



125

1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGCTCTCCAGGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

NcoI (560)
BstEII (555)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATTGGCTTGGGAAGCGAGGCGCGAACCCGGCCCCGAAG
1 M A W E A R R E P G P R R
601 GGCCGCCGTCGGGAGACGGTATGCTGTTGCTGTGCCTGGGGTCCCACCGGCCGCCCTACAACGTGGACTGAGAGCGCGCTGTTTACCAGGGC
13 A A V R E T V M L L L C L G V P T G R P Y N V D T E S A L L Y Q G
701 CCCACAACACGCTGTTCCGGTACTCGTCTGCTGCACAGCCACGGGGCAACCGATGGCTCCTAGTGGGTGCGCCACTGCCAAGTGGCTCGCAACG
47 P H N T L F G Y S V V L H S H G A N R W L L V G A P T A N W L A N
801 CTTAGTATCAATCCCGGGGATTTACAGATGCAGATCGGAAAGAAATCCCGCCAGAGCTGCGAACAGCTCCAGCTGGTAGCCCTAATGAGAAACC
80 A S V I N P G A I Y R C R I G K N P G Q T C E Q L Q L G S P N G E P
901 TGTGAAAGACTTGTGGAAAGAGAGACAATCAGTGGTGGGGTCCACTTTCCAGACAGCCAGGAGAAAATGGATCCATCGTACTTGTGGGCAT
113 C G K T C L E E R D N Q W L G V T L S R Q P G E N G S I V T C G H
1001 AGATGAAAAATATATTTTACATAAAGAATGAAAATAAGCTCCCACTGGTGGTGTATGGAGTGGCCCTGATTTACGAACAGAAGTGGTAAAAGAA
147 R W K N I F Y I K N E N K L P T G G C Y G V P P D L R T E L S K R
1101 TAGCTCCGTGTTATCAAGATTATGTAAAAATTTGGAGAAAAATTTGCATCATGTCAAGCTGGAATATCCAGTTTTTACACAAGGATTTAATTGTGAT
180 I A P C Y Q D Y V K K F G E N F A S C Q A G I S S F Y T K D L I V M
1201 GGGGCCCCAGGATCATCTTACTGGACTGGCTCTTTTTGTACAATAAATAACAAGGCTTTTTAGACAAAACAAAATCAAGTAAAA
213 G A P G S S Y W T G S L F V Y N I T T N K Y K A F L D K Q N Q V K
1301 TTTGGAAGTTATTTAGGATATTCAGTCGGAGCTGGTCAATTTCCGAGCCAGCATACTACCGAAGTAGTCGGAGGAGCTCCTCAACATGAGCAGATTGGTA
247 F G S Y L G Y S V G A G H F R S Q H T T E V V G G A P Q H E Q I G
1401 AGGCATATATATTCAGCATTGATGAAAAAGAACTAAATATCTTACATGAAATGAAAGTAAAAAGCTTGGATCGTACTTTGGAGCTTCTGTCTGTGCTGT
280 K A Y I F S I D E K E L N I L H E M K G K K L G S Y F G A S V C A V
1501 GGACCTCAATGCAGATGGCTTCTCAGATCTGCTGCTGGGAGCACCATGCAGAGCACCATCAGAGAGGAAGGAGTGTTTGTGTACACTCAACTCTGGC
313 D L N A G A D F S D L V G A A P M Q S T I R E E G R V F V Y I N S G
1601 TCGGAGCAGTAATGAATGCAATGGAACAAAACCTCGTTGGAAGTGACAAATATGCTGCAAGATTTGGGGAATCTATAGTTAATCTTGGCGACATTGACA
347 S G A V M N A M E T N L V G S D K Y A A R F G E S I V N L G D I D
1701 ATGATGGCTTTGAAGATGTTGCTATCGGAGCTCCACAAGAAGTACTTGAAGGTGCTATTTATATTTACAATGGCCGTGCAGATGGGATCTCGTCAAC
380 N D G F E D V A I G A P Q E D D L Q G A I Y I Y N G R A D G I S S T
1801 CTTCTCACAGAGAATTGAAGGACTTCAGATCAGCAAATCGTTAAGTATGTTTGGACAGTCTATATCAGGACAAAATGATGCAGATAAATAGGCTATGTA
413 F S Q R I E G L Q I S K S L S M F G Q S I S G Q I D A D N N G Y V
1901 GATGTAGCAGTGGTGTCTCGGCTGATTCTGCTGCTTGGTAAAGGACAAGACTGTAGTAATTTGACGCTTCTTAAGCCACCTGAGTCAGTAA
447 D V A V G A F R S D S A V L L R T R P V V I V D A S L S H P E S V
2001 ATAGAACGAAATTTGACTGTGTTGAAAAATGGATGGCCTTCTGTGTGCATAGATCTAACACTTTGTTTCTCATATAAGGGCAAGGAAGTCCAGGTTACAT
480 N R T K F D C V E N G W P S V C I D L T L C F S Y K G K E V P G Y I
2101 TGTTTTGTTTATAACATGAGTTTGGATGTGAACAGAAAGGCAGAGTCTCCACCAAGATTCTATTTCTTCTAATGGAACCTCTGACGTGATTACAGGA
513 V L F Y N M S L D V N R K A E S P P R F Y F S S N G T S D V I T G
2201 AGCATACAGGTGTCCAGCAGAGAAGCTAACTGTAGAACACATCAAGCATTATGCGGAAAGATGTGCGGGACATCCTCACCCCAATTCAGATTGAAGCTG
547 S I Q V S S R E A N C R T H Q A F M R K D V R D I L T P I Q I E A
2301 CTTACCACCTTGGTCTCATGTCATCAGTAAACGAAGTACAGAGGAATCCCACTTCCAGCAATCTTTCAGCAGAAGAAAAGAACATAATGAA
580 A Y H L G P H V I S K R S T E E F P P L Q Q P I L Q Q K K E K D I M K
2401 AAAAAGCATAAATTTGCAAGGTTTTGTGCCATGAAAATTTGCTGCTGATTACAGGTTTTGCAAAAGATTGGGTTTTTGAAGCCCATGAAAATAAA
613 K T I N F A R F C A H E N C S A D L Q V S A K I G F L K P H E N K
2501 ACATATCTTGTGTTGGGAGTATGAAGACATTGATGTTGAATGTGCTTGTTTAATGCTGGAGATGATGCATATGAAACGACTCTACATGTCAAACCTAC
647 T Y L A V G S M K T L M L N V S L F N A G D D A Y E T T L H V K L
2601 CCGTGGTCTTTATTTTATTAAGATTTTAGAGCTGGAAGAGAAGCAATAAAGTGTGAAGTACAGATAAAGTCTGGCGTGGTACAAGTACTGACTGCAGTAT
680 P V G L Y F I K I L E L E E K Q I N C E V T D N S G V V Q L D C S I
2701 TGGCTATATATGTAGATCATCTCAAGGATAGATTAGCTTTCTCGGATGTGAGTCACTCAGCAGAGCGGAAGAGGACTCAGTATCAGATG
713 G Y I Y V D H L S R I D I S F L L D V S S L S R A E E D L S I T V
2801 CATGCTACCTGTGAAAATGAAGAGGAAATGGACAATCTAAAGCACAGCAGAGTACTGTAGCAATACCTTTAAAATATGAGGTTAAGCTGACTGTTTCATG
747 H A T C E N E E E M D N L K H S R V T V A I P L K Y E V K L T V H
2901 GGTGTGAAACCAACTTCATTTGTGTATGGATCAAATGATGAAAATGAGCCTGAAACGTGCATGGTGGAGAAAATGAACTTAACTTTCCATGTTATCAA
780 G F V N P T S F V Y G S N D E N E P E T C M V E K M N L T F H V I N
3001 CACTGGCAATAGTATGGTCCCAATGTTAGTGTGAAATAATGGTACCAAAATCTTTAGCCCCAACTGATAAGCTGTTCAACATTTTGGATGTCAG
813 T G N S M A P N V S V E I M V P N S F S A P Q T D K L F N I L D V Q
3101 ACTACTACTGGAGAATGCCACTTTGAAAATTAACAAGAGTGTGTGCATTAGAGCAGCAAAAGAGTGAATGCAGACCTTGAAGGCATAGTCCAGTTCT
847 T T T G E C H F E N Y Q R V C A L E Q Q K S A M Q T L K G I V Q F

T T T G E C H F E N Y Q R V C A L E Q Q K S A M Q T L K G I V Q F
3201 TGCCAAGACTGATAAGAGGCTATTGTACTGCATAAAAGCTGATCCACATTGTTAAATTTCTTGTAATTTGGGAAAATGGAAAGTGGAAAAGAAGC
880▶ L S K T D K R L L Y C I K A D P H C L N F L C N F G K M E S G K E A
3301 CAGTGTTTCATATCCAACCTGGAAGGCCGCCATCCATTTTAGAAATGGATGAGACTTCAGCACTCAAGTTTGAATAAGAGCAACAGGTTTTCCAGAGCCA
913▶ S V H I Q L E G R P S I L E M D E T S A L K F E I R A T G F P E P
3401 AATCCAAGAGTAATTGAACTAAACAAGGATGAGAATGTTGCGCAGTTCTACTGGAAGGACTACATCATCAAAGACCCAAACGTTATTTCCACATAGTGA
947▶ N P R V I E L N K D E N V A H V L L E G L H H Q R P K R Y F T I V
3501 TTTATTTCAAGTAGCTTGCTACTTGACTTATTGTACTTCTGTTGATCTCATATGTTATGTGGAAGGCTGGCTTCTTTAAAAGACAATACAAATCTATCCT
980▶ I I S S S L L L G L I V L L L I S Y V M W K A G F F K R Q Y K S I L

NheI (3699)

3601 ACAAGAAGAAAACAGAAGAGACAGTTGGAGTTATATCAACAGTAAAAGCAATGATGATTAAGGACTTCTTTCAAATGAGAGAATGGAAAACAGACTGAC
1013▶ Q E E N R R D S W S Y I N S K S N D D •

MscI (3705)

3701 CTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTAT

3801 TGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAAACAACATTCATTCTTTTATGTTTCAGGTTTCCAGGGGAGGTGTGGGAGGTTTTT

3901 TAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAAACTTTAACCTCCAATCAAGCCTCTACTGAACTCTTTCTG

4001 AGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTTC

4101 CCAAGTTTGAAGTAGCTCTTCAATTTCTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCA

4201 TTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAA

4301 CCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTG

4401 CCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCT

4501 CATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCCTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTGAC

4601 CCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGTCCTCATAGAGCATG

4701 GTGATCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTGTATTATACTATG

4801 CCGATATACTATGCCGATGATTAATTGTCAAACACGCTGGATGGCGTCTCCAGCTTATCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCC

4901 ACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTACTAGTCAAACAAACTCCATTGAC

5001 GTCAATGGGGTGGAGACTTGGAAATCCCCTGAGTCAAACCGCTATCCAGCCATTGATGTACTGCCAAAACCGCATCATGGAATAGCGATGACT

5101 AATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCATTACCCTGATTGACGTCATAGGGGGCG

5201 TACTTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCCTAAATACTCCACCCATTGACGTCATGGAAGTCCCTATTGGCGTACTAT

5301 GGGAAACATACGTCATTATTGACGTCATGGGCGGGGCTGTTGGGCGTCCAGCAGGCGGGCATTACCCTAAGTTATGTAACGCTGCAGGTTAATTA

5401 AGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAGGCCGCTTGGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAA

5501 AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACC

5601 CTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCC

5701 GCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAAGACAGACTT

5801 ATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCCTGAAGTGGTGGCCTAACTACGGCTACACT

5901 AGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCCGCTGGTAGCG

6001 GTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGGTCTGACGCTCAGTGAACGA

6101 AAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAAATATCTTTATTTTTCATTACATCTGTGTGGTT

6201 TTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGAAGTGCAGGTGCCAG

6301 AACATTTCTCTATCGAA