



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTGCCGCCCTCCCGCTGTGGTGCCTCTGAAGCTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGAGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTGTTTCGTTT

NcoI (560)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCTACCTGAGATCACCGGTGAGGGCTCTGGGGCGGTGCGGGCGAGCGGGCC
601 GCTGCTGGGCGCGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCTGGGGCGTTGCGGGCGGGCGCGCGGTGCGGGCGGTGCTGCTGACGGAGTGTG
13▶ L L G A R R A A A A A A A A G A F A G R R A A C G A V L L T E L L
701 GAGCGCGCGCTTTCTACGGCATCACGTCCAACCTGGTCTATTCTGAACGGGGCGCGTTCTGCTGGGAGGGCGCGCAGGCCAGCGAGGGCGTGTGC
47▶ E R A A F Y G I T S N L V L F L N G A P F C W E G A Q A S E A L L
801 TCTTCATGGGCTCACCTACCTGGGCTCGCGTTCGGAGGCTGGTGGCCGACGCGCGGTGGGCCGGGCGCGCCATCCTGCTGAGCCTGGCGCTA
80▶ L F M G L T Y L G S P F G G W L A D A R L G R A R A I L L S L A L Y
901 CCTGCTGGGCATGCTGGCCTTCCCGTGTGGCGCGCCCGCCACGCGAGCCGCTGCGGTTCCGCGCGCTCAACTGCACGGCGCCTGGTCCC
113▶ L L G M L A F P L L A A P A T R A A L C G S A R L L N C T A P G P
1001 GACGCGCGCGCGCTGCTGCTCACCAGCCACCTTCGCGGGGCTGGTGTGGTGGGCTGGCGTGGCCACCGTCAAGGCCAATCAGCCCTTCGCGG
147▶ D A A A R C C S P A T F A G L V L V G L G V A T V K A N I T P F G
1101 CCGACCAGTTAAAGATCGAGGTCGGAAGCCACTAGGAGATTTTTAATTGGTTTTATTGGAGCATTAACTGGGAGCGATCTGTCTTAGGTGGCAT
180▶ A D Q V K D R G P E A T R R F F N W F Y W S I N L G A I L S L G G I
1201 TGCTATATTCAGCAGAAGCTCAGCTTTGCTACTGGTATGCGATCCCACTGTCTGCGTCCGCTTGTCTTTGTGGTCTTCTCTGTGGCCAGAGCGTT
213▶ A Y I Q Q N V S F V T G Y A I P T V C V G L A F V V F L C G Q S V
1301 TTCATACCAAGCTCTGATGGCAGTGCCTTACCAGACATGTTCAAGATACTGACGTATTCCTGCTGTTCCAGAAGCGAAGTGGAGAGCGCCAGAGTA
247▶ F I T K P P D G S A F T D M F K I L T Y S C C S Q K R S G E R Q S
1401 ATGGTGAAGGCATTGGAGTCTTTCAGCAATCTTCTAAACAAAGTCTGTTTGATTGATGTAAGATGTCTCATGGTGGGCCATTTACAGAAGAGAAAGTGA
280▶ N G E G I G V F Q Q S S K Q S L F D S C K M S H G G P F T E E K V E
1501 AGATGTAAAAGCTCTGGTCAAGATTGTCCTGTTTTCTGGCTTTGATACCTACTGGACAGTGTATTTCAAATGCAGACAACATATGTTTTACAGAGT
313▶ D V K A L V K I V P V F L A L I P Y W T V Y F Q M Q T T Y V L Q S
1601 CTTCAATTTGAGGATCCAGAAATTTCAAATATTACAACACTCCTCACAGCTCCCTGCGAGCTGGTGGCCATGTTGTAGTGTGTCTATCTCTCTGC
347▶ L H L R I P E I S N I T T P H T L P A A W L T M F D A V L T I L L
1701 TCATCCCTCTGAAGGACAACTGGTGCATCCCATTTTGAGAAGACATGGCCTGCTCCCATCCCTGAAGAGGATGCGCGTGGGCVLTTCTTTGTCA
380▶ L I P L K D K L V D P I L R R H G L L P S S L K R I A V G M F F V M
1801 GTGCTCAGCCTTTGCTGCAGGAATTTGGAGAGTAAAAGGTGAACCTTGTAAAAGAGAAAACCATTAATCAGACCATCGGCAACGTCGTCTACCATGCT
413▶ C S A F A A G I L E S K R L N L V K E K T I N Q T I G N V V Y H A
1901 GCCGATCTGCTGCTGTGGTGGCAGGTGCCGAGTACTTGTGATTGGGATCAGCGAGATCTTGAAGTATCGCAGGCCTGGAATTTGCATACCTCAGCTG
447▶ A D L S L W W Q V P Q Y L L I G I S E I F A S I A G L E F A Y S A
2001 CCCCCAAGTCCATGCAGAGTGCATAATGGGCTTGTCTTTCTTCTGCGTGGGTCGTTCTGTTGGGTTCTGGACTGTGGCAGTGGTGTCTATCAA
480▶ A P K S M Q S A I M G L F F F S G V G S F V G S G L L A L V S I K
2101 AGCCATCGGATGGATGAGCAGTACACAGACTTTGGAATATTAACGGCTGCTATTGAACTATTACTTTTTCTTCTGGCTGCTATTCAAGGACTACC
513▶ A I G W M S S H T D F G N I N G C Y L N Y Y F F L L A A I Q G A T
2201 CTCCTGTTTTCTCATTATTTCTGTAATATGACCATCATCGAGACCATCAGCGATCAAGAGCCAATGGCGTGGCCACCAGCAGGAGGGCCTGACCTT
547▶ L L L F L I I S V K Y D H H R D H Q R S R A N G V P T S R R A •

NheI (2313)

2301 CCTGAGGCCATGTCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAA
2401 ATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAACAACAACATTGCATTATTTATGTTTCAGTTTCAGGGGAGG
2501 TGTGGGAGGTTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACT
2601 TGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGAGTTTAAAG
2701 TATAGTGTATTTCCCAAGGTTTGAAGTACTGCTCTTCATTTCTTTATGTTTTAAATGCAGTGCCTCCACATTCCTTTTTAGTAAAATATTCAGAAATA
2801 ATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTT
2901 AGGGAACAAGGAACCTTAAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGG
3001 TTTTGACCAGCTTGCATTCTCAATGAGCACAAGCAGTCAAGGAGCATAGTCAAGATGAGTCTCTGCACATGCCACAGGGGCTGACCACCCTGAT
126▶ 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 P S V V R I
3101 GGATCTGCCACCTCATCAGAGTAGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCA
93▶ 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 G I A I A E

3201 GCACAGACAGTGACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGCTTGTGT
59 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 G V H H K N D
3301 CCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCATGGTGGCCCTCTATAGTGAGT
26 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
3401 CGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACAAACGAGCTCTGCT
3501 TATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTTACTAGTCAAAC
3601 AAATCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGG
3701 TAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGAC
3801 GTCAATAGGGGGCTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCT
3901 ATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACG
4001 CTGCAGGTTAAITTAAGAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCT
4101 GACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCT
4201 CTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTC
4301 GGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCG
4401 GTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA
4501 ACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAC
4601 CACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGAC
4701 GCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAATCAGCGGCCGCAATAAAATATCTTTATTTTCATTACA
4801 TCTGTGTGTTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAAACAAAACGAAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCA
4901 AGTGCAGGTGCCAGAACATTTCTCTATCGAA