



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCCCTGTTGCTCAACTCTACGCTTTTGTTCGTTT

BspEI (558)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGCTCCGGACACCATGGCGCCGGCGGAGGCTCGCGGCCGCTCCC
GGCTGGATCTCCGTGCTGGGCTGGGCTGCGCTGTGCTCTGTGCGGGGCGGGCCACTGTGGAGTGGCAGCCATGAGTGGAAAAACTAATTTTG
11▶ G W I S V L G W G L A L C S L C G A G P L W S G S H E W K K L I L
701 ACCGAGCACTGGCCCCAACAGTATGCAAGGAGTTAACAGCTGCCAAGACTCTGTGATTACTGGACAATACATGGACTATGGCCGATAGAGCAGAAG
45▶ T Q H W P P T V C K E V N S C Q D S L D Y W T I H G L W P D R A E
801 ATTGTAACCACTCTGGCACTTAACTTAGATGAGATTAAGGACCTTTTGCAGACATGAAGATCTACTGGCCGATGTGATTACCGGCTTCTAATCG
78▶ D C N Q S W H F N L D E I K D L L R D M K I Y W P D V I H R S S N R
901 CAGCCAATTCTGAAACATGAGTGGGTTAAACACGGCACCTGTGCTGCCAGGTAGACGCCCTCAATCCGAGAAGAAGTACTTTGGGAAGAGCCTGGAT
111▶ S Q F W K H E W V K H G T C A A Q V D A L N S E K K Y F G K S L D
1001 CTGTACAAGCAGATTGACCTCAACAGTGTGCTACAAAAATTTGGGATCAAGCCATCCATCAACTACTACCAGCTTGACAGATTTCAAAGATGCACTTACCA
145▶ L Y K Q I D L N S V L Q K F G I K P S I N Y Y Q L A D F K D A L T
1101 GAATCTATGGTGTGGTGCCTAAAATCCAGTGCCTTATGCCAGAACAGGAGAGAGCGTGCAGACCGTTGGCCAGATAGAGCTGTGCTTACCAAGGAGGA
178▶ R I Y G V V P K I Q C L M P E Q G E S V Q T V G Q I E L C F T K E D
1201 CTTACATTTGCGAACTGCACTGAGCCAGGAGAGCAGCTGTCTCCAGGCAGGAAGCCTGGCTGGCCATGGAGGCTTACACATGGGATGATGGTCTGT
211▶ L H L R N C T E P G E Q L S S R Q E A W L A M E A S T H G M M V C

NheI (1349)

1301 GAAGACGGTCCAATCTTCTACCTCCACCTACAAAGACCCAACATTGATGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACA
245▶ E D G P I F Y P P P T K T Q H •
1401 ACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAAACA
1501 ATTGCAATTCATTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTTACAATGTGGTATGGAATTCATAAATACAGC
1601 ATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGC
1701 TGTTTGACGCTCACCTCTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTGAAGTAGCTCTTCAATTTCTTATGTTTTAAATGCACTGACC
1801 TCCCACATTCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAAATCCAGATGCTCAAGG
1901 CCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCC
2001 TGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAG
138▶ 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 D S I L
2101 CTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGC
105▶ 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 F D K Q
2201 CCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAAGCAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCT
71▶ 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 E G T K
2301 TGGTCTGATGGCCGCCGACATGGTCTTGTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGAT
38▶ T R I A A 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
2401 GTTGAAGTCTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCT
5▶ N F T K M
2501 CCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGA
2601 CATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCAC
2701 GCCAATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGTCATGTACTG
2801 GGCATAATGCCAGGCGGGCATTACCCTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCCAAGTGGGCAGTTTACCCT
2901 AAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGCTGTTGGGCGGTC
3001 AGCCAGGCGGGCATTACCCTAAGTTATGTAACGCCCTGCAGGTTAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCC
3101 GCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGAT

3201 ACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCGGACCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCTTCGGGAAGCGTGGC
3301 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGACCGCTGC
3401 GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATG
3501 TAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG
3601 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT
3701 CAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAAT
3801 CAGCGCCGCAATAAAAATATCTTTATTTTCATTACATCTGTGTGGTTTTTTTGTGTAATCGTAACATAACGCTCTCCATCAAAACAAAACGAAA
3901 CAAAACAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAAGTGCCAGAACATTTCTCTATCGAA