



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTATCATGAAAGGCTTTGGTCTTGCCTTTGGACTGTTCTCCGC
1 M K G F G L A F G L F S A
601 TGTGGTTTTCTCTCTGGACTCCTTTAACTGGGCTCAAGACCTCCATTTGGGAAGCTGTGTGATTACTGCAAACCTACAGGCAATACAAAAGGAATTT
13 V G F L L W T P L T G L K T L H L G S C V I T A N L Q A I Q K E F
701 TCTGAGATTCGGGATAGTGTGCAAGCTGAAGATACAAATATTGACATCAGAATTTAAGGACGACTGAGTCTTTGAAAGACATAAAGTCTTTGGATAGGT
47 S E I R D S V Q A E D T N I D I R I L R T T E S L K D I K S L D R
801 GCTGCTTCCTTCGTCATCTAGTGAAGTCTATCTGACAGGGTATTCAAAGTCTACCAGACCCTGACCACCATACCTGAGAAAAGATCAGCAGCCTCGC
80 C C F L R H L V R F Y L D R V F K V Y Q T P D H H T L R K I S S L A
901 CAACTCCTTTCTATCATCAAGAAGGACCTCTCAGTCTGTCTTCTCACATGGCATGTCATTGTGGGAAGAAGCAATGGAGAAAATACAAACCAATCTG
113 N S F L I I K K D L S V C H S H M A C H C G E E A M E K Y N Q I L
1001 AGTCACTTCATAGAGTTGGAAGTTCAGGACGCGGTGTAAGGCTTTGGGAGAACTAGGCATCTTCTGAGATGGATGGAGGAGATGCATAGATGAAAG
147 S H F I E L E L Q A A V V K A L G E L G I L L R W M E E M L •

NheI (1138)

1101 TGGAGAGGCTGCTGAGAACACTCCTGTCCAAGAATCTCGTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCA
1201 GTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAAATTCATTCA
1301 TTTATGTTTCAGGTTACAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCATAAATACAGCATAGCAAACT
1401 TTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTCAGCC
1501 TCACCTCTTTTCATGGAGTTTAAAGATATAGTGTATTTCCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCC
1601 CTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAAT
1701 ATCCCCAGTTTAGTAGTTGACTTAGGGAACAAGGAACCTTTAATAGAATTTGACAGCAAGAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTG
141 • N R T Y K
1801 AGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACA
134 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 E R C M
1901 TGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCCGTTGCTCAC
101 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 G N S V
2001 AGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGCTGATG
68 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 T R I
2101 GCCGCCCCGACATGGTGTCTGTGCTCATAGAGCATGGTATCTTCTCAGTGGCAGCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGTCT
34 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 N F T K
2201 TCATGGTGGCCCTCTATAGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCT
1 M
2301 GACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAA
2401 GTCCCGTTGATTTACTAGTCAAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCGTGAGTCAAACCGCTATCCACGCCATTGATG
2501 TACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAGGTCATGACTGGGCATAATGCC
2601 AGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGACTGCAAGTGGGCGATTACCGTAAATACTCCAC
2701 CCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATTGACGTCAATGGGCGGGGCTGTTGGGCGGTGAGCCAGGCGGG
2801 CCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCTTGTGTC
2901 GTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTT
3001 CCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCCGGAAGCGTGGCGCTTCTCATA
3101 GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCCTGCGCTTATCCGG

3201 TAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT
3301 ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTG
3401 GTAGCTCTTGATCCGGCAAACAACACCAGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCC
3501 TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCA
3601 ATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACT
3701 AGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA