



PvuI (7) SgfI (6) EcoNI (96)
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) PvuII (239) EcoNI (287) Bsu36I (291)
201 GTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACCGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCGGTTGAGTGCCTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441) NaeI (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGTCTTTGTTTCGTTT

KasI (535) AgeI (552) NcoI (568)
501 TCTGTTTGGCGGCTTACAGATCCAAGCTGTGACCGCGGCTACCTGAGATCACCGGTAGGAGGGCCACCATGGTGTTCCTAATGTGGACTGAAGAGA
601 CAAATTCCTATCCTTTTAAACATAATCCTAATTTCAAACCTCCTTGGGGCTAGATGGTTTCTAAAACTGCGCTGTGATGTCACTCTGGATGTTCCAA
110 Q I L I L F N I I L I S K L L G A R W F P K T L P C D V T L D V P
701 AGAACCATGTGATCGTGGACTGCACAGACAAGCATTTGACAGAAATCCTGGAGGTATCCACGAAACACCACGAACCTCACCTCACCATTAAACACAT
440 K N H V I V D C T D K H L T E I P G G I P T N T T N L T L T I N H I

ClaI (849) BspLU11I
801 ACCAGACATCTCCCAGCGTCTTTACAGACTGGACCATCTGGTAGAGATCGATTTCCAGATGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAC
770 P D I S P A S F H R L D H L V E I D F R C N C V P I P L G S K N N
901 ATGTGCATCAAGAGGCTGCAGATTAACCCAGAAGCTTGTGGACTCACTTATTTAAATCCCTTACCTGGATGAAAACAGCTACTAGAGATACCGC
110 M C I K R L Q I K P R S F S G L T Y L K S L Y L D G N Q L L E I P
1001 AGGGCCTCCCGCTAGCTTACAGCTTCTCAGCCTTGGGCCAAACATCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCCAAATAGAAATACT
140 Q G L P P S L Q L L S L E A N N I F S I R K E N L T E L A N I E I L
1101 CTACCTGGGCAAAAAGCTGTTATTATCGAAATCCTGTTATGTTTCAATAGAGAAAAGATGCCTTCTAACTTGACAAAAGTTAAAAGTCTCTCC
170 Y L G Q N C Y Y R N P C Y V S Y S I E K D A F L N L T K L K V L S
1201 CTGAAAGATAACAATGTCACAGCGCTCCCTACTGTTTTGCCATCTACTTTAACAGAATATATCTCTACAACAACATGATTGCAAAAACCAAGAAGATG
210 L K D N N V T A V P T V L P S T L T E L Y L Y N N M I A K I Q E D
1301 ATTTAATAAACCCTCAACCAATTCAAAATCTTGACCTAAGTGGAAATGGCCCTGTTGTTATAATGCCCCATTTCTGTGGCCGCTGTAAAAATAATC
240 D F N N L N Q L Q I L D L S G N C P R C Y N A P F P C A P F C A K N N S
1401 TCCCTACAGATCCCTGTAATGCTTTTGTGCGCTGACAGAATTAAGTTTTACGTCTACACAGTAACTCTCTTACAGCATGTGCCCAAGATGGTTT
270 P L Q I P V N A F D A L T E L K V L R L H S N S L Q H V P P R W F
1501 AAGAACATCAACAACTCCAGGAAGTGGATCTGTCCAAAACCTTCTGGCCAAAAGAAATTTGGGGATGTAAATTTCTGCATTTTCTCCCGCCTCATCC
310 K N I N K L Q E L D L S Q N F L A K E I G D A K F L H F L P S L I
1601 AATTGGATCTGTCTTCAATTTTGAATTCAGGCTATCQGTGCATGATGAATCTATCAAGCAATTTTCTTCACTGAAAAGCTGAAAATTTGCGGAT
340 Q L D L S F N F E L Q V Y R A S M N L S Q A F S S L K L K I L R I
1701 CAGAGGATATGCTTTAAAGAGTTGAAAAGCTTAACTCTCGCCATTACATAATCTTCAAAATCTTGAAGTTCTTGATCTTGGCACTAACTTTATAAAA
370 R G Y V F K E L K S F N L S P L H N L Q N L E V L D L G T N F I K

