



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCCTCCCGCCTGTGGTGCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

SphI (560)

AgeI (552)

501 TCTGTTCTGCGCAGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTCAGCATGCAAAGAATCCTGATGGCCCTGGCCAAAGCTGCTG
1 M Q R I L M A P G Q S C C
601 TGCCTGGCCATCCTGCTGGCAATTGTGAACCTCCAACATGGTGGATGATTATCATGTCACCAGCTCAGCATCCAGAAAGGAGGGGCGACTGGACCTCACC
13 A L A I L L A I V N F Q H G G C I H V T S S A S Q K G G R L D L T
701 TGACTTTGTGGCACAAGAAAGACGAAGCTGAGGGGCTAATACTCTTCTGGTCAAAGACAATCCTTGGAAGTGTCCCTGAGACCACTTAGAACAGC
47 C T L W H K K D E A E G L I L F W C K D N P W N C S P E T S L E Q
801 TTAGGGTTAAAAGGGATCCTGAGACAGATGGCATCACTGAAAAGTCACTCAGTTGGTGTCCACATAGAACAAAGCTACACCATCAGACAGTGGGACCTA
80 L R V K R D P E T D G I T E K S S Q L V F T I E Q A T P S D S G T Y
901 CCAGTGTGTGCCAGAAGCCAGAAACCAGAAATCTACATTCATGGCCACTTCTCTCCGTTCTAGTCACAGGGAACCACAGAGATAAGACAGAGACAA
113 Q C C A R S Q K P E I Y I H G H F L S V L V T G N H T E I R Q R Q
1001 AGTGCACACCCTGACTTCAGCCATATCAACGGCACTCTCAGTTCAGGCTTCTGCAAGTAAAGGCTTGGGGGATGTTGGTACCAGCCTGGTGGCCCTTC
147 R S H P D F S H I N G T L S S G F L Q V K A W G M L V T S L V A L

NheI (1158)

1101 AAGCTCTATATACCTTGTGACAGAGAGAAAAAAGACTTCAGAGGAAGGTACAGCGAGGCTAGCTGGCCAGACATGATAAGATAACATTGATGAGTTTGGAA
180 Q A L Y T L •
1201 CAAACCACAAC TAGAATGCAGTGAATAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTA
1301 ACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAATTCTA
1401 AAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATG
1501 TGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGTATTTTCCCAAGGTTTGAAGTACTGCTTTCATTTCTTTATGTTTAAAT
1601 GCACTGACCTCCACATTCCCTTTTATGAAAATATTCAGAAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGA
1701 TGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGC
1801 TTTAGTTCCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAGCAGTCAGGAGCATAGTC
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 D
1901 AGAGTAGGCTCTGACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTCCCTGACAGCCACAATGGTGTCAAAG
108 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 F
2001 TCCTTCTGCCCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCT
74 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 E
2101 CCCAGTCTTGGTCTGATGGCGCCCCGACATGGTGTCTGTGCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTG
41 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 Q
2201 CTGAGAGATGTTGAAGTCTTCAATGATGGCCCTCTATAGTGAAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGG
8 Q S I N F T K M
2301 ATGGCGTCTCCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAG
2401 TTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACC
2501 GCTATCCACGCCATTGATGTAAGTCCGCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCCAAGTAGGAAAGTCCCATAGGT
2601 CATGTAAGTGGCATAATGCCAGGCGGGCCATTTACCGTCAATGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGTACTGCAAGTGGGCA
2701 GTTTACCGTAAATACTCCACCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATTATTGACGTCAATGGGGGGGGTCTG
2801 TGGCGGTGACCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCTGAGGTTAATAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT
2901 AAAAAGGCCCGGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTTGGCGAAACCCGACAGGAC
3001 TATAAAGATAACCAGGCGTTTCCCGTGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGG
3101 AAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGTCTGCTCCAAGCTGGGCTGTGTGACGAACCCCGTTCAGCCC

3201 GACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG
3301 CGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT
3401 TACCTTCGGAAAAGAGTTGGTAGCTTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAA
3501 AAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAA
3601 CATTTAATCAGCGGCCCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAC
3701 AAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA