



1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATGCCACAGTCCCGGAGAAGTTGGGGGAGGGTTCGGCAATTGAACGGGTGCCTA  
101 GAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCC  
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGCC  
301 GCCATCCACGCCGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCTCCTGAAGTGCCTCCGCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACC  
401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCTGCTTGTCTCAACTCTACGCTTTGTTTCGTTT

**NcoI (560)**  
**AgeI (552)**

501 TCTGTTCTGCGCGTTACAGATCCAAGCTGTGACCGCGCTACCTGAGATCACCGGTCCAGTGGTGACAAGATCCTGAGGGCAAAGGGAAGCAATT  
1 M A D K I L R A K R K Q F  
601 TATCAACTCAGTGAGTATAGGGACAATAAATGGATTGTTGGATGAACCTTTAGAGAAGAGAGTGTGAATCAGGAAGAAATGGATAAAAATAAACTTGA  
13 I N S V S I G T I N G L L D E L L E K R V L N Q E E M D K I K L A  
701 AACATTACTGCTATGGACAAGGCACGGGACCTATGTGATCATGTCTCTAAAAAGGGCCCCAGGCAAGCCAAATCTTTATCACTTACATTTGTAATGAAG  
47 N I T A M D K A R D L C D H V S K K G P Q A S Q I F I T Y I C N E  
801 ACTGCTACCTGGCAGGAATTCTGGAGCTTCAATCAGCTCCATCAGCTGAAACATTTGTTGCTACAGAAGATTCTAAAGGAGGACATCCTTCATCCTCAGA  
80 D C Y L A G I L E L Q S A P S A E T F V A T E D S K G G H P S S S E  
901 AACAAAGGAAGAACAAGAAGATGGCACATTTCCAGGACTGCTGGGACCTCAAGTTTTGCCCTTTAGAAAAAGCCAGAAGTTATGGAAAGAA  
113 T K E E Q N K E D G T F P G L T G T L K F C P L E K A Q K L W K E  
1001 AATCCTTCAGAGATTTATCCAATAATGAATACAACCACTCGTACACGTCTTGCCTCATTATCTGCAACACAGAGTTTCAACATCTTTCTCCGAGGGTTG  
147 N P S E I Y P I M N T T T R T R L A L I I C N T E F Q H L S P R V  
1101 GAGCTCAAGTTGACCTCAGAGAAATGAAGTTGCTGCTGGAGGATCTGGGTATACCGTGAAGTGAAGAAAAATCTCACAGCTCTGGAGATGGTGAAGA  
180 G A Q V D L R E M K L L L E D L G Y T V K V K E N L T A L E M V K E  
1201 GGTGAAAGAAATTTGCTGCCTGCCAGAGCACAAGACTTCTGACAGTACTTTCTTGTATTGATGTCTCATGGTATCCAGGAGGGAATATGTGGGACCACA  
213 V K E F A A C P E H K T S D S T F L V F M S H G I Q E G I C G T T  
1301 TACTCTAATGAAGTTTCAGATATTTAAAGTTGACACAATCTTTCAGATGATGAACACTTTGAAGTGCCCAAGCTTGAAGACAAGCCCAAGGTGATCA  
247 Y S N E V S D I L K V D T I F Q M M N T L K C P S L K D K P K V I  
1401 TTATTCAGGCATGCCGTGGAGAGAAACAAGGAGTGGTGTGTTAAAAGATTGAGTAAGAGACTCTGAAGAGGATTTCTTAACGGATGCAATTTTTGAAGA  
280 I I Q A C R G E K Q G V V L L K D S V R D S E E D F L T D A I F E D  
1501 TGATGGCATTAAAGAGGCCCATATAGAGAAAGATTTTATTGCTTTCTGCTCTTCAACACCAGATAATGTGCTTGGAGACATCCTGTGAGGGGCTCACTT  
313 D G I K K A H I E K D F I A F C S S T P D N V S W R H P V R G S L  
1601 TTCATTGAGTCACTCATCAAACATGAAGAATATGCCTGGTCTTGTGACTTGGAGGACATTTTCAGAAAGTTTCGATTTTCATTGAAACAACAGAAAT  
347 F I E S L I K H M K E Y A W S C D L E D I F R K V R F S F E Q P E

**NheI (1771)**

1701 TTAGGCTACAGATGCCACTGCTGATAGGGTGACCCTGACAAAACGTTTCTACCTCTTCCCGGACATTAAGCTAGCTGGCCAGACATGATAAGATACAT  
380 F R L Q M P T A D R V T L T K R F Y L F P G H •  
1801 TGATGAGTTTGGACAAACCAACTAGAAATGCAAGTGAAGAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGC  
1901 AATAACAAGTTAAACAACAATTGCATTATTTTATGTTTCAGGTTTCCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTG  
2001 GTATGGAATCTAAAATACAGCATAGCAAACTTAACTCCAAATCAAGCCTCTACTTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGG  
2101 GCTGTTGCCAATGTGCATTAGCTGTTTGACGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGATTTTCCCAAGGTTTGAAGTACTTCTCATTCT  
2201 TTATGTTTTAAATGCACTGACCTCCACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATA  
2301 GGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATGGACAGCAAGAAA  
2401 GCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTATCTCAATGAGCACAAAGCAGT  
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  
2501 CAGGAGCATAGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCAC  
112 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  
2601 AATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCAATGGCAATGGCTTACAGCACAGACAGTACCTGCCAATGTAGGCCTCAATGTGGACA  
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  
2701 GCAGAGATGATCTCCCGAGTCTTGGTCTGTATGGCCGCCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCA  
45 A S I E 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  
2801 GCTCCAGATCCTGCTGAGAGATGTTGAAGTCTTTCATGGTGGCCCTCTATAGTGAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGT  
12 E L D Q Q S I N F T K M  
2901 CAAACAGCGTGGATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCTACCGCCATTTGCGT  
3001 CAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAACCTCCATTGACGTCAATGGGGTGGAGACTTGGAAATCCC  
3101 CGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTCTGCAAAACCGCATCATGTAATAGCGATGACTAATACGTAGATGTAAGTCAAGTAGGAA

3201 AGTCCATAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTAC  
3301 TGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAAT  
3401 GGGCGGGGTCGTTGGGCGTCAAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAA  
3501 GGCCAGGAACCGTAAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA  
3601 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCT  
3701 TTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACC  
3801 CCCCGTTCAGCCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC  
3901 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTC  
4001 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCCTGGTAGCGGTGGTTTTTTTGGTTGCAAGCAGCAGAT  
4101 TACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGAAAACTCACGTTAAGGGATTTTGGTCATG  
4201 GCTAGTTAATTAACATTTAATCAGCGGCCCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGTAACTAACATACGC  
4301 TCTCCATCAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCCAGTGC AAGTGCAGGTGCCAGAACATTTCTCTATCGAA