



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

SphI (560)

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGTTCAGCATGCGACGGCTGCTGATCCCTCTGGCCCTGTGGCTGGG  
13 M R R L L I P L A L W L G  
601 CGCGGTGGGCGTGGGCGTCGCCGAGCTACGGAAGCCAGCGCCGGGGCTGCAGGTGGCCCTGGAGGAATTCACAAGCACCCGCGCCGTGCAGTGGGCC  
13 A V G V G V A E L T E A Q R R G L Q V A L E E F H K H P P V Q W A  
701 TTCAGGAGACCAGTGTGGAGAGCGCCGTGGACACGCCCTCCAGCTGGAATATTTGTGAGGCTGGAATTTAAGCTGCAGCAGACAAGTGC CGGAAGA  
47 F Q E T S V E S A V D T P F P A G I F V R L E F K L Q Q T S C R K  
801 GGGACTGGAAGAAACCCGAGTGCAAAGTCAAGCCCAATGGGAGGAAACGAAATGCCTGGCCCTGCATCAAAGTGGCTCTGAGGACAAAGTTCTGGGCCG  
80 R D W K K P E C K V R P N G R K R K C L A C I K L G S E D K V L G R  
901 GTTGGTCCACTGCCCCATAGAGACCCAAGTTCTGCGGGAGGCTGAGGAGCACCAGGAGACCCAGTGCCTCAGGGTGCAGCGGGCTGTTGAGGACCCCCAC  
113 L V H C P I E T Q V L R E A E E H Q E T Q C L R V Q R A G E D P H

NheI (1075)

1001 AGCTTCTACTTCCCTGGACAGTTTCGCTTCTCCAAGGCCCTGCCCGCAGCTAAGCCAGCACTGAGATGCGTGGTCTAGTGGCCAGACATGATAAGAT  
147 S F Y F P G Q F A F S K A L P R S •  
1101 ACATTGATGAGTTTGACAAACCACAACCTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAG  
1201 CTGCAATAAACAAGTTAAACAACAACAATTGCATTCAATTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGTTTTTAAAGCAAGTAAACCTCTACAAA  
1301 TGTGGTATGGAATTTCTAAATACAGCATAGCAAACCTTAACTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATC  
1401 AGGGGCTGTTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTCTTTTCATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTGCTTTCAT  
1501 TTCTTTATGTTTTAAATGCAGTACCTCCACATTCCTTTTTAGTAAATATTCAGAATAATTTAAATACATCATTGCAATGAAAAAATGTTTTTTT  
1601 ATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAA  
1701 GAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAG  
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  
1801 CAGTCAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACTCATCAGAGTAGGGGTGCCTGACAG  
113 C D P A Y 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  
1901 CCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGCCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCCTCAATGTG  
80 V I T D 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  
2001 GACAGCAGAGATGATCTCCCACTTGGTCTGATGGCCGCCCGACATGGTGTCTGTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCC  
47 V A S I 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  
2101 ACCAGCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTCATGGTGGCCCTCTATAGTGTGATTATACTATGCCGATATACTATGCCGATGATTA  
13 V L E L D Q Q S I N F T K M  
2201 TTGTCAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATT  
2301 GCGTCAATGGGGCGGAGTTGTTACGACATTTTGAAAGTCCCGTTGATTTACTAGTCAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGAAAA  
2401 TCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTGCACAAACCGCATCATCATGTAATAGCGATGACTAATACGTAGATGTAAGTGCACAA  
2501 GGAAGTCCCATAAAGTTCATGTAAGTGGCATAATGCCAGGCGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACTTGTAT  
2601 GTAAGTCCCAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATATTGACGT  
2701 CAATGGGCGGGGTCGTTGGCGGTGAGCCAGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAATGTGAGCAAAAGGCCAGC  
2801 AAAAGGCCAGAACCGTAAAAAGGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCAGCAAAATCGACGCTCAAGTCAAGGTTG  
2901 GCGAAACCCGACAGGACTATAAAGATACCAGCGTTTTCCCGTGAAGCTCCCTCGTGCCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCC  
3001 GCCTTTCTCCCTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACG  
3101 AACCCCCGTTACGCCGACCGCTGCGCCTTATCCGGTAACCTATCGTCTTGTGAGTCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG

3201 TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGC  
3301 GCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGCAAGCAGC  
3401 AGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACACGTTAAGGGATTTGGT  
3501 CATGGCTAGTTAATTAACATTTAAATCAGCGGCCGAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACAT  
3601 ACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA