



100

1 GGATCTGCATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCTGCTTGTCTAACTCTACGCTTTTGTTCGTTT

**NcoI (560)**

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGCGCCTACCTGAGATCACCGGTACCATGGAGGTGGTGGACCCGAGCAGCTGGGCATGTTAC  
1 M E V V D P Q Q L G M F T  
601 GGAGGGCAGCTGATGTCGGTGGGTATGGACACGTTTATCCACCGCATCGACTCCACCGAGGTCATCTACCAGCCGCGCCGAAGCGGGCAAAGCTCATC  
13 E G E L M S V G M D T F I H R I D S T E V I Y Q P R R K R A K L I  
701 GGCAAGTACCTGATGGGGACCTGCTGGGGAAAGGCTTTACGGCAAGGTGAAGGAGGTGCTGGACTCGGAGACGCTGTGCAGGAGGGCCGTCAGATCC  
47 G K Y L M G D L L G E G S Y G K V K E V L D S E T L C R R A V K I  
801 TCAAGAAGAAGAGTTGCGAAGGATCCCAACGGGAGGCCAAGCTGAAGAAGGAAATTAACACTACTGAGGAGGTTACGGCACAAAAATGTCATCCAGCT  
80 L K K K K L R R I P N G E A N V K K E I Q L L R R L R H K N V I Q L  
901 GGTGGATGTTATACAACGAAGAGAAGCAGAAAATGTATATGGTGTGGACTGCTGTGGCATGCAGGAAATGCTGGACAGCTGCCGGAGAAG  
113 V D V L Y N E E K Q K M Y M V M E Y C V C G M Q E M L D S V P E K  
1001 CGTTTCCAGTGTGCCAGGCCACGGTACTTCTGTACAGTGAATGACGGCCTGGAGTACCTGCATAGCCAGGGCATTGTGCAACAAGGACATCAAGCCGG  
147 R F P V C Q A H G Y F C Q L I D G L E Y L H S Q G I V H K D I K P  
1101 GGAACCTGCTGCTCACCACCGGTGGCACCTCAAATCTCCGACTGGGCGTGGCCGAGGACTGCACCCGTTCCGGCGGACGACACCTGCCGGACCAG  
180 G N L L L T T G G T L K I S D L G V A E A L H P F A A D D T C R T S  
1201 CCAGGGTCCCGGCTTTCCAGCCGCCGAGATTGCCAACGGCCTGGACACCTTCTCCGGCTCAAGTGGACATCTGGTCCGCTGGGGTACCCTCTAC  
213 Q G S P A F Q P P E I A N G L D T F S G F K V D I W S A G V T L Y  
1301 AACATCACCGGTCTGTACCCCTTGAAGGGGACAACATCTACAAGTTGTTGAGAATCGGGAAGGGAGCTACGCCATCCCGGGCAGCTGTGGCC  
247 N I T T G L Y P F E G D N I Y K L F E N I G K G S Y A I P G D C G  
1401 CCCCCTCTGTACCTGCTGAAAGGGATGCTTGAGTACGAACCGCCAAGAGGTTCTCCATCCGGCAGATCCGGCAGCACAGCTGGTCCGGAAGAAACA  
280 P P L S D L L K G M L E Y E P A K R F S I R Q I R Q H S W F R K K H  
1501 TCCTCCGGTGAAGCACCAGTCCCATCCACCGAGCCAGACACCAAGGACCGGTGGCGCAGCATGACTGTGGTCCGCTACTTGGAGGACCTGCACGGC  
313 P P A E A P V P I P P S P D T K D R W R S M T V V P Y L E D L H G  
1601 GCGGACGAGGACGAGGACCTTCTGCATCGAGGATGACATCATCTACACTCAGGACTTACGGTCCCGGACAGGTCCAGAAAGAGGAGGCCAGTACA  
347 A D E D E D L F D I E D D I I Y T Q D F T V P G Q V P E E E A S H  
1701 ATGACAGCGCCGGGCTCCCAAGGCCGTGTATGAACGGCACAGAGCGCGCAGCTGAGCACAAATCCAGGCGGAGGGCCCGGCCCAACCC  
380 N G Q R R G L P K A V C M N G T E A A Q L S T K S R A E G R A P N P

**NheI (1872)**

1801 TGCCCGCAAGGCTGCTCCGCCAGCAGCAAGATCCGCCGGTGTGGCCTGCAAGCAGCAGTGAAGGCTGGCCCTAGCTGGCCAGACATGATAAGATACA  
413 A R K A C S A S S K I R R L S A C K Q Q •  
1901 TTGATGAGTTTGACAAACCACAACCTAGAATGCAGTAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTG  
2001 CAATAAACAAGTTAACAACAACAATTGCATTATTTTATGTTTCAGGTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGT  
2101 GGTATGGAATTCTAAATACAGCATAGCAAACCTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGG  
2201 GGCTGTTGCCAATGTGCATTAGCTGTTGCAGCCTCACCTTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTTGAAGTACTGCTTTCATTT  
2301 TTTATGTTTTAAATGCACTGACCTCCACATCCCTTTTTAGTAAATATTAGAAATAATTTAAATACATCATTGCAATGAAAAAATGTTTTTTATT  
2401 AGGCAGAATCCAGATGCTCAAGGCCCTCATAATATCCCCAGTTTAGTAGTTGACTTAGGGAACAAGGAACCTTAAATAGAAATTGGACAGCAAGAA  
2501 AGCGAGCTTCTAGCTTTAGTTCCTGGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCATCTCAATGAGCACAAAGCAG  
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  
2601 TCAGGACATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCA  
112 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 V  
2701 CAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTCCAGCAGACAGTACCCCTGCAATGTAGGCCTCAATGTGGAC  
79 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  
2801 AGCAGAGATGATCTCCAGTCTTGGTCTGATGGCCGCCGACATGGTGTCTTGTGCTCATAGAGCATGGTGATCTTCTCAGTGGCGACCTCCACC  
46 A S I I E G T K T R I A A G V H K N D E Y L M T I K E T A V E V  
2901 AGTCCAGATCCTGCTGAGAGATGTTGAAGGCTTTCATGGTGGCCCTCTATAGTGAGTCGTATTATACTATGCCGATATACTATGCCGATGATTAATTG  
12 L E L D Q Q S I N F T K M  
3001 TCAAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTGGC  
3101 TCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAACCTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCC

3201 CCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGGA  
3301 AAGTCCCATAAGGTCATGTAAGTCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGGA  
3401 CTGCCAAGTGGGAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTATTATTGACGTCAA  
3501 TGGGCGGGGTGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCTGACGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAA  
3601 AGGCCAGGAACCGTAAAAAGCCGCTTGTGGCGTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCG  
3701 AAACCCGACAGGACTATAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCTGCCGTTACCGGATACCTGTCCGCC  
3801 TTTCTCCCTTCGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGATGATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCAGAAC  
3901 CCCCCGTTAGCCGACCGCTGCGCTTATCCGTAATATCGTCTTGTAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA  
4001 CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGGCT  
4101 CTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTGTTGCAAGCAGCAGA  
4201 TTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGAACGAAAACACTCACGTTAAGGATTTTGGTCAT  
4301 GGCTAGTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTTCATTACATCTGTGTGGTTTTTTGTGTGAATCGTAACTAACATACG  
4401 CTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAAGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA