



1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGCGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCAGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGTTGAGTCCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTCCTGACCTGCTTGTCTCAACTCTACGCTTTTGTTCGTTT

SphI (560)
AgeI (552)

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGTCAGCATGCCCTGGAGTCCCTGGCTGCGCTGAGCTGGCGCT
601 GGTGCTGAGGCTGCTGGCTTTGCTGTGGCCCCCTGGACGGGGTGAAGCGTGCAGCTGTGCGCTGCGCACCCCGCAGCAGCTTCTGCCACTCGGCTCTA
13▶ V L R L L A L L W P P G R G E A C S C A P A H P Q Q H F C H S A L
701 GTGATACGGCCAAAATATCCAGTGAGAAGGTAGTCCCTGCCAGCAAAGACCTGTGACTCAAAAAGTATCCGGTATGAAATCAAAACAGATAAAGA
47▶ V I R A K I S S E K V V P A S K D P A D T Q K L I R Y E I K Q I K
801 TGTTAAAGGTTTCGAGAAGCCAAAGGATATTCAGTATGCTACACGCCATTTGACTCTTCCCTGTGGTGTGAAGCTAGAAACCAACAGTCACAAGCA
80▶ M F K G F E K A K D I Q Y V Y T P F D S S L C G V K L E T N S H K Q
901 GTATCTTTGACTGGCCAGATTCTCAGTGATGGAAGTCTTTCATCCATCTGTGCACTACATTGAGCCCTGGGAGGACCTGTCTTGGTGCAGAGGGAG
113▶ Y L L T G Q I L S D G K V F I H L C N Y I E P W E D L S L V Q R E
1001 AGCCTGAATCATCACTACCACAGAACTGTGGCTGCCAAATCACCACCTTGTATGAGTGCATGTACCATCTCAGCCCCAATGAGTGTCTCTGGACGG
147▶ S L N H H Y H Q N C G C Q I T T C Y A V P C T I S A P N E C L W T
1101 ACTGGCTGCTGGAACGGAAGCTCTATGGGTACCAGGCCAGCACTATGTCTGCATGAAGCATGTTGATGGCATCTGCAGCTGGTACCAGGCCACCTACA
180▶ D W L L E R K L Y G Y Q A Q H Y V C M K H V D G I C S W Y R G H L H

NheI (1261)

1201 CCTCCGAAGGAGTACGTTGACATCATCCAGCCCTAGTAGGGACCAGAACCCATCCTTGGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTT
213▶ L R K E Y V D I I Q P •

1301 GGACAAACCACAAGTGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAG
1401 TTAACAACAACAATTGCATTATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATT
1501 CTAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTGGCA
1601 ATGTGCATTAGCTGTTTGACGCCTCACCTTCTTCATGGAGTTAAGATATAGTGTATTTCCCAAGGTTTGAAGTACTCTTCATTTCTTTATGTTTA
1701 AATGCACTGACCTCCACATTCCCTTTTATGATAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCC
1801 AGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTCT

1901 AGCTTTAGTTCCTGGTGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAGCAGTCAGGAGCATA
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
2001 GTCAGAGATGAGCTCTCTGCACATGCCACAGGGGTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCA
109▶ 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 V I T D
2101 AAGCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCAGCAGACAGTACCCTGCCAATGTAGGCCCTCAATGTGGACAGCAGAGATGA
75▶ 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 A S I I
2201 TCTCCCCAGTCTGGTCTGATGGCCGCCCCGACATGGTGCTTGTGCTCTCATAGAGCATGGTATCTTCTCAGTGGCGACCTCCACCAGCTCCAGATC
42▶ 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 L E L D
2301 CTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCG
9▶ Q Q S I N F T K M
2401 TGGATGGCTCTCCAGCTTATCTGACGGTTCATAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTGCGTCAATGGGGC

2501 GAGTTGTTACGACATTTTGGAAAGTCCCCTGTTTACTAGTCAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAA
2601 ACCGCTATCCAGCCATTGATGTAAGTGCACAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCCATAA
2701 GGTCATGTAAGGCATAATGCCAGGCGGCCATTTACCGTATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCCAAGTGG
2801 GCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAGTCCCTATTGGCGTACTATGGGAACATACGTATTATTGACGTCAATGGCGGGGGT
2901 CGTTGGCGGTGACCCAGCGGGCCATTTACCGTAAGTTATGTAACGCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC
3001 CGTAAAAGGCCGCTTGGTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCAGAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG
3101 GACTATAAAGATACCAGCGCTTCCCTGGAAGTCCCTCGTGCCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTC

3201 GGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAG
3301 CCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
3401 GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCC
3501 AGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGGCAAGCAGCAGATTACGCGCAGA
3601 AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAAT
3701 TAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAA
3801 AACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCCAGTGAAGTGCAAGTGCCAGGTGCCAGAACATTTCTCTATCGAA