



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGCTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCGGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTATCATGAGCTCTGAAAATTGTTTCGTAGCAGAGAACAGCTC
1 M S S E N C F V A E N S S
601 TTTGCATCCGAGAGTGGAACAAGAAATGATGCCACCAAGTCCCATTTCTCAACACGTCATGAAGGGTCTTCCAAGTTCCTGTCCTGTGTGTAATG
13 L H P E S G Q E N D A T S P H F S T R H E G S F Q V P V L C A V M
701 AATGTGGTCTTCATCACCATTTAATCATAGCTCTCATTGCCTTATCAGTGGGCAATACAATTGTCCAGGCAATACACATTCTCAATGCCATCAGACA
47 N V V F I T I L I A L I A L S V G Q Y N C P G Q Y T F S M P S D
801 GCCATGTTTCTCATGCTCTGAGGACTGGGTTGGTACCAGAGGAAATGCTACTTTATTTCTACTGTGAAGAGGAGCTGGACTTCAGCCAAAATGCTTG
80 S H V S S C S E D W V G Y Q R K C Y F I S T V K R S W T S A Q N A C
901 TTCTGAACATGGTCTACTCTTGCTGTATTGATTCTGAAAAGGACATGAACCTTCTAAAACGATACGAGGTAGAGAGGAACACTGGGTTGGACTGAAA
113 S E H G A T L A V I D S E K D M N F L K R Y A G R E E H W V G L K
1001 AAGGAACCTGGTACCCATGGAAGTGGTCAAATGGCAAAGAATTTAAACAACTGGTTCAACGTTACAGGGTCTGACAAGTGTGTTTTCTGAAAAACACAG
147 K E P G H P W K W S N G K E F N N W F N V T G S D K C V F L K N T

NheI (1173)

1101 AGGTCAGCAGCATGGAATGTGAGAAGAAATTTACTGGATATGTAACAAACCTTACAATAATAAGGAAACATGCTAGCTGGCCAGACATGATAAGATAC
180 E V S S M E C E K N L Y W I C N K P Y K •
1201 ATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAGAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCT
1301 GCAATAAAACAAGTTAACAACAACAATTGCATTCAATTTATGTTTCAGGTTGAGGGGAGGTGTGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATG
1401 TGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAG
1501 GGGCTGTTGCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGAAGTACTGCTTTCATT
1601 CTTTATGTTTTAAATGCAGCTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAATGTTTTTAT
1701 TAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGA
1801 AAGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAGCA
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
1901 GTCAGGAGCATAGTCAGAGATGAGTCTCTGCACATGCCACAGGGGCTGACCACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCC
113 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
2001 ACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGACACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGA
79 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 V
2101 CAGCAGAGATGATCTCCCAAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTGTTGCTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACTCCAC
46 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 V
2201 CAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTATGGTGGCCCTCTATAGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAAT
13 L E L D Q Q S I N F T K M
2301 GTCAAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTG
2401 GTC AATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATC
2501 CCCGTGAGTCAAACCGCTATCCACGCCATTGATGACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGG
2601 AAAGTCCATAAGGTCATGTAAGTGGCATAATGCCAGGCGGGCATTACCCTGATTGACGTCAATAGGGGGGCTACTTGGCATATGATACACTTGATGT
2701 ACTGCCAAGTGGGAGTTTACCCTAAATACTCCACCCATTGACGTCAATGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATATTGACGTCA
2801 ATGGGCGGGGTCGTTGGGCGGTGAGCCAGGCGGGCATTACCCTAAGTTATGTAACGCTGAGGTTAAITAAAGAACATGTGAGCAAAAGGCCAGCAA
2901 AAGGCCAGGAACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGC
3001 GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGC
3101 CTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGGGCTGTGTGCACGAA

3201 CCCCCGTTAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA
3301 ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGC
3401 TCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGTGGTAGCGGTGGTTTTTTTGTTCGCAAGCAGCAG
3501 ATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCA
3601 TGGCTAGTTAATTAACATTTAATCAGCGGCCCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATAC
3701 GCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA