



100

1 GGATCTGCATCGCTCCGGTGCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGGCTCTCCAGGCTTTGCCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

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

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGGAGATTTAGATTTCACAGAAGCCTACCCACAAAC
13▶ T E F D Y G D S T P C Q K T A V R A F G A G L L P P L Y S L V F I
701 ATTGGAGTGGTGGCAATGTCCTAGTATTCTGGTGCATGCAGCATAGGAGGCTTCAAAGCATGACCAGCATCTACCTGTTCAACCTGGCTGTCTCTG
47▶ I G V V G N V L V I L V L M Q H R R L Q S M T S I Y L F N L A V S
801 ATCTGGTCTTCTTTTCACTTTACCTTTCTGGATTGACTACAAGTTGAAAGACGACTGGATTTTGGTGATGCCATGTCAAGCTTCTCTCTGGGTTTTA
80▶ D L V F L F T L P F W I D Y K L K D D W I F G D A M C K L L S G F Y
901 TTACCTGGGTTTATACAGTGAGATCTTCTTTATCATCTGTTGACGATTGACAGATACCTGGCCATTGTCCATGCTGTGTTTGCCTGAGGGCCGAAC
113▶ Y L G L Y S E I F F I I L L T I D R Y L A I V H A V F A L R A R T
1001 GTTACTTTTGGCATCATACCAGTATTATCACCTGGGCCCTAGCCATCTTAGCTTCCATGCCTGCCTTATACTTTTTAAAGCCAGTGGGAGTTCACCTC
147▶ V T F G I I T S I I T W A L A I L A S M P A L Y F F K A Q W E F T
1101 ACCGTACTGTAGCCCTATTTCCCTACAAGAGCCTGAAGCAGTGAAGAGGTTTCAAGCTCAAAGCTAAACCTTCTTGGACTAATTTGCCTCTGTT
180▶ H R T C S P H F P Y K S L K Q W K R F Q A L K L N L L G L I L P L L
1201 AGTCATGATAATCTGCTATGCAGGGATCATCAGAATCTGCTCAGAAGACCCAGTGAAGAAGGTCAAAGCCGTGCGTCTGATATTTGCTATTACTCTT
213▶ V M I I C Y A G I I R I L L R R P S E K K V K A V R L I F A I T L
1301 CTATTCTCTCTCTGGACCCCTACAATCTGAGTGTATTGTTTCTGCTTCCAAGATGTTCTATTACCAATCAGTGTGAGCAGAGTAAGCAACTGG
247▶ L F F L L W T P Y N L S V F V S A F Q D V L F T N Q C E Q S K Q L
1401 ACCTGGCCATGCAGGTGACTGAGGTGATTGCCTACCCACTGTTGTGTAACCCAATCATTTATGTTTTGTGGGTGAACGGTTCTGGAAGTACCTTCG
280▶ D L A M Q V T E V I A Y T H C C V N P I I Y V F V G E R F W K Y L R
1501 GCAGCTGTTTCAAAGGCATGTGGCTATACCACTGGCAAATGGCTGCCTTCTCTCTGTGGACCAACTAGAAAGGACCAGTCTATATCTCCATCCACA
313▶ Q L F Q R H V A I P L A K W L P F L S V D Q L E R T S S I S P S T

NheI (1661)

1601 GGAGAACATGAGCTCTCTGCTGGCTTCTGATTGAGCCATAGGTGTAACCAAGGAAGGTGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTT
347▶ G E H E L S A G F •
1701 GGACAAACCAACTAGAAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAG
1801 TTAACAACAACAATTGCATTATTTTATGTTTCAGGTTTCAAGGGGAGGTGTTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATT
1901 CTAAAATACAGCATAGCAAACTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCA
2001 ATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTCTTCAATTTCTTTATGTTTTA
2101 AATGCACTGACCTCCACATTCCCTTTTATGAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCC
2201 AGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTGTGGACTTAGGGAACAAAGAACCTTAAATAGAAATGGACAGCAAGAAAGCGAGCTTCT
2301 AGCTTTAGTTCTGGTACTTGGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAAGCAGTCAAGGAGCATA
141▶ • 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 D P A Y
2401 GTCAGAGATGAGCTCTGACATGCCACAGGGGTGACCACCTGATGGATCTGCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCA
109▶ 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 V I T D
2501 AAGTCTTCTGCCGTTGCTCAGCAGACCCAATGGCAATGGCTTACGACAGACAGTACCTGCAATGTAGGCCTCAATGTGGACAGCAGAGATGA
75▶ 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 V A S I I
2601 TCTCCCAGTCTTGGTCTGATGGCCGCCCCGACATGGTGTCTGTTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATC
42▶ E G T K T R I A A G V V H K N D E Y L M T I K E T A V E V L E L D
2701 CTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGTGATGCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCG
9▶ Q Q S I N F T K M
2801 TGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGC
2901 GAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAA
3001 ACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCAAGTAGGAAAGTCCCATAA
3101 GGTGATGACTGGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCGTACTTGGCATATGATACACTGATGACTGCAAGTGG

3201 GCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGT
3301 CGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC
3401 CGTAAAAAGGCCGCTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG
3501 GACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTC
3601 GGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAG
3701 CCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
3801 GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCC
3901 AGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGA
4001 AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAAT
4101 TAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAA
4201 AACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGAAGTGCAAGTGCCAGGTGCCAGAACATTTCTCTATCGAA