



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

101 GAGAAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) HindIII (245) EcoNI (287)
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCCTACCTGAGGCC

301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCCTCCCGCCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCTCAACTCTACGTCTTTGTTTCGTTT

KasI (535) **AgeI (552)** **SphI (560)**
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCCTAAGCTGAGATCACCGGTGAGCAGTGCAGTGGGATGGCTGCTGCCCTGCTGTGGGACAG

601 GTGCTGGCTACGGACGGTTATTCGCTCTCAGTACTGGTTCAGTGCAGTGCAGGAGGGCCTGTGCGTCTTTGTAGCCTGCCAAGTCCAGTACCCCAAT
13▶ C L A T D G Y S L S V T G S V T V Q E G L C V F V A C Q V Q Y P N

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

801 AAGCTCAAGTTCGATTCTATCTCATGGGAAAGGAAAATAGTCATAATTGCTCCCTGGACATCAGAGATGCACAGAAAATGCACACAGGACGTAAGTCTTCT
80▶ E A Q G R F Y L M G K E N S H N C S L D I R D A Q K I D T G T Y F F

ScaI (789)
901 CAGATTGGATGGTCTGTGAAATACAGTTTTGAGAAGTCAATGCTCTGTGCTTGTGATAGCTCTGACTGAGGTCCCTAACATCCAAGTCCACATCTACC
113▶ R L D G S V K Y S F Q K S M L S V L V I A L T E V P N I Q V T S T

1001 CTGGTGTCTGGCAATTCTACAAACTGCTCTGTCTGTGCCCTGGCCTGTGAGCAGGGGACACCCCCATCTTTTCTGGATGTCATCTGCCCTCACTT
147▶ L V S G N S T K L L C S V P W A C E Q G T P P I F S W M S S A L T

PshAI (966)
1101 CCCTGGGCCACAGGACCACCTCTCCTCAGAGCTGAACCTCACACCCAGGCTCAGGACAATGGACCAACCTCACCTGCCAGGTGAACCTTACCTGGCAC
180▶ S L G H R T T L S S E L N L T P R P Q D N G T N L T C Q V N L P G T

StuI (1147)
1201 TGGTGTACTGTGGAGAGGACTCAACAGCTCAGTGTCTATGCTCCACAGAAGATGACCATCAGGGTGTCTGGGAGATGACACAGGAACATAAAGT
213▶ G V T V E R T Q Q L S V I Y A P Q K M T I R V S W G D D T G T K V

1301 CTGCAGAGTGGAGCATCTCTGCAGATCCAAGAGGGTGTGCTCCAGCCTGGTCTGTATGGCTGACAGCAATCCCTGCGTGTGAGTTGGGAGGCC
247▶ L Q S G A S L Q I Q E G E S L S L V C M A D S N P P A V L S W E R

SacII (1430)
1401 CGACCCAAAAGCCCTTCCAGCTCTCTACCCCGCGAACTGCAACTGCCACGGCGGAGTTGGAGGACCAGGGAAAATACATCTGCCAAGCTCAGAACAG
280▶ P T Q K P F Q L S T P A E L Q L P R A E L E D Q G K Y I C Q A Q N S

SandI (1557)
1501 TCAGGGTGTCTCAGACAGCCTCTGTGAGCCTCAGCATAAGGAGCCTTCTACAGCTGCTGGGACCCTCCTGCTCCTCGAGGGTCAAGGTCTGCACTGCAGC
313▶ Q G A Q T A S V S L S I R S L L Q L L G P S C S F E G Q G L H C S

1601 TGTCCTCAGAGCCTGGCTGCCCTCCCTGCCTGGCGGCTGGGGAGGGGTGCTGGAGGGAAACAGCAGCAATGGCTCCTTACGTTAAGTCCA
347▶ C S S R A W P A P S L R W R L G E G V L E G N S S N G S F T V K S

EcoRI (1744) **HindIII (1775)**
1701 GCTCCGACAGGACAGTGGGCAACAGCTCCCTGATCCTCAGCATGGAATTCAGCTCTAACCACAGACTCAGCTGTGAAGCTTGGAGTGACAATAGAGTTCA
380▶ S S A G Q W A N S S L I L S M E F S S N H R L S C E A W S D N R V Q

XcmI (1805)
1801 GAGAGCCACTATCTTGTGGTGTACAGGCCAAAGTCTCACAGGCAGGCAATCAGAAACAGTAGAGGAACGGTCTGGGGCCATCTGGGAGCTGGT
413▶ R A T I L L V S G P K V S Q A G K S E T S R G T V L G A I W G A G

1901 CTTATGGCCTTGTGCTGTCTGCCTCTGCCTCATTTTTTACAGTGAAGTCTCAGGAAGAAATCAGCCCTGAAAGTAGCAGTACGAAAGGCAACC
447▶ L M A L L A V C L C L I F F T V K V L R K K S A L K V A A T K G N

XcmI (2051)
2001 ATCTTGCCAGAACCCTGCTCCACCATAAATAGTGAAGCATAAATTCTCAACCAATTGCGTTGGGATACCAATCCAGGGTCAATGAAATGAAACCAGG
480▶ H L A K N P A S T I N S A S I T S S N I A L G Y P I Q G H L N E P G

2101 GTCACAGACCCAGAAAGACAGCCACCCCTTGCACAGTCCAGACACCCAAAAGGATGAGCCTGAATCCACTATGCCTCCCTCTCCTTCCAAGGCCG
513▶ S Q T Q K E Q P P L A T V P D T Q K D E P E L H Y A S L S F Q G P

Bst1107I (2239) **NheI (2297)**
2201 ATGCCCCGAAGCCTCAGAACACGGAGGCTATGAAATCTGTATACACAGAGATCAAGATCCACAAGTGTGAAGTTCTTTGCTCCATGTGCCAGTGCCT
547▶ M P P K P Q N T E A M K S V Y T E I K I H K C •

MscI (2303)
2301 AGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAGAAAATGCTTTATTTGTAATTTGTGATGCTATTG

HpaI (2435) **MfeI (2446)**
2401 CTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAACAACAACATTGCATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGTTTTTTTA

EcoRI (2531)
2501 AAGCAAGTAAAACCTTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAG

2601 GGATGAATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCC

2701 AAGGTTTGAAGTACTCTTCATTCTTTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATT
2801 GCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGGAACC
2901 TTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCC
111 • N R T Y K L P I L E E I T T K V L K G
3001 ATTCACTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCA
121 N M E I L 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
3101 TCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTGACCC
87 D S Y P H 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
3201 TGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTGGTCTGATGGCCGCCCGACATGGTCTTGTGCTCATAGAGCATGGT
54 G I Y A E 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
3301 GATCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCC
21 I K E T A V E V L E L D Q Q S I N F T K M
3401 GATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCAC
3501 CGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGT
3601 CAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAA
3701 TACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTA
3801 CTTGGCATATGATACACTTGATGTACTGCCAAGTGGGAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGG
3901 GAACATACGTCATTATTGACGTCAATGGGCGGGGTCTGTTGGCGGTGACCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAAATTAAG
4001 AACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAA
4101 ATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCT
4201 GCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGATAGTATCTCAGTTCGGTGTAGGTCGTTCCG
4301 TCCAAGCTGGGCTGTGTGCACGAACCCCCGTTAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTAT
4401 CGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAG
4501 AAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGCAAAACAAACCACCGCTGGTAGCGGT
4601 GGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAA
4701 ACTCACGTTAAGGGATTTGGTTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTT
4801 TTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAA
4901 CATTCTCTATCGAA