



PvuI (7)
SgfI (6)
1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAAGTGGCGGGGTAAGTGGAAAGTGTGCTGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) **PvuII (239)** **Bsu36I (291)**
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGGCC
301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCCGTCTAGGTAAGTTAAAGCTCAGGTCGAGACC

NgoMIV (441)
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

KasI (535) **AgeI (552)** **BspHI (560)**
501 TCTGTTTGTGCGCGGTTACAGATCCAAGCTGTGACCGCGCCCTACCTGAGATCACCGGTCATCATGAAGTTATATGATTTTCTGGTTAACACTGGAAGTAC
1 M K L Y V F L V N T G T T

SphI (648) **DraIII (685)**
601 TCTAACATTTGACTGAACTTACAGTGCAAACTGTGGCAGACCTTAAGCATGCCATTCAAAGCAAATACAAGATTGCTATTCAACACCGAGTGTGGTG
13 L T F D T E L T V Q T V A D L K H A I Q S K Y K I A I Q H Q V L V

NsiI (714)
701 GTCATGGAGGAGAATGCATGGCTGCAGATCGAAGAGTGTGTACTACAGTGTGGGACGGATACAAATCCAATTTTTCTTTTAAACAAAGAAATGATCT
47 V N G G E C M A A D R R V C T Y S A G T D T N P I F L F N K E M I
801 TATGTGATCGTCCACCTGCTATTCTAAACTACCTTTTCGACAGAAAATGACATGGAATAAAAAGTTGAAGAATCTCTTATGATGCCTGCAGTTTTCA
80 L C D R P P A I P K T T F S T E N D M E I K V E E S L M M P A V F H
901 TACTGTTGCTTCAAGGACACAGCTTGCATTGGAATGTGAAGTTGCCAAGAAACTTTGTTCTTTTGTGAAGTCTTGTACATGATGAACATCTTCAA
113 T V A S R T Q L A L E M Y E V A K K L C S F C E G L V H D E H L Q
1001 CACCAAGGCTGGGCTGCAATCATGGCCAACCTGGAGGACTGTTCAAATTCATACCAAAAAGCTACTTTTCAAGTTTGAAGTATTTATTCAAATATCTGC
147 H Q G W A A I M A N L E D C S N S Y Q K L L F K F E S I Y S N Y L
1101 AGTCCATAGAAGACATCAAGTTAAAACCTACTCATTAGGAACTGCAGTTTCAGTAATGGCCAAAGATCCACTGTTGGAGTGCCTAACACAGACATAGTTA
180 Q S I E D I K L K L T H L G T A V S V M A K I P L L E C L T R H S Y
1201 CAGAGAATGTTGGGAAGACTGGATTCTTACCTGAACATGAAGACTCAGAAAAAGCTGAGAGAAAAGATCCACTGAACTGGTGTCTCTCCTGATATG
213 R E C L G R L D S L P E H E D S E K A E T K R S T E L V L S P D M
1301 CCTAGAACAACCTAACGAATCTTTGTTAACTCATTTCCTCAAGTCAAGTGAACATGTGTCCCAGATACCGCAGATGCTGAAAGTGGCAAGAAATAGGG
247 P R T T N E S L L T S F P K S V E H V S P D T A D A E S G K E I R

ScaI (1411)
1401 AATCTTGTCAAAGTACTGTTTCATCAGCAAGATGAAACTACGATTGACTAAAGATGGTGATCTGCCCTTTTTAATGTCTCTTTGTTAGACTGGATAAA
280 E S C Q S T V H Q Q D E T T I D T K D G D L P F F N V S L L D W I N
1501 TGTTCAAGATAGACCTAATGATGTGGAATCTTTGGTCAGGAAGTCTTTGATTCTATGAGCAGGCTTGATCCAAGGATTATTCGACCATTTATAGCAGAA
313 V Q D R P N D V E S L V R K C F D S M S R L D P R I I R P F I A E

NheI (1690)
1601 TGCCGTCAAACCTATTGCCAACTTGATAATCAGAATATGAAAGCCATTAAGGACTTGAAGATCGGCTCTACGCCCTGGACCAGATGATTGCTAGCTGTG
347 C R Q T I A K L D N Q N M K A I K G L E D R L Y A L D Q M I A S C

NsiI (1775)
1701 GCCGACTGGTGAATGAACAGAAAGACTTGGCTCAGGGATTTTTAGCTAATCAGAAGAGAGCTGAAAACCTAAAGGATGCATCTGTATTACCTGATTTATG
380 G R L V N E Q K E L A Q G F L A N Q K R A E N L K D A S V L P D L C
1801 CCTGAGTCACGCAATCAGTTGATGATTATGTTGCAAAATCATAGAAAACCTGTTAGATATTAAGCAGAAGTGTACCCTGCCAAAACAAGAACTAGCAAA
413 L S H A N Q L M I M L Q N H R K L L D I K Q K C T T A K C Q C L E L A N
1901 AACCTACATGTAGCTGAGTGAAGTGTGCTTTGTAATGCTTACGTGATCAGATGAGAGAAAGTTACAAGCTTTGCTCCGCTCGTAATAGAGCTGT
447 N L H V R L K W C C F V M L H A D Q D G E K L Q A L L R L V I E L
2001 TAGAAAGAGTCAAATGTTGAAGTCTTAGTACAGTTCCTCAGATGTACTGCTTAGCTGTTGTTGAGGTTGAAGAAGAAAATGTTCAAAAACCTA
480 L E R V K I V E A L S T V P Q M Y C L A V V E V V R R K M F I K H Y
2101 CAGGGAGTGGGCTGGTCTTAGTCAAAGTGAAGAGATTATGAAGCAGAAAACTAAAAGGGAAATCCTTTGGGAAATATTAGGAAGTCTTTT
513 R E W A G A L V K D G K R L Y E A E K S K R E S F G K L F R K S F
2201 TTAAGAAATCGTCTGTTTAGGGGACTGGACTCTGGCCCCCTTCTTTTGTACTCAAAAAGCTCGAAAAGTTGACTGTGAACCTCCAGATATTTCAATA
547 L R N R L F R G L D S W P P S F C T Q K P R K F D C E L P D I S L

Bsu36I (2349)
2301 AAGATTTACAGTTTCTGCAATCATTGTCCTTCGGAAGTTCAGCCATTCCTCAGGGTTCCTTACTTTGTGACTTTGAACTCTACACCAGCATGTACT
580 K D L Q F L Q S F C P S E V Q P F L R V P L L C D F E P L H Q H V L

BglII (2460)
2401 TGCTCTACATAATTTGGTAAAAGCAGCACAAAGTTGGATGAAATGTCACAGACCATTACAGATCTACTGAGTGAACAAAAGGCATCTGTGAGTCAGACA
613 A L H N L V K A A Q S L D E M S Q T I T D L L S E Q K A S V S Q T

BamHI (2592)
2501 TCCCCACAGTCTGCTTCTTACCAAGGATGAAAGTACAGCAGGAATTACAACACTACTCCTCACCGAGAAGTCCCTCCACCCTGACTGTTCCAGGATCCCT
647 S P Q S A S S P R M E S T A G I T T T T S P R T P P P L T V Q D P
2601 TATGTCCTGCAGTTTGTCCCTTAGAAGAATTATCTCCAGATAGTATTGATGCACATACGTTTGTATTTGAAACTATTCCTCCATCCAAACATAGAACAGAC
680 L C P A V C P L E E L S P D S I D A H T F D F E T I P H P N I E Q T
2701 TATTCACCAAGTTCTTTAGACTTGGATTCTTAGCAGAAAGTCTGAAATCAGATTTTATGCTGCTGTGAATGAGTTTGAATAGAAGAAAATTTGCG
713 I H Q V S L D L D S L A E S P E S D F M S A V N E F V I E E N L S
2801 TCTCCTAATCCTATAAGTATCCACAAAGCCAGAAATGATGGTGAATCACTTTATTCATCAGTTATCAATGCGATAGACAGTAGACGAATGCAGGATA
747 S P N P I S D P Q S P E M M V E S L Y S S V I N A I D S R R M Q D

Bst1107I (2993)
2901 CAAATGTATGTGGTAAGGAGATTTTGGAGATCATACTTCTGTAATGTCAGTTGAAAAGATGTAGAGTTGTTGCCCAAGACTCTCACTTCAGTATACA
780 T N V C G K E D F G D H T S L N V Q L E R C R V V A Q D S H F S I Q

3001 ACCATTAAGGAAGACCTTTGCCACTTTAGAACATTTGTACAAAAAGAACAGTGTGACTTCTCAAATTCATTAATGTACAGCAGTAGAATAAGAAAC
813▶ T I K E D L C H F R T F V Q K E Q C D F S N S L K C T A V E I R N
3101 ATTATTGAAAAAGTAAATGTTCTCTGGAATAAACAATAAAGAAAAACATCAAAAAAGAACTACTGTCTTTAAAAATGAATATGAAGGTAACCTTGACG
847▶ I I E K V K C S L E I T L K E K H Q K E L L S L K N E Y E G K L D
3201 GACTAATAAAGGAACTGAAGAGAATGAAAAACAAATTAATAAATTAAGGGAGAGTTAGTATGCCTTGAGGAGGTTTTACAAAAATAAGATAATGAATT
880▶ G L I K E T E E N E N K I K K L K G E L V C L E E V L Q N K D N E F
3301 TGCTTTGGTTAAACATGAAAAAGAGCTGTAATCTGCCTGCAGAAATGAAAAGGATCAGAAGTTGTTAGAGATGAAAAATAATGCACCTCAAAATTTG
913▶ A L V K H E K E A V I C L Q N E K D Q K L L E M E N I M H S Q N C
3401 GAAATTAAGAAGCTGAAGCAGTACGAGAAATAGTGTAGAAGACTTAAAAAGCTCCATGTTGAAAAATGATGAGAAGTTACAGTTATTGAGGGCAGAAC
947▶ E I K E L K Q S R E I V L E D L K K L H V E N D E K L Q L L R A E
3501 TTCAGTCTTGGAGCAAAGTCATCTAAAGGAATTAGAGGACACACTTCAGGTTAGGCACATACAAGAGTTTGAGAAGGTTATGACAGACCACAGAGTTTC
980▶ L Q S L E Q S H L K E L E D T L Q V R H I Q E F E K V M T D H R V S
3601 TTTGGAGGAATTAATAAGGAAACCAAAATTAATAACAAGAACTCTCAATGAAATTCAGGAAAAAGAAAAACAGTTACAGGAATTA
1013▶ L E E L K K E N Q Q I I N Q I Q E S H A E I I Q E K E K Q L Q E L
3701 AAACCTCAAGGTTTCTGATTTGTGACAGACGAGATGCAAGTTAGAGTTGAACTTTCGTTGAAGGAAGCAGAACTGATGAAATAAAAATTTTGTGGAAG
1047▶ K L K V S D L S D T R C K L E V E L A L K E A E T D E I K I L L E
3801 AAAGCAGAGCCCAGCAGAAGGAGACCTTGAAATCTCTTGAACAAGAGACAGAAAAATTTGAGAACAGAAATAGTAAACTCAACAAAAAGATTTCAGGA
1080▶ E S R A Q Q K E T L K S L L E Q E T E N L R T E I S K L N Q K I Q D
3901 TAATAATGAAAAATTACAGTGGCTTAGCAGAGCTAAGAATTAATGACAATTTGAAAAAGATCAGTGTATTCCGAGTTAATTAGTAGACATGAAGAA
1113▶ N N E N Y Q V G L A E L R T L M T I E K D Q C I S E L I S R H E E
4001 GAATCTAATATACTTAAAGCTGAATTAACAAAGTAACATCTTGCATAACCAAGCATTGAAATAGAAAAAACCTAAAAGAACAAATAATTGAACTGC
1147▶ E S N I L K A E L N K V T S L H N Q A F E I E K N L K E Q I I E L
4101 AGAGTAAATTTGAGTTCAGAAATGAGTCTTGAAGACAAAAAGATGAAAAATTTACCAACAAGAAGAGAAATACGAAGCTATTATCCAGAACCTTGA
1180▶ Q S K L D S E L S A L E R Q K D E K I T Q Q E E K Y E A I I Q N L E
4201 GAAAGACAGACAAAAATGGTCAAGCAGCCAGGAGCAAGACAGAGAACAGTTAATTGAGAAGCTTAATTGTAAAAAGATGAAGCTATTGAGACTGCCCTA
1213▶ K D R Q K L V S S Q E Q D R E Q L I Q K L N C E K D E A I Q T A L
4301 AAAGAATTTAAATTTGGAGAGAGAAGTTGTTGAGAAAGAGTTATTAGAAAAAGTTAAACATCTTGAAGATCAAATAGCAAAAAAGTCTGCCATTGACTCTA
1247▶ K E F K L E R E V V E K E L L E K V K H L E N Q I A K S P A I D S

XbaI (4467)

4401 CCAGAGGAGATCTTCAAGCTTAGTTGCTGAACCTCAAGAAAAGCTTCAGGAAGAAAAAGCTAAGTTTCTAGAACAACCTGAAGAGCAAGAAAAAGAAA
1280▶ T R G D S S S L V A E L Q E K L Q E E K A K F L E Q L E E Q E K R K

BstBI (4520)

4501 GAATGAAGAAATGCAAAATGTTGCAACATCTTTGATTGCGGAACAACAGACCAATTTTAACTGTTTTAACAAGAGAGAAAAATGAGAAAAGAAAAACATA
1313▶ N E E M Q N V R T S L I A E Q Q T N F N T V L T R E K M R K E N I
4601 ATAAATGATCTTAGTGATAAGTTGAAAAGTACAATGCAGCAACAAGAAGGGGATAAGATTTGATAGAGTCACTTTCTGAAGATCGAGCTCGTTTGTCTG
1347▶ I N D L S D K L K S T M Q Q Q E R D K D L I E S L S E D R A R L L

FspI (4736)

ApaLI (4799)

4701 AGGAAAAGAAAAAGCTTGAAGAAGAAGTCAAGTAAAGTTGCGCAGTAGCAGTTTTGTTCTTCCACATATGTAGCTACAGCCCCAGAACTTTATGGAGCTTG
1380▶ E E K K K L E E E V S K L R S S S F V P S P Y V A T A P E L Y G A C
4801 TGCACCTGAACTCCCAGGTGAATCAGATAGATCCGCTGTGAAAACAGCAGATGAAGGAAGAGTGGATTGAGCAATGGAGACAAGCATGATGTCTGTACAA
1413▶ A P E L P G E S D R S A V E T A D E G R V D S A M E T S M M S V Q

SspI (4903)

PshAI (4997)

4901 GAAAAATTCATATGTTGTCTGAAGAAAAACAGCGATAATGCTGTTAGAACGAACATTGCAATTGAAAGAAGAAGAAAAATAACGGTTAAATCAAAGAC
1447▶ E N I H M L S E E K Q R I M L L E R T L Q L K E E E N K R L N Q R
5001 TGATGTCTCAGAGCATGTCTTCAGTATCTTCAAGGCATCTGAAAAGATAGCTATTAGAGATTTTCAGGTGGGAGATTTGGTACTCATCATCCTAGACGA
1480▶ L M S Q S M S S V S S R H S E K I A I R D F Q V G D L V L I I L D E
5101 ACGCCATGACAATTATGTGTTATTTACTGTTAGTCTACTTTATATTTTCTACATTGAGAGTCTTACCTGCCTGGATCTCAAACAGGTGAGGGTGTCT
1513▶ R H D N Y V L F T V S P T L Y F L H S E S L P A L D L K P G E G A

XbaI (5209)

5201 TCAGGTGCATCTAGAAGACCCTGGTACTCGGAAAAGTAAAGGAAAAGAACTGTCAAGCCAAAAAGGCACAAAACAGATTTAAAGTTCCTTTGGGGA
1547▶ S G A S R R P W V L G K V M E K E Y C Q A K K A Q N R F K V P L G

AvrII (5376)

5301 CAAAGTTTACAGAGTGAAGCCGTATCATGGAATAAGAAAGTATAACTTATGGACAAAATTAATACATTCTATGACCTAGGACTAGCTGGCCAGACATG
1580▶ T K F Y R V K A V S W N K K V •

5401 ATAAGATACATTGATGAGTTGGACAAACCACAACACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCA
5501 TTATAAGCTGCAATAAACAGTTAAACAACAACAAATTCATTCTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCT

EcoRI (5616)

5601 CTACAAATGTGGTATGGAATCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCAT
5701 AGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGAAGTACG

SspI (5855)

5801 TCTTCATTTCTTTATGTTTTAAATGCACCTGCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAAATAAT

EcoO109I (5930)

5901 GTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTTG
6001 ACAGCAAGAAAGCAGCTTCTAGCTTTAGTTCTGGTGTACTTGAAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCATCTCAATGAG
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
6101 CACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGC
116▶ 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 (6294)

6201 CTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCT
82 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
6301 CAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGCTCTGATGGCCGCCGACATGGTGCTTGTTCCTCATAGAGCATGGTGATCTTCTCAGTGGC
49 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 (6444)

6401 GACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCTGATTATACTATGCCGATATACTATGCCGA
16 V E V L E L D Q Q S I N F T K M
6501 TGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCCTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCG

SpeI (6657)

6601 CCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCCTTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATGGGGTGGAGAC

SnaBI (6785)

6701 TTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTG
6801 CCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATA
6901 ACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTA

PacI (7076)

SdaI (7068)

7001 TTGACGTCAATGGGCGGGGTCGTTGGCGGTGACCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAAATAAGAACATGTGAGCAAAA
7101 GGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTC
7201 AGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCTGCCGTTACCGGATA
7301 CCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGT

ApaLI (7400)

7401 GTGCACGAACCCCCGTTGACCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAG
7501 CCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGCCTAACTACGGCTACACTAGAAGAACAGTATTTGG
7601 TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGC
7701 AAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGA

EagI (7836)

NotI (7835)

PacI (7816)

7801 TTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTTGTGTAATCGTAA
7901 CTAACATACGCTCTCCATCAAAAACAAACGAAACAAACAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGACAGGTGCCAGAACATTTCTCTATCGAA