PmeI (1817) BglII (1847)
1801 ATTGCTAACCTCAGCATGTTTAAACAATTTAAAGACTGAAAGCATAGATCTTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCT
410 I A N L S M F K Q F K R L K V I D L S V N K I S P S G D S S E V G
1901 TCTGCTCAATGCCAGAAGCTTGTGAGAAAGTTATGAACCCAGGTCCTGGAACAATTACATTATTTTCAGATATGATAAGTATGCAAGGAGTTGCAGATT
440 F C S N A R T S V E S Y E P Q V L E Q L H Y F R Y D K Y A R S C R F
2001 CAAAAACAAGAGGCTTCTTTCATGCTGTTAATGAAAGCTGCTACAAGTATGGGAGACCTTGGATCAAGTAAAAATAGTATATTTTTGTCAAGTCC
470 K N K E A S F M S V N E S C Y K Y G Q T L D L S K N S I F F V K S

XcmI (2159)
2101 TCTGATTTTCAGCATCTTTCTTCTCAATGCCTGAATCTGTGAGGAAATCTCATTAGCCAACTCTTAATGGCAGTGAATCCAACCTTTAGCAGAGC
510 S D F Q H L S F L K C L N L S G N L I S Q T L N G S E F Q P L A E

XmnI (2256)
2201 TGAGATATTTGGACTTCTCAACAACCGGCTTGATTTACTCCATTCAACAGCATTTGAAAGAGCTTCACAAACCTGGAAGTTCTGGATATAAGCAGTAATAG
540 L R Y L D F S N N R L D L L H S T A F E E L H K L E V L D I S S N S

NdeI (2328)
2301 CCATTATTTTCAATCAGAAGGAATTAACCTTACCAAGAACCTAAAGGTTCTGCAGAACTGATGATGAACGACAATGACATCTCTTCC
570 H Y F Q S E G I T H M L N F T K N L K V L Q K L M M N D N D I S S

XcmI (2404) NcoI (2413)
2401 TCCACCAGCAGGACCATGGAGAGTGAAGTCTTTAGAATCTGGAATTCAGAGGAAATCACTTAGATGTTTTATGGAGAGAAGGTGATAACAGATACTTAC
610 S T S R T M E S E S L R T L E F R G N H L D V L W R E G D N R Y L
2501 AATTATTCAGAATCTGCTAAAATTAGAGGAATTAGACATCTCAAAAATCCCTAAGTTTCTTGCCTTCTGGAGTTTTGATGGTATGCCTCAAATCT
640 Q L F K N L L K L E E L D I S K N S L S F L P S G V F D G M P P N L
2601 AAAGAATCTCTTTGGCCAAAATGGGCTCAAATCTTTCAGTTGGAAGAAATCCAGTGTCTAAAGAACCTGGAACCTTTGGACCTCAGCCACAACCAA
670 K N L S L A K N G L K S F S W K K L Q C L K N L E T L D L S H N Q

PshAI (2702)
2701 CTGACCCTGTCCTGAGAGATTATCCAACCTGTTCCAGAAGCCTCAAGAATCTGATTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTAC
710 L T T V P E R L S N C S R S L K N L I L K N N Q I R S L T K Y F L

EcoRV (2818)
2801 AAGATGCCTTCCAGTTGGGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCAGAAAATGTCCTCAACAATCTGAAGAT
740 Q D A F Q L R Y L D L S S N K I Q M I Q K T S F P E N V L N N L K M

2901 GTTGCTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGGTTTGTCTGGTGGGTTAACCATACGGAGGTGACTATTCTTACCTGGCCACA
777▶ L L L H H N R F L C T C D A V W F V W V N H T E V T I P Y L A T

3001 GATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGCCAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCT
811▶ D V T C V G P G A H K G Q S V I S L D L Y T C E L D L T N L I L F

3101 CACTTTCCATATCTGTATCTCTCTTTCTCATGGTGTGATGACAGCAAGTACCTCTATTTCTGGGATGTGGGTATATAACCATTTCTGTAAGGCCAA
844▶ S L S I S V S L F L M V M M T A S H L Y F W D V W Y I Y H F C K A K

3201 GATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTATGACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCT
877▶ I K G Y Q R L I S P D C C Y D A F I V Y D T K D P A V T E W V L A

3301 GAGCTGGTGGCCAAACTGGAAGACCCAAGAGAGAAACATTTTAAATTTATGCTCGAGGAAGGACTGTTTACCAGGGCAGCCAGTTCTGGAAAACTTT
911▶ E L V A K L E D P R E K H F N L C L E E R D W L P G Q P V L E N L

3401 CCCAGAGCATAACAGCTTAGCAAAAAGACAGTGTGTGTGATGACAGACAAGTATGCAAAGACTGAAAATTTAAGATAGCATTTTACTTGTCCCATCAGAG
944▶ 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 F Y L S H Q R

3501 GCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTCAGAAAGTCCAAGTTCCTCCAGCTCCGAAAAGGCTCTGTGGGAGT
977▶ L M D E K V D V I I L I F L E K P F Q K S K F L Q L R K R L C G S

3601 TCTGTCCTTGAGTGGCCAAACCCGCAAGCTCACCCATACTTCTGGCAGTGTCTAAAGAAGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGG
1011▶ 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 H V A Y S Q

3701 TGTTCAAGGAAACGGTCTAGCCCTTCTTTGCAAAACACAACCTGCCTAGTTTACCAAGGAGAGGCTGGCTGTTTCTAGCTGGCCAGACATGATAAGATA
1044▶ V F K E T V •

3801 CATTGATGAGTTTGGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGC

3901 TGCAATAAACAAGTTAACAAACAATTGCATTCATTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAAT

4001 GTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCA

4101 GGGCTGTGGCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTGCTTTCATT

4201 TCTTTATGTTTTAAATGCACTGACCTCCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTA

4301 TTAGGCAGAAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGAACCTTTAATAGAAATTTGGACAGCAAG

4401 AAAGCGAGCTTCTAGCTTTAGTTCTCTGGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGC
144▶ • N R T Y K L P I L E E I T T K V L K G N M E I L V F C

4501 AGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGC
113▶ 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 R V A

4601 CACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGG
80▶ 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 I H

4701 ACAGCAGAGATGATCTCCCAAGTCTGGTCTGATGGCCGCCGACATGGTCTTGTGTCCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCA
46▶ 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 V E V

4801 CCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCTATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAAT
13▶ L E L D Q Q S I N F T K M

4901 TGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATT

5000 GCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTGATTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAA

5099 ATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGT

5199 AGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGA

5299 TGTACTGCCAAGTGGCAGTTTACCCTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACG

5399 TCAATGGGCGGGGTCTGTTGGCGGTGACGACAGCGGGCCATTTACCGTAAGTTATGTAACGCC T G 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 G G C C
5497 AGCAAAAGGCCAGGAACCGTAAAAAGCCCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAG

5597 GTGGCGAAACCCGACAGGACTATAAAGATACCAGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTG

5697 TCCGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGC
 5797 ACGAACCCCGTTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC
 5897 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATC
 5997 TCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGC
 6097 AGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACTCACGTTAAGGGATTTT

EagI (6228)

PacI (6208) SwaI (6217)

NotI (6227)

6197 GGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCGCAATAAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTTGTGTGAATCGTAACTAA
 6297 CATACGCTCTCCATCAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA