



PvuI (7)
SgfI (6) 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA
EcoNI (96)
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTACTGGCTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

Psp1406I (203) 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGGCCGAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCTTACAGCGCCCGCCCTACCTGAGGCC
EcoNI (287)
301 GCCATCCACGCGGTTGAGTCGCGTTTCTGCCGCTCCCGCCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC

NgoMIV (441) 401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGTCTTTGTTTCGTTT

KasI (535) 501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCATCATGAGTGGGCTCTACAGGATCCTGGTCCAGCTGGAACA
AgeI (552) 1 M S G L Y R I L V Q L E Q
BamHI (578)
601 AAGTCCATATGTGAAGACCGTGCCTTTGAACATGAGAAGGGATTTTTTTTTCTGTTGTAACCTGGATGCCTAAGACAGTCAAGATGAATGGGAGCAGC
13 S P Y V K T V P L N M R R D F F F L V V T W M P K T V K M N G S S
701 TTTGTGCCATCTCTACAGCTCCTGCTCATGTTAGTAGGATTTCTCTGCCGCTGTGGCAGAGACATATGGGTTCAACAAGTGACACAGTATGAATTTG
47 F V P S L Q L L L M L V G F S L P P V A E T Y G F N K C T Q Y E F

DraIII (805) 801 ATATTCACCATGTGCTCTGCATTAGGAAGAAGATCACCAACTTGACAGAGGCCATTAGTGACATACCTAGATATACTACTCACCTAACCTCACACACAA
80 D I H H V L C I R K K I T N L T E A I S D I P R Y T T H L N L T H N
901 CGAAATCAAGTCCCTCCCTCCCTGGAGTTTACCAATCTGTCTGCTTGGTGGACTTGAGACTAGAGTGAAGTCAATTTGGAAGATCGACGAAAGTGCC
113 E I Q V L P P W S F T N L S A L V D L R L E W N S I W K I D E G A
1001 TTTAGGGGACTTGAAAATTTGACTCTGCTGAATTTAGTGGAAAATAAGATTCAAAGTGTGAATAACTCTTTGAGGGCCTGTCCAGCCTGAAGACCTGC
147 F R G L E N L T L L N L V E N K I Q S V N N S F E G L S S L K T L
1101 TCCTGAGCCATAATCAGATTACCATATTACAAAAGACGCCTTCACTCCTTAATCAAATTTGAAATATTTGAGCCTATCTGAAAACAACATTAGCGATTT
180 L L S H N Q I T H I H K D A F T P L I K L K Y L S L S R N N I S D F

AvrII (1288) 1201 TTCTGGTATTCTTGAAGCAGTCCAGCATCTCCATGCCTGGAGCGCCTTGATCTAACTAACACAGCATCATGTAAGTGGACCACAGCCCTAGGTCAGT
213 S G I L E A V Q H L P C L E R L D L T N N S I M Y L D H S P R S L
1301 GTTCTCTGACCCACCTGAGTTTGGAGGAAACAACTAAGGAGTTAAACTCTCTGCTTTGTCACTTACCTAACTAACCAATCTAAGTGTCTCCGGGA
247 V S L T H L S F E G N K L R E L N F S A L S L P N L T N L S A S R
1401 ATGGCAATAAAGTCAATCAGAATGTGTATCTTAAACTCTGCCCAACTTAAAGCTTGAATCTGAGTGGAAACGGTGATAAAATTTGAAAATCTTTCCGC
280 N G N K V I Q N V Y L K T L P Q L K S L N L S G T V I K L E N L S A

PstI (1507) 1501 CAAACACCTGCAGAACTAAGAGCTATGGATCTCAGTAATTTGGGAGCTTAGACATGGTCACTTAGATATGAAAAGTGTGTCACCTGCTCGAAACCTA
313 K H L Q N L R A M D L S N W E L R H G H L D M K T V C H L L G N L

Tth111I (1550)
PshAI (1690) 1601 CCCAAGTTAGAGACTGGTTTTTCAGAAAAATGTCACCAATGCAGAGGGCATTAAAGCAGCTGGCAAAGTGCACCAGGCTCTTGTTCCTTGACCTGGGTC
347 P K L E T L V F Q K N V T N A E G I K Q L A K C T R L L F L D L G
1701 AAAACAGTGACTTGATTTATCTCAATGACAGCGAGTTCAATGCTTCCAGTCTCCAAAAGTGAATTTGAACAAGTGCCAGCTCTCCTTCATCAACAA
380 Q N S D L I Y L N D S E F N A L P S L Q K L N L N K K C Q L S F I N N

BglII (1836) 1801 CAGGACCTGGAGTTCCTTGCAAGAACTTGACCAGCCTAGATCTGAGCCACAACAAGTTTAAAGCTTTCCAGATTTTGCAATTTCCCCCTTGAAGCACTTG
413 R T W S S L Q N L T S L D L S H N K F K S F P D F A F S P L K H L
1901 GAGTTCTCTCTTTCAAGAAACCCATCAGAACTCAATAATCTGGCCTTTAGTGGGCTATTTGCACTGAAGGAGCTCAACTTTGGTGCATTTGGGA
447 E F L S T L S R N P I T E L N N L A F S G L F A L K E L N L A A C W

BglII (2052) 2001 TTGTAACAATTGACAGGTATTCCTTACTCAGTTTCCAACTTAGAGGTCTTAGATCTTGAGGACAACAATATTCGGACTCTCAACCACGGAACCTTCGG
480 I V T I D R Y S F T Q F P N L E V L D L G D N N I R T L N H G T F R
2101 ACCTCTGAAAAAAGTCCAGTCTTTGATTCTTTCCACAAGTCCGCTTAAAGTCTGGAGCCAAAATTCGTTTTCTGGTCTCACTAACCTACGTTCCCTTGAC
513 P L K K L Q S L I L S H N C L K I L E P N S F S G L T N L R S L D

BsrGI (2204) 2201 CTGATGTACAACGCTTGTGCATATTTTCATGAACACCTTTTCTCGGGCCTTGAAAAGCTTCTGATTTTGAAGCTTGGTTTAAATAGATCACATATGAAA
547 L M Y N S L S Y F H E H L F S G L E K L L I L K L G F N K I T Y E
2301 CTACTAGGACCTTCAGTATCCTCCATTATAAAGCTCAAGTCTTTGAAACAGCTTAACTTAGAAGGACAAAGACACGGGATTGAGTTGTTCCGAGCAA
580 T T R T L Q Y P P F I K L K S L K Q L N L E G Q R H G I Q V V P S N
2401 CTTTTTCAAAGGATTGGGTAGTTTGCAGGAGTACTCTTAGGAAAAAATCCCTCTGTATTCTGGACCACCAATTTGACCTCTGATTAACCTCACA
613 F F Q G L G S L Q E L L L G K N P S V F L D H H Q F D P L I N L T

EcoRV (2506) 2501 AAGTTGGATATCTCAGGAACAAAAGATGGAGATCGAAGCCTCTATTTAAATGCTTCTTATTCCAAAAGCTCAAAGGCTAAAGATCCTCCGCTTGAAA
647 K L D I S G T K D G D R S L Y L N A S L F Q N L K R L K I L R L E

BsaBI (2524) 2601 ATAACAACCTTAGAGTCACTGGTCTCCAGCTTACAGAGCCTCCAGGCTTTTTCTTTAAGATTCAACAACCTTGAAGTCAATTAATCAAAG
680 N N N L E S L V P D M F S S L Q S L Q V F S L R F N N L K V I N Q S
2701 TCATCTGAAGAATCTGAAATCACTGATGTTTTTGTGTCATGGGAACAACTTCAAGTGCACCTGTGACAATTTGTTGTTCAAGAAGTGGTCAATGAAC
713 H L K N L K S L M F F D V Y G N K L Q C T C D N L W F K N W S M N

BspEI (2825) 2801 ACAGAGGAGTCCACATCCCCTCCGAGCTATCCCTGTCAGCAGCCAGGACGCGAGTTTACTTATAGATTTTGTGATGATGCCATGTGTAATTTG
747 T E E V H I P F L R S Y P C Q Q P G S Q S L L I D F D D A M C N F

2901 ACTTGGGAAAGGTCTACTTCTTATGTTCCCTTCAGTATGGTCCCTCAGCACCATGGTTTTCTCTTGGTTCAGTACCAAGATGATTGCATCTCTGTGGTATGG
780▶ D L G K V Y F L C S F S M V L S T M V F S W F S T K M I A S L W Y G
BsrGI (3002) Acc65I (3020)

3001 TTTGTACATATGTAGGGCCTGGTACCTCACTAAATGGCACAAAACGGAGAAGAAGTTCTTATATGATGCGTTTGTCTCTTTCTCGGCCACTGATGAGGCA
813▶ L Y I C R A W Y L T K W H K T E K K F L Y D A F V S F S A T D E A
Bst1107I (3103) EcoRV (3193)

3101 TGGGTATACAAAAGAGCTTGTCCAGCCCTAGAACAAGGAGCCAGACCCTTTAAACTCTGTCTTCCACACGGGATTTTGAACCAGGCATTGATATCT
847▶ W V Y K E L V P A L E Q G S Q T T F K L C L H Q R D F E P G I D I
3201 TCGAGAACATCCAGAATGCTATTAACACAAGCAGGAAAACCTTGTGCGTAGTCAGTAACCACTACCTACACAGTGAATGGTGCCGACTTGAAGTCCAGCT
880▶ F E N I Q N A I N T S R K T L C V V S N H Y L H S E W C R L E V Q L
NheI (3301)

3301 GGCTAGCATGAAGATGTTTTATGAGCACAAAGGATGTCATTATCTTGATCTTCTTGAAGAGATTCCGAACTATAAGCTGTCCAGCTACCACCGACTCAGG
913▶ A S M K M F Y E H K D V I I L I F L E E I P N Y K L S S Y H R L R
BstXI (3454)

3401 AAACCTATAAAACAAACAGACGTTTATCACTTGGCCGGACAGTGTTCACCAGCAGCCACTTTTCTGGGCTCGCATCAGAAATGCGTTGGCAAGGAGACTG
947▶ K L I N K Q T F I T W P D S V H Q Q P L F W A R I R N A L G K E T
BamHI (3535)

3501 TGGAGAAAAGAAAATACACATCTAATTGTTGTTGAGGGATCCCTATCCCTATGATGTGCCAGACTATGCTGGCTATCCATATGATGTTCTGATTATGCTGG
980▶ V E K E N T H L I V V E G S Y P Y D V P D Y A G Y P Y D V P D Y A G
MscI (3638)

3601 ATACCCCTTATGATGTGCCAGACTATGCCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAA
1013▶ Y P Y D V P D Y A •
NheI (3632)

3701 AAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACCAAGTTAAACAACAACATTGCATTCATTTTATG
HpaI (3770)

3801 TTTCAGGTTCCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACC
EcoRI (3866)

3901 TCCAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCACCT
4001 TCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGTTTGAAGTACTGCTCTTCATTTCTTTATGTTTTAAATGCAGTACCTCCCACATTCCTTTTT
4101 AGTAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCC
4201 CAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTCGGTGTACTTGAGGGGG
1414 • N R T Y K L P

4301 ATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCAC
1324 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 G C

4401 AGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGA
994 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 S
StuI (4544)

4501 CCCAATGGCAATGGCTTTCAGCACAGACAGTACCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCC
664 G I A I A E A C V T V R G I Y A E I H V A S I I E G T K T R I A A
4601 CCGACATGGTGCTTGTTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGA
324 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 M
AseI (4752)

4701 TGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGT
4801 TCACTAAACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCG
SpeI (4907)

4901 TTGATTTACTAGTCAAAAACAACTCCCATTGACGTCAATGGGGTGAGACTTGAAATCCCCGTGAGTCAAACCCTATCCACGCCATTGATGTACTGC
SnaBI (5035)

5001 CAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGG
5101 GCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCTCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTG
5201 ACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTGCTTGGGCGGTGACCCAGGCGGGCCATTT
PacI (5326)

5301 ACCGTAAGTTATGTAACGCTGCAGGTTAAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGGCCGCTTGGCTGGCGTTTTT
PstI (5319) SdaI (5318) BspLU11I (5336)

5401 CCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCT
5501 GGAAGTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC
5601 GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACTA
5701 TCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG

5801 TTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCT

5901 CTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTGTGGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGAT

6001 CTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATC AGCGGCCGCAATAAAA

PacI (6066)

EagI (6086)

NotI (6085)

6101 TATCTTTATTTTCATTACATCTGTGTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAA

6201 ATAGGCTGTCCCCAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA