

NheI (4936)

4901 CAAGGGAACAACCTTGACTGGCTCCTTTAACTGGATGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAACCACAACCTAGAATGCAGT
1447▶ Q G N N L T G S F •

MfeI (5085)

5001 GAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAAATTGCATTCATT
5101 TATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTTACAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTT
5201 AACCTCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTC
5301 ACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTTGAAGTACTCTTCAATTTCTTTATGTTTAAATGCACTGACCTCCACATTCCCT

Swal (5423)

EcoO109I (5484)

5401 TTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATAT
5501 CCCCCAGTTTAGTGTGACTTAGGGAACAAAGAACCTTTAATAGAAATTTGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAG
5601 GGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATG
134▶ 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

BstXI (5713)

5701 CCACAGGGGTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAG
100▶ 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 A

StuI (5848)

5801 CAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGC
67▶ S G I A I A F A C V T V R G I Y A F I H V A S I I F G T K T R I A

5901 CGCCCCGACATGGTGCTTGTTCCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTC
 34 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

AseI (6056)

6001 ATGATGGCCCTCCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTG
 0 M

6100 ACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAG

SpeI (6211)

6200 TCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATG

SnaBI (6339)

6299 TACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATGTACTGGGCATAATGCC

NdeI (6444)

6399 AGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCAC

6499 CCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGCGGTCAGCCAGGCGGG

SdaI (6622) PacI (6630) BspLU11I (6640)

6599 CCATTTACGTAAGTTATGTAACGCCTG 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 C A A A A G G C C A G G A A C C G T A A A A A G G C C G T T G C T G

6697 GCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT

6797 TTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCA

ApaLI (6954)

6897 TAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTTCGCTCCAAGTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGACCGCTGCGCCTTATCC

6997 GGTAACATCTGCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTG

7097 CTACAGAGTTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGT

7197 TGGTAGCTCTTGATCCGGCAAACAAACCACGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGAT

EagI (7390)

PacI (7370) SwaI (7379) NotI (7389)

7297 CCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCG

7397 CAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAA

7497 CTAGCAAAATAGGCTGTCCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA