



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCGCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT

BspHI (560)

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTCATCATGAAGTGGCACATGATCATCTCTGGGCTTATTGTGGT
1 M N W H M I I S G L I V V
601 AGTGCTTAAAGTTGTTGGAATGACCTTATTTCTACTTTATTTCCACAGATTTTTAAACAAAAGTAACGATGGTTTACCACCACCAGGTCCTATGGAACA
13 V L K V V G M T L F L L Y F P Q I F N K S N D G F T T T R S Y G T
701 GTCTCACAGATTTTTGGGAGCAGTTCCTCAAGTCCCAACGGCTTATTACCACAAGGAGCTATGGAACAGTCTGCCCCAAAGACTGGGAATTTTATCAAG
47 V S Q I F G S S S P S P N G F I T T R S Y G T V C P K D W E F Y Q
801 CAAGATGTTTTTCTTATCCACTTCTGAATCATCTTGAATGAAAGCAGGGACTTTTGCAAAGGAAAAGGATCCACATTGGCAATTGTCAACACGCCAGA
80 A R C F F L S T S E S S W N E S R D F C K G K G S T L A I V N T P E
901 GAAACTGAAGTTTCTTCAGGACATAACTGATGCTGAGAAGTATTTTATTGGCTTAATTTACCATCGTGAAGAGAAAAGGTGGCGTTGGATCAACAACCTCT
113 K L K F L Q D I T D A E K Y F I G L I Y H R E E K R W R W I N N S
1001 GTGTTCAATGGCAATGTTACCAATCAGAATCAGAATTTCAACTGTGCGACCTTGGCCTAACAAAGACATTTGATGCTGCATCATGTGACATCAGCTACC
147 V F N G N V T N Q N Q N F N C A T I G L T K T F D A A S C D I S Y

NheI (1129)

1101 GCAGGATCTGTGAGAAGAATGCCAAATGAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAACCACAAGTGAATGCAGTGAAAAAA
180 R R I C E K N A K •
1201 ATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTT
1301 CAGGTTACAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCTCTACAATGGGTATGGAATTCTAAAATACAGCATAGCAAAACTTTAACCTCC
1401 AAATCAAGCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCT
1501 TTCATGGAGTTTAAAGATATAGTGATTTTTCCCAAGGTTTGAAGTAGCTCTTCAATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTATG
1601 AAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAG
1701 TTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTAAATAGAAATTGGACAGCAAGAAAGCAGGCTTCTAGCTTTAGTTCTGGTGTACTTGGGGGGATG
141 • N R T Y K L P I
1801 AGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAGAGATGAGCTCTGCGACATGCCACAGG
131 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 G C P
1901 GGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCC
98 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
2001 AATGGCAATGGCTCAGCACAGCAGTGACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGGTCTGTATGGCCGCCCG
65 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 A A G
2101 ACATGGTCTTGTGCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTTCATGGTGG
31 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 M
2201 CCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCA
2301 CTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTG
2401 ATTTACTAGTCAAACAACCTCCATTGACGTCAATGGGTGGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAA
2501 AACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCACTGACTGGGCATAATGCCAGGCGGGCC
2601 ATTTACCGTCATTGACGTCAATAGGGGCGTACTTGGCATATGATACTTGTACTGCCAAGTGGGCGAGTTTACCATAAATACTCCACCCATTGACG
2701 TCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTTGACGTCAATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTACC
2801 GTAAGTTATGTAACGCCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCGCGTTGCTGGCGTTTTTCCA
2901 TAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGGCGTTTCCCCTGG
3001 AGCTCCCTCGTGCCTCTCTGTTCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCT
3101 GTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAACCCCGTTCCAGCCGACCGCTGCGCTTATCCGGTAACTATCG

3201 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTC
3301 TTGAAGTGGTGGCCTAACTACGGCTACACTAGAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTT
3401 GATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTCGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTT
3501 TTCTACGGGGTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATAT
3601 CTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATA
3701 GGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA