



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGGCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

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

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGGAGCAGAGGAAGGGTCTGGCTGGCCTCTTCTGGT
601 GATCTCTCTTCTTCAAGGCACTGTAGCCAGACAAATAAAGCAAAGAAATTTGGTACAAGTGGATGGCAGCCGAGGAGACGGTCTGTACTTCTGACTTGT
13▶ I S L L Q G T V A Q T N K A K N L V Q V D G S R G D G S V L L T C
701 GGCTTGACTGACAAGACTATCAAGTGGCTTAAAGACGGGAGCATAATAAGTCTCTAAATGCAACTAAAAACACATGGAATCTGGGCAACAATGCCAAAG
47▶ G L T D K T I K W L K D G S I I S P L N A T K N T W N L G N N A K
801 ACCCTCGAGGCACGTATCAGTGTCAAGGAGCAAAGGAGACGTCAAACCCCTGCAAGTGTATTACAGAATGTGTAAAACTGCATTGAGCTAAACATAGG
80▶ D P R G T Y Q C Q G A K E T S N P L Q V Y Y R M C E N C I E L N I G
901 CACCATATCCGGCTTTATCTTCGCTGAGGTCATCAGCATCTTCTTCTTGGTGTATATCTCATTGCGGGACAGGATGGAGTTCGCCAGTCAAGA
113▶ T I S G F I F A E V I S I F F L A L G V Y L I A G Q D G V R Q S R
1001 GCTTCAGACAAGCAGACTCTGTTGCAAAATGAACAGCTGTACCAGCCCTCAAGGACCGGAAATATGACCAGTACAGCCATCTCCAAGGAAACCAACTGA
147▶ A S D K Q T L L Q N E Q L Y Q P L K D R E Y D Q Y S H L Q G N Q L

NheI (1127)

1101 GGAAGAAGTGAACCTCAGCAGGACTCAGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAAT
180▶ R K K •
1201 GCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTATTTTATGTTTCA
1301 GGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACCTTTAACCTCCAA
1401 ATCAAGCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTT
1501 CATGGAGTTAAGATATAGTGTATTTTCCCAAGTTTGAAC TAGCTCTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCTTTTATAGTAA
1601 AATATTAGAAATAATTTAAATACATCATTGCAATGAAAAATAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTT
1701 TAGTAGTTGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTCTGGTGTACTTGAGGGGGATGAG
141▶ • N R T Y K L P I L
1801 TTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGG
131▶ 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 P
1901 CTGACCACCTGATGGATCTGTCACCTCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCAGCAGAGACCCAA
97▶ 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 G I
2001 TGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGATGGCGCCCGAC
64▶ 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 G V
2101 ATGGTGTGTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCTCATGGTGGCC
31▶ 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
2201 CTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACT
2301 AAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGAT
2401 TTAGTAGTCAAACAACCTCCATTGACGTCAATGGGTGGAGACTTGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAAAA
2501 CCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGTAAGTGGGCATAATGCCAGGCGGGCCAT
2601 TTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTG
2701 AATGAAAGTCCCTATTGGCGTACTATGGAAACATACGTCATTATTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTACCGT
2801 AAGTTATGTAACGCTGCAGGTTAATAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATA
2901 GGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAG
3001 CTCCCTCGTGGCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCATAGTCAAGCTGT
3101 AGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTC

3201 TTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT
3301 GAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGA
3401 TCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTT
3501 CTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCT
3601 TTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGG
3701 CTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA