



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BstEII (555)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCCTACCTGAGATCACCGGTACCATTGGATTCTGAAAACCTGTTCTATAACGGAAAATAGCTC
1 M D S E N C S I T E N S S
601 TTCACATCTGGAGAGAGGGCAGAAGGACCATGGCACCAGTATACATTTGAGAAGCATCAAGGATCCATTCAAGTTTCTATCCCTGGGCTGTGTTA
13 S H L E R G Q K D H G T S I H F E K H H E G S I Q V S I P W A V L
701 ATAGTGGTCCATCACGTCTTAATAATAGCTCTCATTGCCTTAAATGTGGGCAAGTACAATTGCCAGGCTGTACGAGAAGTTGGAATCATCTGACC
47 I V V L I T S L I I A L I A L N V G K Y N C P G L Y E K L E S S D
801 ACCATGTTGCTACCTGCAAGAATGAGTGGATTTACATAAAGAGGACATGTTACTTCTCTCCACCACAACCAAGAGTTGGGCCTTGGCCCAACGCTCTTG
80 H H V A T C K N E W I S Y K R T C Y F F S T T T K S W A L A Q R S C
901 TTCTGAAGATGCTGCTACTCTTGCTGTAATTGATTAGAAAAGGACATGACGTTTCTGAAGCGATATTCTGGTGAAGTGAACATTGGATTGGCTGAAA
113 S E D A A T L A V I D S E K D M T F L K R Y S G E L E H W I G L K
1001 AATGAAGTAATCAGACATGAAAATGGCAAATGGCAAAGAAATTAACAGCTGGTCAACTTGACGGGGTCTGGGAGGTGCGTGTCCGTGAACCAAAAA
147 N E A N Q T W K W A N G K E F N S W F N L T G S G R C V S V N H K

NheI (1162)

1101 ATGTTACCGCTGTGGACTGTGAGGCAAACCTCCACTGGTCTGCAGCAAGCCCTCCAGATGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTT
180 N V T A V D C E A N F H W V C S K P S R •
1201 TGGACAAACCACAACCTAGAATGCAGTAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAA
1301 GTTAACAACAACAATTGCATTATTTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAAT
1401 TCTAAAATACAGCATAGCAAACCTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGGC
1501 AATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGTTTGAAGTACTCTTCATTCTTTATGTTTT
1601 AAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATC
1701 CAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCGAGCTTC
1801 TAGCTTTAGTTCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCAT
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
1901 AGTCAGAGATGAGCTCTGACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTG
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
2001 AAAGTCTTCTGCCGTTGCTCAGCAGACCCAATGGCAATGGCTTACAGACAGCAGTACCCTGCCAATGTAGGCTCAATGTGGACAGCAGAGATG
76 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
2101 ATCTCCCCAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGCTCCTCATAGAGCATGGTGTCTTCTCAGTGGCAGCTCCACCAGCTCCAGAT
42 I E G T K 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
2201 CCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGTGCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGC
9 Q Q S I N F T K M
2301 GTGGATGGCGTCTCCAGCTTATCTGACGGTTACTAAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTGCGTCAATGGGGC
2401 GGAGTTGTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGTGGAGACTTGGAAATCCCCGTGAGTCA
2501 AACCGCTATCCACGCCATTGATGTAAGTCCGATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCAAGTAGGAAAGTCCATA
2601 AGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTG
2701 GGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGG
2801 TCGTTGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCTGACGTTAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA
2901 CCGTAAAAAGGCCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAATAATGACGCTCAAGTCAGAGGTGGCGAAACCCGACA
3001 GGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTT
3101 CGGGAAGCGTGGCGTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCAGAACCCCGTTCA

3201 G C C C G A C C G T G C G C T T A T C C G G T A A C T A T C G T C T T G A G T C C A A C C C G T A A G A C A C G A C T T A T C G C C A C T G G C A G C A G C C A C T G G T A A C A G G A T T A G C
3301 A G A G C G A G G T A T G T A G G C G G T G C T A C A G A G T T C T T G A A G T G G T G G C C T A A C T A C G G C T A C A C T A G A A G A A C A G T A T T T G G T A T C T G C G C T C T G C T G A A G C
3401 C A G T T A C C T T C G G A A A A G A G T T G G T A G C T C T T G A T C C G G C A A C A A C C A C C G C T G G T A G C G G T G G T T T T T T G T T G C A A G C A G C A G A T T A C G C G C A G
3501 A A A A A A G G A T C T C A A G A A G A T C C T T T G A T C T T T T C T A C G G G T C T G A C G C T C A G T G G A A C G A A A A C T C A C G T T A A G G G A T T T T G G T C A T G G C T A G T T A A
3601 T T A A C A T T T A A A T C A G C G G C C G C A A T A A A A T A T C T T T A T T T T C A T T A C A T C T G T G T G T T G G T T T T T T G T G A A T C G T A A C T A A C A T A C G C T C T C C A T C A
3701 A A C A A A A C G A A A C A A A A C A A A C T A G C A A A A T A G G C T G T C C C C A G T G C A A G T G C A G G T G C C A G A A C A T T T C T C T A T C G A A