



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTTCACGCGCCCGCCCTACCTGAGGCC
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

SphI (560)

AgeI (552)

501 TCTGTTCTGCGCAGTTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGGTCAGCATGCTGTTGGAACCCGGCAGAGGCTGCTGTGCCCTGGC
1 M L L E P G R G C C A L A
601 CATCTGTGGCAATTGTGGACATCCAGTCTGGTGGATGCATTAACATCACCAGCTCAGCTTCCAGGAAGGAACGCGACTAACTTAATCTGTACTGTA
13 I L L A I V D I Q S G G C I N I T S S A S Q E G T R L N L I C T V
701 TGGCATAAGAAAGAGAGGCTGAGGGGTTTGTAGTGTTTTGTGCAAGGACAGGTCTGGAGACTGTTCTCTGAGACCAGTTTAAACAGCTGAGACTTA
47 W H K K E E A E G F V V F L C K D R S G D C S P E T S L K Q L R L
801 AAAGGGATCCTGGGATAGATGGTGGTGAATATCATCTCAGTTGATGTTACCATAAGCCAAGTACACCGTTGCACAGTGGGACCTACCAGTGTG
80 K R D P G I D G V G E I S S Q L M F T I S Q V T P L H S G T Y Q C C
901 TGGCAGAAGCCAGAAGTCAAGTATCCGCTTACAGGCCATTTTTCTCCATTCTATTACAGAGACAGGAACTACACAGTACAGGATTGAAACAAAGA
113 A R S Q K S G I R L Q G H F F S I L F T E T G N Y T V T G L K Q R
1001 CAACACCTTGAGTTCAGCCATAATGAAGGCACTCTCAGTTCAGGCTTCTACAAGAAAAGGCTGGGTAATGCTGGTACCAGCCTTGTGCCCTTCAAG
147 Q H L E F S H N E G T L S S G F L Q E K V W V M L V T S L V A L Q

NheI (1126)

1101 CTTTGAAGCCTTGTGCCAAAAGAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGAAAAAATG
180 A L •
1201 CTTTATTTGTAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTAACAACAACATTGCATTATTTTATGTTTCAG
1301 GTTCAGGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAA
1401 TCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTC
1501 ATGGAGTTAAGATATAGTGTATTTTCCAAGGTTTGAAGTACTCTTCAATTTCTTTATGTTTAAATGCACTGACCTCCACATTCCCTTTTAGTAAA
1601 ATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTT
1701 AGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGT
141 • N R T Y K L P I L
1801 TCCTCAATGGTGGTTTTGACCAGCTTGCATTCATCTCAATGAGCACAAAGCAGTCAGGAGCATAGTCAAGATGAGCTCTCTGCACATGCCACAGGGGC
130 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 S
1901 TGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAAT
97 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 I
2001 GGCAATGGCTTACGACAGACAGTACGACCTGCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCCAGTCTTGGTCTGATGGCCGCCCGACA
64 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 V
2101 TGGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCTGCTGAGAGATGTTGAAGTCTTCAATGATGGCCC
30 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 TCCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTTCACTA
2301 AACGAGCTCTGCTTATATAGACCTCCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATT
2401 TACTAGTCAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTACTGCCAAAAC
2501 CGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTACGTAAGTCCCATAAGGTCATGTAAGTGGCATAATGCCAGGCGGGCCATT
2601 TACCGTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGTACTGCAAGTGGGAGTTCACCGTAAATACTCCACCCATTGACGTCA
2701 ATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCAATATTGACGTCAATGGGCGGGGTCGTTGGGCGGTACAGCCAGGCGGGCCATTACCCTA
2801 AGTTATGTAACGCTGACGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCGCTTGTGCGTTTTTCCATAG
2901 GCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTTGGCGAAACCCGACAGGACTATAAAGATACCAGGCTTTCCCCCTGGAAGC
3001 TCCCTCGTGGCTCTCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTA
3101 GGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCT

3201 TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTG
3301 AAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT
3401 CCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTC
3501 TACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTT
3601 TATTTTCATTACATCTGTGTGTTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGC
3701 TGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA